EMBARRASSINGLY PRELIMINARY

COMMENTS WELCOME

PLEASE DO NOT CITE

PREPARED FOR HANDBOOK OF LABOR ECONOMICS, VOLUME 4

Imperfect Competition in the Labor Market

Alan Manning

Centre for Economic Performance
London School of Economics
Houghton Street
London WC2A 2AE

October 2009

Abstract

Introduction

In recent years, it has been increasingly recognized that many aspects of labour markets are best analyzed from the perspective that there is some degree of imperfect competition. At its most general, 'imperfect competition' should be taken to mean that employer or worker or both get some rents from an existing employment relationship. If an employer gets rents, then this means that the employer will be worse off if a worker leaves i.e. the marginal product is above the wage and worker replacement is costly. If a worker gets rents then this means that the loss of the current job makes the worker worse off – an identical job cannot be found at zero cost. If labour markets are perfectly competitive then an employer can find any number of equally productive workers at the prevailing market wage so that a worker who left could be costlessly replaced by an identical worker paid the same wage. And a worker who lost their job could immediately find another identical employer paying the same wage so would not suffer losses.

A good reason for thinking that there are rents in the employment relationship is that people think jobs are 'a big deal'. For example, when asked open-ended questions about the most important events in their life over the past year, employment-related events (got job, lost job, got promoted) come second after 'family' events (births, marriages, divorces and death) - see Table 1 for some British evidence on this. This evidence resonates with personal experience and with more formal evidence – for example, the studies of Jacobson, Lalonde and Sullivan (1983) and von Wachter, Song and Manchester (2009) all suggest substantial costs of job loss. And classic studies like Oi (1962) suggest non-trivial costs of worker replacement.

This chapter reviews recent developments in thinking about imperfect competition in labour markets. The plan is as follows. The next section outlines the main sources of rents in the employment relationship. The second section discusses estimates of the size of rents in the employment relationship. The third section then consider models of how the rents in the employment relationship are split between worker and employer, the question of wage determination and the fourth section considers evidence on this point. I argue that this all adds up to a persuasive view that imperfect competition is pervasive in labour markets. But, up to this point, we have not considered the 'so what' question – what is the value-added of this perspective? – that is

the subject of the fifth section. The sixth section then reviews a number of classic topics in labour economics – the effect of regulation, the gender pay gap – and other areas – economic geography, macroeconomics – where the perspective of imperfect competition can be shown to make a difference.

Much work in this area is phrased in terms of canonical models – one might mention the search and matching models of Pissarides (1990) or Mortensen and Pissarides (1994) or the wage-posting model of Burdett and Mortensen (1998). New developments are thought of as departures from these canonical models. This paper does not do that, partly for the reason that excellent surveys can be found elsewhere (e.g. Rogerson, Shimer and Wright, 2005, or Mortensen and Pissarides, 1999, on search models) but also because of a belief that this model-based approach to the topic is not always helpful. Although the use of very particular models encourages precise thinking, that precision relates to the models and not the world and can easily become spurious precision when the models are very abstract with assumptions designed more for analytical tractability than realism. So, this survey is based on the belief that it can be useful to think in very general terms about general principles and that one can say useful things without having to couch them in a complete but necessarily very particular model.

1. The Sources of Imperfect Competition

As will be discussed below there are different ways in which economists have sought to explain why there are rents in the employment relationship. This section will argue they are best understood as having a common theme – that, from the worker perspective, it takes time and/or money to find another employer who is a perfect substitute for the current one and that, from an employer perspective, it is costly to find another worker who is a perfect substitute for the current one. And, that, taken individually, these explanations of the sources of rents often do not seem particularly plausible but, taken together, they add up to a convincing description of the labour market.

1.1 *Frictions and idiosyncracies*

To illustrate this, consider search models (for relatively recent reviews see Mortensen and Pissarides, 1999; and Rogerson, Shimer and Wright, 2005). In these

models it is assumed that it takes time for employers to be matched with workers – in some versions, the time can be influenced by expenditure. These models have become the workhorse model in much of macroeconomics (see Rogerson and Shimer, xxxx) because one cannot otherwise explain the dynamics of unemployment. But, taken literally, this model is not very plausible. It is not hard to find an employer – I can probably see 10 from my office window. But, what is hard is to find an employer who is currently recruiting[1] who is the same as my current one i.e. a perfect substitute for my current job. This is because there is a considerable idiosyncratic component to employers across a vast multitude of dimensions that workers care about. This idiosyncratic component might come from non-monetary aspects of the job (e.g. one employer has a nice boss, another a nasty one, one has convenient hours, another does not) or from differences in commuting distances or from many other sources. A good analogy is our view of the heavens: the stars appear close together but this is an illusion caused by projecting three dimensions onto two. Neglecting the multitude of dimensions along which employers differ that matter to workers will seriously overestimate our impression of the extent to which jobs are perfect substitutes for each other from the perspective of workers.

One other commonly given explanation for why there may be rents in the employment relationship is 'specific human capital'. Although this is normally thought of as distinct from the reasons given above, it is better thought of as another way in which employers may not be perfect substitutes for each other – in this case in terms of the quality of the match or the marginal product of the worker. This comes out clearly in the discussion of specific human capital provided by Lazear (2003). He struggles with the problem of what exactly are specific skills coming up with the answer that "it is difficult to generate convincing examples where the firm-specific component [of productivity] approaches the general component" question. He goes on to argue that all skills are general skills but that different employers vary in how important those skills are in their particular situation. So, a worker with a particular package of general skills will not be

1 It is an interesting question why not all employers are recruiting all the time if the typical employment relationship has rents. Manning (2003, chapter 10) offers an answer to this apparent conundrum – it is costly to create jobs and employers do not create jobs they do not expect to be able to fill. Vacancies, in this view, are best seen as 'accidents'.

faced with a large number of employers requiring exactly that package. As Lazear (2003, p2) makes clear this relies on employers being thin on the ground otherwise a large supply of employers demanding exactly your mix of skills would be available and the market would be perfectly competitive. Again, it is the lack of availability of employers who are perfect substitutes that can be thought of as the source of the rents2.

1.2 *Institutions and Collusion*

So far, the discussion has concentrated on rents that are inevitable. But rents may also arise from man-made institutions that artificially restrict competition. This implicit or explicit collusion may be by workers or employers. Traditionally it is collusion by workers in the form of trade unions that has received the most attention. However, this chapter does not discuss the role of unions at all because it is covered in another chapter (Farber, xxxx).

Employer collusion has received much less attention. This is in spite of the fact that Adam Smith had this to say about the practice of economists to see bargaining power of workers everywhere: "we rarely hear, it has been said, of the combinations of masters; though frequently of those of workmen. But whoever imagines, upon this account, that masters rarely combine, is as ignorant of the world as of the subject". Employer collusion where it exists is thought to be very specific labour markets e.g. US professional sports (see Boal and Ransom, 1997, for adiscussion). More, recently Priest (2010) has argued that the 'problems' of the labour markets for medical interns and legal clerks (which have led to the use of matching algorithms proposed by Roth, 1990) are in fact the consequences of employer collusion on wages in a labour market with very heterogeneous labour. But although it is clear that employers do not en masse collude to set wages there may be more subtle but nevertheless effective ways to do it. For example, Naidu (2010) explores the effect of legislation in the post-Bellum South that punished (almost exclusively white) employers if they enticed (almost exclusively black) workers away from other employers. Although it might appear at first sight to be white employers who suffer from this legislation, Naidu (2010) presents evidence that, by

2 Of course, specific human capital may change over time so that it is entirely plausible that rents rise with job tenure, a critical idea in Becker's thinking (Becker 1963).

reducing competition for workers, it was blacks who were made worse off by this. The legislation can be thought of as a way for employers to commit not to compete for workers, leading to a more collusive labour market outcome.

But, it may also be the case that economists have traditionally under-estimated the ability of employers to collude. As the physical location of employers is important to workers, it is likely that, for many workers, the employers who are closest substitutes from the perspective of workers are also geographically close making communication and interaction between them easy. Manning (2009) gives an example of a model in which employers are on a circle (as in Bhaskar and To, 1999) and collude only with the two neighbouring employers in setting wages. Although there is no collusion spread over the whole market, Manning (2009) shows that a little bit of collusion can go a long way leading to labour market outcomes a long way from perfect competition. One way of putting the question is 'Do managers of neighbouring fast food restaurants talk to each other or think about how the other might react if wages were to change?'. I think the honest answer is that we just don't know

2. How much imperfect competition? The Size of Rents

A natural question to ask is how important is imperfect competition in the labour market? As explained in the introduction, this is really about the size of rents earned by employer and worker from an on-going employment relationship. The experiment one would like to run is to randomly and forcibly terminate employment relationships and examine how the pay-offs of employer and worker change. We do not have that experiment and, if we did, it would not be that easy to measure the pay-offs which would not just be in the current period but in the future.

Nonetheless we can make some attempt to measure the size of rents and this section illustrates the way in which we might attempt to do that. We consider a number of approaches. First we seek to exploit the idea that the larger the size of rents, the more expenditure on rent-seeking activity we would expect to see – we use this idea from both worker and employer perspectives. Before we review these estimates, one should be aware that there is almost certainly huge variation in the extent of rents in the labour market so that one has to bear in mind that the estimates that follow are not from random

samples and should not automatically be regarded as representative of the labour market as a whole. In the absence of a large number of estimates we do not have a clear idea how to resolve this problem.

2.1 *The Costs of Recruitment*

First, consider how we might attempt to measure rents from the perspective of employers. If an employer and worker are forcibly separated then a good estimate of the size of the rents is the cost of replacing the worker with an identical one – what we might call the marginal hiring cost. This marginal hiring cost is quite a general principle but lets see it worked out in a specific model, the Pissarides (1990) matching model. Denote by J the value of a filled job and J_v the value of a vacant job – the size of the rents accruing to an employer can be measured by $(J - J_v)$. The value function of a vacant job must be given by:

$$rJ_v = -c + \theta(J - J_v) \tag{1}$$

Where r is the interest rate, c is the per-period cost of a vacancy and θ is the rate at which vacancies are filled. As firms can freely create vacant jobs (it is a filled vacancy that can't be costlessly created we will have $J_v = 0$ in equilibrium in which case (1) can be re-arranged to give us:

$$(J - J_v) = \frac{c}{\theta} \tag{2}$$

Which can be interpreted as the per period vacancy cost times the expected duration of a vacancy. This can be interpreted as the marginal cost of a hire. This latter principle can be thought of as a general one.

The specific model outlined here suggests a very particular way of measuring the rents accruing to employers – measure the cost of advertising a job and the expected duration of a vacancy. Both of these numbers are probably small, at least for most jobs. However, the way in which costs of vacancies are modeled here is not the best. Actual studies of the costs of filling vacancies find that the bulk of the costs are not in generating applicants as this model suggests but in selecting workers from applicants and training those workers to be able to do the job.

Even once one has got an estimate of hiring costs, h, one needs to scale it in some way to get an idea of how important they are. The natural way to do that would is to relate it to the wage, w. However, salary is a recurrent cost whereas the hiring cost is a one-off cost. How large are hiring costs depend in part on how long the worker will be with the firm. Given this it is natural to multiply the hiring costs by the interest rate plus the separation rate i.e. to use the measure (r+s)h/w. Because separation rates are often about 20% and much bigger than real interest rates, this is effectively equal to multiplying the hiring costs by the separation rate, (s*h/w) which can also be thought of as dividing the hiring cost by the expected tenure of the worker, to give the hiring cost spread over each period the firm expects to have the worker. Another way of looking at the same thing is the share of wage payments over the whole job tenure that is spent on recruiting and training them. In a steady-state this will be equal to the ratio of total hiring costs to the current wage bill as the total hires must be equal to sN with total hiring costs sNh, compared to total wage bill wN.

It is hard to get direct data on hiring costs and the estimates we do have are for very different times and places and from very different data sets. They are summarized in Table 2. Not all of the estimates measure all aspects of hiring costs and not all the studies contain enough information to enable one to compute the ideal measure described above. For example, the French studies of Abowd and Kramarz (2003) and Kramarz and Michaud (2009) exclude the amount of time spent by workers in the firm on the recruitment process. Given these problems what is perhaps remarkable is that the original Oi (1962) estimates seem in the right ballpark – with hiring costs at about 5% of the total.

One distinction that will be shown to be of critical importance later is between average and marginal hiring costs. Suppose that the total cost of R recruits is given by:

$$C = h_0 R^{\beta} \tag{3}$$

Then there is the following relationship between marginal hiring cost and the average hiring cost:

$$marginal\ hiring\ cost = \beta * average\ hiring\ cost \tag{4}$$

If β is above 1 so there are increasing marginal costs of recruitment, the marginal cost will be above the average cost. Some of the estimates referred to in Table 2 are clearly average costs though others are not clear.

We do have some little bits of evidence on marginal hiring costs. Manning (2006) and Blatter, Muhlemann and Schenker (2009) both report increasing marginal costs. However Abowd and Kramarz (2003) and Kramarz and Michaud (2009) report decreasing marginal costs as they estimate hiring to have a fixed cost component. However this last result may be because they exclude the costs of recruitment where one would expect marginal costs to be highest. The finding in Barron and Bishop (1985) that large firms have higher hiring costs might also be interpreted as evidence of increasing marginal costs as large firms can only get that way by lots of hiring.

2.2 *The Search Activity of the Unemployed*

Now consider things from the perspective of workers. One cannot use a similar methodology because, while it is reasonable to assume that vacant jobs are in potentially infinite supply, one cannot make the same assumption about unemployed workers. But, lets consider the value function for an unemployed worker who faces a wage offer distribution and can influence the arrival rate of job offers through the expenditure of time and money on job search as well as the choice of a reservation wage. The value of being unemployed, V^u, can then be written as:

$$rV^u = \max_{(r,\lambda_u)} b + \lambda_u \int_r \left[V(w) - V^u \right] dF(w) - c(\lambda_u) \tag{5}$$

Where r is the reservation wage and F(w) is the wage offer distribution. Lets take the first order condition for the job offer arrival rate:

$$c'(\lambda_u) = \int_r \left[V(w) - V^u \right] dF(w) \tag{6}$$

This shows us that the incentive for workers to generate wage offers is related to the rents they will get from those offers. Let us re-arrange (6) to give us:

$$\frac{\int_r \left[V(w) - V^u \right] dF(w)}{1 - F(r)} = \frac{c'(\lambda_u)}{1 - F(r)} = \frac{1}{\lambda_u \left(1 - F(r)\right)} \lambda_u c'(\lambda_u) = d_u c(\lambda_u) \frac{\partial \log c_u}{\partial log \lambda_u} \tag{7}$$

Where d_u is the expected duration of unemployment. This is observable and we can also get estimates of the total expenditure on time and effort of the unemployed on getting work $c(\lambda_u)$. The first two terms are the total expected costs of getting a job for the unemployed. The final term can be interpreted as the inverse of the elasticity of the unemployment duration with respect to search effort. This is perhaps harder to estimate. The intuition for this is simple – if workers typically get rents from jobs we would expect to see them willing to expend considerable amounts of time and money to get a job.

Again, we would like to normalize these costs. If most of the costs of job search are time costs then it is natural to normalize the amount of time spent looking for a job as a fraction of the amount of time one would spend in the job when one gets it so that the measure of the importance of rents becomes:

$$\frac{d_u c(\lambda_u)}{d_e h_e} \frac{\partial \log c_u}{\partial log \lambda_u} \tag{8}$$

What does the evidence suggest about the amount of time and money spent by the unemployed in trying to gain employment.

Probably the most striking fact about the job search activity of the unemployed is often how small is the amount of time they seem to spend on it. A good example would be the cross-country comparison of Krueger and Mueller (2008) who use time-use surveys to conclude that the average unemployed person spends approximately 4 minutes a day on job search in the Nordic countries, 10 minutes in the rest of Europe, and 30 minutes in North America. There is considerable variation in estimates e.g. Holzer (1988) finds an average of 128 minutes for unemployed American youth [need to expand this]. Taking these numbers at face value and valuing them at the average wage and using the formula in (8) one would be tempted to conclude that the rents offered by employment could not be large.

However, there are a number of reasons to be cautious about this conclusion. First, (8) suggest that the amount of search effort will be also be influenced by how effective it is at the margin. If increases in search time lead to little improvement in job offer arrival rates, a small amount of job search is consistent with large rents. Ideally we would like to have some experimental evidence on what happens when we force individuals to increase job search activity. Although there are a large number of studies,

many experimental or quasi-experimental, that seek to estimate the affect of programmes designed to assist with job search on various outcomes for the unemployment, many of these job search assistance programs combine more checking on the job search activity of the unemployed with help to make search more effective. For current purpose we would like only the former. One study that seems to come close is Klepinger, Johnson and Joesch (2002) which investigate the effect of Maryland doubling the number of required employer contacts from 2 to 4. This doubling of required contacts significantly reduced the number of weeks of UI receipt by 0.7 weeks on a base of 11.9 so a doubling in the required number of contacts reduces unemployment durations by 6%. Assuming that the doubling of the number of contacts doubles the cost leads to a very small implied elasticity of 0.04. There are a number of reasons to be cautious – we do not have evidence about how much employer contacts were actually increase and second, when individuals are forced to comply with increased employer contacts they would not choose for themselves, they will probably choose low-cost but ineffective contacts. These would tend to lead to lower estimates of the elasticity. On the other hand exits from UI are not the same as exits to employment and the employment outcomes are not so favourable though with a great deal of imprecision. It could be that stock-flow matching (Coles and Smith, 1998; Shimer, 2006) provides a plausible reason for why, at the margin, there is little return to extra job search.

There are also a number of non-experimental studies that seek to relate unemployment durations to job search intensity with mixed results that suggest caution in interpretation. For example, Holzer (1987) reports estimates for the effect of time spent on a variety of search methods on the probability of gaining new employment (though he also controls for the number of search methods used) – many of the estimated effects are insignificant or even 'wrongly-signed'.

Secondly, the formula in (8) assumes that the cost of time in job search and employment can be equated. However, the time cost of job search may be higher than one might think as Krueger and Mueller (2008) find that levels of sadness and stress are high for the unemployed while looking for a job and levels of happiness are low. If these emotional costs are high, the cost of job search will be higher than one otherwise have thought.

Thirdly, job search uses money as well as time and, while the unemployed have a lot of time on their hands, they are short of money. Studies like Card, Chetty and Weber (200x) suggest that the unemployed are unable to smooth consumption across periods of employment and unemployment so that the marginal utility of income for the unemployed may be much higher than for the employed. For example, in the UK evaluation of the Job Seeekers' Allowance one-third of UI recipients reported that their job search was limited because of the costs involved with the specific costs most commonly mentioned being travel, stationery, postage and phone. If time and money are complements in the job search production function, low expenditure will tend to be related to low time spent.

Finally, DellaVigna and Paserman (2005) investigate the effect of hyperbolic discounting in a job search model. They present evidence that, in line with theoretical predictions, the impatient engage in lower levels of job search and have longer unemployment durations. If this is the right model of behavior one would have to up-rate the costs of job search by the degree of impatience to get an estimate of the size of rents from jobs.

So, the bottom line is that the unemployment do not seem to expand huge amounts of effort into trying to get employment so that one might conclude that the rents are not large. However, we have discussed reasons why such a conclusion might be hasty. And we do have other evidence that the unemployed are worse off than the employed in terms of well-being – see, for example, Clark and Oswald (1994), Krueger and Mueller (2008). I would be hesitant to conclude that the rents from employment are small for the unemployed because of the low levels of search activity as I suspect that if one told a room of the unemployed that their apathy showed they did not care about having a job, one would get a fairly rough reception. When asked to explain low levels of search activity, one would be much more likely to hear the answer 'there is no point' i.e. they say that the marginal return to more search effort is low.

This method gives us an estimate of the average rents accruing to an unemployed worker in the first job they get after a spell of unemployment. One incredibly important factor that the discussion so far has ignored is heterogeneity. One way to convince oneself of the importance of this is to note that there are large numbers of people without

a job who are not looking for one. For this group – that labour market statistics would normally classify as the inactive – the expected rents from the employment relationship are too small to justify job search. Once one recognizes the existence of heterogeneity one needs to worry about the population whose rents one is trying to measure. The methodology here might be useful to tell us about the rents for the unemployed but we would probably expect that the average rents for the unemployed are lower than for the employed.

2.3 *The Costs of Job Loss*

So, we would like to have a measure of rents for the employed. The experiment one would like to run is to consider what happens when workers are randomly separated from jobs. There is a literature that considers exactly that question – studies of displaced workers (Jacobson, Lalonde and Sullivan, 2003; von wachter, Manchester and Song, 2009). The more recent studies perhaps suggest a loss in the PDV of earnings of about 17.5% from job loss [check and confirm]. One concern is the difficulty of finding good control groups e.g. the reason for displacement is presumably employer surplus falling to less than zero. But , for some not totally explained reason, it seems that wages prior to displacement are not very different for treatment and control groups – it is only post-displacement that one sees the big differences. Under this assumption one can equate these estimates to loss of worker surplus.

2.4 *Conclusions*

The methods discussed in this section can be used to give us ballpark estimates of the extent of imperfect competition in labour markets. Because we have discussed estimates of the rents accruing to employers and workers, one might also think about using these estimates to give us some idea of how the rents are split between worker and employer. However, because none of the estimates come from the same employment relationship, that would be an unwise thing to do. The next section discusses models of the balance of power between employers and workers and these are reviewed in the next section.

3. Models of Wage Determination

When there are rents in the employment relationship, one has to model how these rents are split between worker and employer i.e. one needs a model of wage determination. This is a very old problem in economics in general and labour economics in particular, going back to the discussion of Edgeworth (1932) where he argued that the terms of exchange in bilateral monopoly were indeterminate. That problem has never been definitively resolved, and we will argue that is probably because it cannot be. In this section we describe the two main approaches found in the literature and compare and contrast them.

2.1 *Ex Post Wage Bargaining and ex Ante Wage-Posting*

Here we briefly review the two main approaches that have been taken in recent years – what we might call ex post wage bargaining and ex ante wage-posting – though we briefly discuss others at the end of the section. In ex post wage bargaining the wage is split after the worker and employer have been matched, according to some sharing rule, most commonly an asymmetric Nash bargain. In ex ante wage-posting the wage is set unilaterally by the employer before the worker and employer meet.

These two traditions have been used in very different ways. The bargaining models are the preferred models in macroeconomic applications (see Rogerson and Shimer, 2010) while microeconomic applications tend to use wage posting. But, what is often not very clear to students entering this area is why these differences in tradition have emerged and what are the consequences. Are these differences based on good reasons, bad reasons or no reasons at all? Here we try to provide an overview which, while simplistic, captures the most important differences.

Although the models used are almost always dynamic, the ideas can be captured in a very simple static model and that is what we do here. The simple static model derives from Hall and Lazear (1984) who discuss a wider set of wage-setting mechanisms than we do here. Assume that there are firms, which differ in their marginal productivity of labour, p. A firm is assumed only to be able to employ one worker.

In ex post wage bargaining models, the wage in a match between a worker with leisure value b and a firm with productivity p is split to maximize an asymmetric Nash bargain:

$$(p-w)^{(1-\beta)}(w-b)^{\beta} \tag{9}$$

leading to a wage equation:

$$w=\beta p+(1-\beta)b \tag{10}$$

Where β can be thought of as the bargaining power of the worker which is typically thought of as exogenous to the model. The match will be consummated whenever there is some surplus to be shared i.e. whenever $p \geq b$ so that there is ex post efficiency.

Now consider a wage-posting model in which employers set the wage before being matched with a worker. To derive the optimal wage in this case we need to make some assumption about the process by which workers and employers are matched – for the moment, assume that is random though alternatives are discussed below. And assume that workers differ in their value of leisure, b – denote the distribution function of this across workers by G(b).

If the firm sets a wage w, a worker will accept the offer if w>b, something that happens with probability G(w). So expected profits after a match with a worker will be given by:

$$\pi(w)=(p-w)G(w) \tag{11}$$

This leads to the following first-order condition for wages:

$$w(p)=\frac{\varepsilon(w(p))}{1+\varepsilon(w(p))}p \tag{12}$$

Where ε is the elasticity of the function G with respect to its argument. Higher productivity firms offer higher wages. An important distinction from ex post wage bargaining is that not all ex post surplus is exploited – some matches with positive surplus i.e. with p>b may not be consummated because b>w. In matches that are consummated the rents are split between employers and workers so employers are unable to extract all surplus from workers.

In this model G(w) can be thought of as the labour supply curve facing the firm in which case can think of as standard model of monopsony in which the labour supply to a firm is not perfectly elastic and and (12) as the standard formula for the optimal wage of a monopsonist. There is a simple and familiar graphical representation of the decision-

making problem for the firm – see Figure 1. In contrast, there is no such simple representation for the outcome of the ex post wage bargaining model3.

One might think that the two wage equations (10) and (12) are very different. But they can easily be made to look more similar. Suppose that the supply of labour can be written as:

$$G(w) = (w - b)^{\varepsilon} \tag{13}$$

Where b is now to be interpreted now not as the specific worker's reservation wage but as the lowest wage any worker will work for. Then the profit function (11) is isomorphic to (9) with the bargaining power becoming the elasticity of the labour supply curve to the firm. In some sense, the bargaining power of workers in the wage-posting model is measured by the elasticity of the labour supply curve to the firm.

The assumption of random matching plays an important role in the nature of the wage-posting equilibrium so it is instructive to consider other models of the matching process. The process described here has been completely random and, taken literally, imply that workers and employers are matched completely at random, something that is absurd. Newly-minted PhDs are much more likely to apply to Harvard than they are to McDonalds. The labour market clearly segments in some way and the main alternative model is one of directed search. Models of directed search essentially assume that the market segments by wages (or expected wages).

For models based on ex post wage bargaining, this is of little consequence but for models based on wage-posting it does make an important difference. The models do not assume – like perfectly competitive models – that an application necessarily leads to a job so there is typically some frictional unemployment in equilibrium. So the expected utility of a worker applying to a particular firm is not just the wage but needs to take account of the probability of getting a job. But – and this is the most important point – all employers must, in equilibrium offer workers the same level of expected utility and can expect to hire as many workers as they want at that so the labour supply curve facing them is perfectly competitive. Armed with this insight it is then no great surprise to discover that

3 Actually, the natural place to look for familiar models which are similar would be trade union models who typically have a bargaining model for wage determination. But the tradition in ex post wage bargaining models of having one worker and one employer tends to limit the analogy.

the equilibrium in these models is quasi-competitive though this conclusion is not hugely robust (Albrecht, Gautier and Vroman, 2006). Is the model of directed search a good one? It is true that the labour market does segment but it goes back to the point that employers are idiosyncratic in lots of ways.

3.2 *The Right Model?*

Shimer, Wright and Rogerson (2005, p984) conclude their survey of search models by writing that one of the unanswered questions is 'what is the right model of wages?' with the two models described above being the main contenders. If we wanted to choose between these two descriptions of the wage determination process, how would we do so? We might think about using theoretical or empirical arguments. As economists abhor unexploited surpluses, theory would seem to favour the ex post wage bargaining models in which no match with positive surplus ever fails to be consummated[4]. One might expect that there would be renegotiation of the wage in a wage-posting model if p>b>w.

However, over a very long period of time, many economists have felt that this account is over-simplistic, that wages, for reasons that are not entirely understood, have some form of rigidity in them that prevents all surplus being extracted from the employment relationship. There are a number of possible reasons suggested for this. Hall and Lazear (1984) argue that imperfections in information at the heart of the reasons for this. Ellingsen and Rosen (2003) argue that wage-posting represents a credible commitment not to negotiate wages with workers something that would cost resources and raise wages. There is also the feeling that workers care greatly about nations of fairness e.g. see Mas (2006) so that this makes it costly to very wages for workers who see themselves as equals. If jobs were only ever destroyed when there was no surplus left to either side, there would be no useful distinction between quits and lay-offs, though most labour economists do think that distinction meaningful. The bottom line is that

4 Though this statement should not be taken to mean that markets as a whole with ex post wage bargaining need be more efficient than those with wage-posting. The efficiency concept referred to here is an ex post notion and labour market efficiency is an ex ante notion. In particular, we later show how, if we model the incentives to get a match, we might get different conclusions.

theory alone does not seem to resolve the argument about the 'best' model of wage determination.

What about empirical evidence? In a recent paper Hall and Krueger (2008) use a survey to investigate the extent to which newly-hired workers felt the wage was a 'take-it-or-leave-it' offer as ex ante wage-posting models would suggest. All those who felt there was some scope for negotiation are regarded as being ex post wage bargaining. They show that both institutions are common in the labour market, with negotiation being more prevalent. In low-skill labour markets wage-posting is more common than in high-skill labour markets, as perhaps intuition would suggest.

Interesting and novel though this study is, the classification is not without its problems. For example, some of those who report a non-negotiable wage may never have discovered that they had more ability to negotiate over the wage than the employer (successfully) gave them the impression there was. Similarly, there are potential problems with assuming that all those without stated ex ante wages represent cases of bargaining. For example, employers with all the bargaining power would like to act as a discriminating monopsonist tailoring their wage offer to the circumstances of the individual worker, not the simple monopsonist the wage-posting model assumes. Hall and Krueger (2008) are aware of this line of argument but argue it is not relevant because wage discrimination would result in all workers in the US being held to their reservation wage, a patently ridiculous claim. But, there is a big leap from saying some monpsonisitic discrimination is practiced to saying it is done perfectly so this argument is not completely compelling.

There is also the problem that the methodology used, while undoubtedly fascinating, primarily counts types of contract without looking at the economic consequences. For example, Lewis (1989, p149) describes how Salomon Brothers lost their most profitable bond-trader because of their refusal to break a company policy capping the salary they would pay. Undoubtedly, this contract should be described as individualistic wage bargaining but there were limits placed on that which resulted in some ex post surplus being lost as suggested by the wage-posting models.

One possible way of resolving these issues would be to look at outcomes. For example, ex post individualistic wage bargaining would suggest, as from (10), that there

would be considerable variation in wages within firms between workers with different reservation wages. On the other hand, ex ante wage bargaining would suggest no wage variation within firms between workers with different reservation wages. Machin and Manning (2004) examine the structure of wages in a low-skill labour market, that of care workers in retirement homes. They find that, compared to all other characteristics of the workers, a much greater share of the total wage variation is between as opposed to within firms. Reservation wages are not observed directly but we might expect to be correlated with those characteristics so ex post wage bargaining would deliver correlations with those variables.5

One could spend an enormous amount of time debating the 'right' model of wage determination. But, we will never be able to resolve it because the labour market is very heterogenous so that no one single model fits all. It is also worth reflecting that, in many regards, these models are quite similar (e.g. they both imply that rents are split between worker and employer) so that it may not make very much difference which model one uses as a modelling device. The main substantive issue in which they differ is in whether one thinks that all ex post surplus is extracted. But, because even ex post efficiency does not mean ex ante efficiency, this may not be such a big difference in practice.

However this is not to say that the choice of model has no consequences because too many economists see the labour market only through the prism of the labour market model with which they are most familiar.

For example, as illustrated above, a wage-posting model naturally leads one to think in terms of the elasticity of the labour supply curve to an individual firm and that one can represent the wage decision using a familiar diagram. It is easy to forge links with other parts of labour economics so it is perhaps not surprising that this has often been the model of choice for microeconmic models of imperefect competition in the labour market. It is much more difficult to forge such links with an ex post bargaining model and the literature that uses such models seems to have developed in a parallel universe to more conventional labour economics and has concentrated on macroeconomic

5 . This is not inconsistent with the conclusions of studies like Lazear and Shaw (2007) who argue that most wage dispersion is within firms as that is primarily about wage dispersion between managers and janitors who differ in their productivity and not among workers who might be expected to have similar levels of productivity.

applications. Though there is perhaps a tendency in recent years for some people to realize that one can also use these models to address issues of microeconomic concern, though more traditional labour economists often view these models as reinventing the wheel and not always a round one at that.

3.3 *Other Perspectives on Wage Determination*

I have described the two most commonly found models of wage determination. But just as I have emphasized that one should not be thought as obviously 'better' than the other, so one should not assume that these are the only possibles. Here we simply review some of the others that can be found in the literature. We make no attempt to be exhaustive (e.g. see Hall and Lazear, 1984, for a discussion of a range of possibilities we do not dicsuss here).

The simple model sketched above only has workers moving into jobs from non-employment because it is a one-period model. In reality, over half of new recruits are from other jobs (Manning, 2003; Nagypal, 2005) so that one has to think about how wages are determined when a worker has a choice between two employers.

In models with ex-post wage bargaining, on-the-job search is a bit tricky to incorporate into standard models because it is not clear how to model the outcome of bargaining when workers have a choice of more than one employer and different papers have taken different approaches e.g. Pissarides (1994) assumes that the fall-back position for workers with two potential employers is unemployment while Cahuc, Postel-Vinay and Robin (2006) propose that the marginal product at the lower productivity firm be the outside option. Shimer (2006) points out that, as the value function for employed workers is typically convex in the wage when there is the possibility of moving to a higher-wage job in the future and derives another bargaining solution, albeit one with many equilibria.

In contrast, models based on wage-posting do not find it hard to incorporate on-the-job search as they typically simply assume that the worker accepts the higher of the two wage offers. But, they do find it difficult to explain why the employer about to lose a worker does not seek to retain them by raising wages. A number of papers look at the institution of offer-matching (Postel-Vinay and Robin, 2002) in which the two employers

engage in Bertrand competition for the worker. However, many have felt that offer-matching is not very pervasive in labour markets and have offered reasons for why this might be the case (see, for example, the discussion in Hall and Lazear, 1984).

4 Estimates of Wage-Splitting

The previous section reviewed theoretical models of the ways in which rents are divided between workers and employers - this section reviews empirical evidence on the same subject.

An earlier section took the approach of trying to get some idea of the size of rents by considering the expenditure on rent-seeking behaviour by employers and non-employed workers. Because it produced estimates of the rents accruing to employer and worker, one could use these estimates to get some idea of how the rents are shared between employer and worker. But, because these estimates are assembled from a few, disparate sources of evidence, we have no study in which we could estimate both employer and worker rents in the same labour market so that estimating how rents are shared by using an estimate of employer rents in one labour market and worker rents in another would not deliver credible evidence. So, in this section we review some other methodologies that can be thought of as seeking to estimate the way in which rents are split between worker and employer.

The part of the literature on imperfect competition in labour markets that has used ex post wage bargaining as the model of wage determination and, consequently, uses an equation like (10) would tend to see rents being split according to the bargaining power of the workers. This is typically not explained within the model (see, for example, the discussion in Cahuc, Postel-Vinay and Robin, 2006, p330) with the parameter calibrated or estimated to help to explain some aspects of labour market data. As these are not direct approaches to estimating the sharing of rents we do not review them here.

In contrast, models that are based on wage-posting, have a monopsony perspective on the labour market and view the elasticity of the labour supply curve facing the employer as the key determinant of how rents are split.

4.1 *Estimating the elasticity of the labour supply curve to an individual employer*

4.1.1 *Experimental and Quasi-Experimental Evidence*

A natural place to start is to think about an experiment that one would like to run to estimate the elasticity of the labour supply curve to a firm. What one would like to do would be to randomly vary the wage paid by the single firm and observe what happens to employment (a fall in response to a wage rise would suggest an outcome on the demand not the supply curve). As yet, the literature does not have a study of such an experiment.

What we do have are a number of quasi-experiments where there have been wage rises in some firms. Typically those experiments have been of public sector firms where there have been perceived to be labour shortages because wages have been set below prevailing market levels. So, they sound like the type of thing where one would expect to be tracing out the elasticity of a labour supply curve. There are a handful of studies like this.

Staiger, Spetz and Phibbs (2010) examine the impact of a legislated rise in the wages paid at Veteran Affairs hospitals. This combined with a plausible argument that these hospitals were allowed to hire as many staff as they wanted (which is required to make sure we are estimating the supply curve) seems as close to an exogenous increase in wages as anything else in the literature. They estimate the short-run elasticity in the labour supply to the firm to be very low - around 0.1 implying an enormous amount of monopsony power possessed by hospitals over their nurses. Falch (2010) investigates the impact on the supply of teachers to individual schools in northern Norway in response to a policy experiment that selectively raised wages in some schools with past recruitment difficulties. He reports an elasticity in the supply of labour to individual firms in the region 1.0-1.9 – higher than the Staiger et al study but still very low. Clotfelter et al, (2008) investigate the impact of bonuses paid to teachers of some subjects in some North Carolina schools on labour turnover. They find that an $1800 bonus which is an approximate 4-5% rise in average earnings reduced the labour turnover rate by 17% percentage points. This implies an elasticity of the quit rate with respect to the wage of -4 to -36.

6 The authors note that this is higher than other estimates of this elasticity though provide some explanation why that might be the case.

Looking at these studies, one clearly comes away with the impression not that it is hard to find evidence of monopsony power but that the estimates are so enormous to be an embarrassment even for those who believe this is the right approach to labour markets. They are too large to be credible.

This means it makes sense to reflect on possible biases. There are a number that come to mind. First, some of these studies only look at the response of employment to wage changes over a relatively small time horizon. As one would expect supply elasticities to be smaller in the short-run, these estimates are not reliable as estimates of the short-run elasticity. There is a simple back-of-the-envelope rule that can be used to link short-run and long-run elasticities. Boal and Ransom (1997) and Manning (2003, chapter 2) show that if the following simple model is used for the supply of labour to a firm:

$$N_t = [1 - s(w_t)] N_{t-1} + R(w_t) \tag{14}$$

Then there is the following relationship between the short-run and long-run elasticities:

$$\varepsilon^s_{Nw} = s(w_t)\varepsilon_{Nw} \tag{15}$$

So that one needs to divide the short-run elasticity by the quit rate to get an estimate of the long-run elasticity. If, for example, labour turnover rates are about 20% then one needs to multiply the estimates of short-run elasticities by 5 to get a better estimate of the long-run elasticity. The method of Clotfelter et al (2008) which looks at the effect on separations would not be expected to be so sensitive to this which perhaps accounts for the higher elasticties found in that study. However, one needs to link the overall labour supply elasticity to the quit elasticity, an issue that is discussed in section 4.1.3 below.

A second issue is whether the wage premia are expected to be temporary or permanent. If they are only temporary then one would not expect to see such a large supply response. In this regard, it is reasonable to think of the wage increases studied by Staiger et al (2010) as permanent, those studied by Falch (2010) as temporary and those studied by Clotfelter et al (2008) as permanent, though they discuss why many teachers perceived them as temporary. It is not clear whether an argument that the wage premia were viewed as only temporary are plausible as explanations of the low labour supply elasticities found.

Here, I suggest that there is another, as yet unrecognised, problem with these estimates of labour supply elasticities. The reason for believing this comes from thinking about estimates of the labour supply elasticities from an alternative experiment – force an employer to raise its employment level and watch what happens to the wages that they pay. This is what is analyzed by Matsudaira (2009) who.analyses the effect of a 1999 California law that required all licensed nursing homes to maintain a minimum number of hours of nurses per patient. This can be thought of as a mandated increase in the level of employment.

According the simplest models of monopsony in which there is a one-to-one relationship between wages and labour supply to the firm, the wage response to the mandated employment increase should give us an estimate of the inverse of the wage elasticity. If the studies of mandated wage increases cited above are correct and the labour supply elasticity is very small, we should see very large wage increases in response to mandated employment changes. This is especially true if the short-run elasticity is very low. In fact, Matsudaira finds that firms that were particularly affected by the mandated increased in employment did not raise their wages relative to other firms who were not affected. As a result, the labour supply to the employer appears very elastic, seemingly inconsistent with studies of mandated wage increases. It is possible that, as these are studies of different labour markets there is no apparent inconsistency but I would suggest that is not the most likely explanation and that the real explanation is a problem with the simple model of monopsony.

How can we reconcile these apparently conflicting findings? The problem with the simple-minded model of monopsony is the following is that it assumes that the only way an employer can raise employment is by raising its wage. A moment's reflection should persuade us that this is not very plausible. There are a number of possible reasons for this - I will concentrate on one in some detail and then mention others.

First, the simple model of monopsony assumes there is nothing the employer can do to hire more workers except raise the wage. But we have already seen that hiring costs money and used estimates of these hiring costs to shed light on the size of employer rents from the employment relationship. If employers want to hire more workers, they can spend more resources on trying to recruit workers e.g. advertising vacancies more

frequently or extensively. The supply of workers to the firm will then be a function not just of the wage but also of the expenditure on recruitment. This model is examined in Manning (2006) who terms it the' generalized model of monopsony' and it can easily explain the paradox described above.

To see how it can do this assume there are constant marginal hiring costs, T(w), which might depend on the wage. If the separation rate is s(w) a flow of s(w)N recruits is necessary for the employer to maintain employment at N which will cost s(w)T(w)N. This represents the per period expenditure on recruitment necessary to keep employment at N if the wage paid is w. Note that, unlike the simple monopsony model, any level of employment is compatible with any level of the wage but that there is an associated recruitment costs. If, in the interests of simplicity, we ignore discounting (the recruitment costs of a worker must be paid up-front but profits accrue in the future), profits of the firm can be written as:

$$\pi = F(N) - wN - s(w)T(w)N \tag{16}$$

First, consider the choices of wage and employment by an unconstrained profit-maximizing firm. The wage will be chosen to satisfy the first-order condition:

$$-1 - s'(w)T(w) - s(w)T'(w) = 0 \tag{17}$$

Denote this choice by w*. The first-order condition for employment will then be:

$$F'(N) = w^* + s(w^*)T(w^*) \tag{18}$$

Now, consider what happens in this model when we mandate wages or mandate employment. Consider, mandated employment first as in the Matsudaira paper. If the government requires an increase in employment, the optimal thing for the firm to do is to increase recruitment activity – the optimal wage (17) remains completely unchanged. This is, to a first approximation, what Matsudaira finds. However, it tells us nothing about the degree of imperfect competition in the labour market which is related to the elasticity of separation rates and recruitment with respect to the wage.

Now consider a mandated increase in the wage. This reduces separations and may reduce the marginal cost of recruitment. But, if it is a small increase from the optimal wage the first-order effect will be to leave employment unchanged – the employer responds by reducing recruitment expenditure. One might explain the small

positive effects on employment found in the literature as being the result of mandated wage increases in public sector firms where wages had been held artificially low.

In the generalized model of monopsony, the two experiments of mandated wage or employment increases are no longer mirror images of each other. A rise in mandated wages which, ceteris paribus, leads to a rise in labour supply to the firm could be met with an off-setting fall in recruitment activity, leaving overall employment unchanged. On the other hand, a rise in mandated employment may be met with a rise in recruitment activity to generate the extra supply with no increase in wages.

We have used a very simple model to break the one-to-one link between wages and employment found in the standard model of monopsony. The change is plausible but does substantially affect how one interprets empirical results. The simple example assumes a constant marginal cost of recruitment – as we shall see, whether or not marginal costs are increasing is important and our evidence on the subject is not as strong as we would like. But the basic point remains.

This is not the only way in which one might seek to reconcile these conflicting empirical findings. Another alternative is to assume that workers are heterogeneous in terms of quality so that employers also face an intensive margin in deciding what is the cut-off quality level for workers. Employers do not simply accept all workers who apply – they reject those they deem of poor quality and how poor one has to be to be rejected is clearly endogenous. An example in the appendix shows how, if the distribution of worker ability in the applicant pool is exponential then firms respond to mandated wage increases by increasing worker quality and not employment and to mandated employment increases by reducing worker quality and not increasing wages. It also shows how a model with non-wage aspects of work deliver the same conclusion.

All of these quasi-experimental studies described above are studies of mandated changes to wages or employment which might be thought to force employers to move along their labour supply curves. But, another empirical strategy is to consider changes in variables which induce moves along the labour supply curve. To identify the labour supply curve (which is all we want here) a variable that shifts the MRPL curve without shifting the supply curve is needed. One can then use this as an instrument for the wage or employment (depending on which way round we are estimating the supply curve) in

estimating the supply curve. But, of course, it requires us to be able to provide such an instrument.

If one is interested in estimating the elasticity of labour supply to an individual firm then the instrument needs to be something that affects the demand curve for that firm but has negligible impact on the labour market as a whole. The reason is that a pervasive labour market demand shock will raise the general level of wages so is likely to affect the labour supply to an individual firm. So, for example, the approach of using demand shocks caused by exchange rate fluctuations (as in Abowd and Lemieux, 1989) does not seem viable here. There are a number of studies that attempt to use firm-level instruments. For example, Sullivan (1989) used the population in the area surrounding the hospital as an instrument affecting the demand for nurses and Beck, Boal and Ransom (1998) use the number of children in a school district as an instrument for the demand for teachers. These represent serious attempts to deal with a difficult problem but their instruments are not beyond criticism. If the main variation in the number of children or the number of patients comes from variation in population it is also likely that the supply of nurses and teachers in an area is proportional to population as well.

4.1.2 *Non-Experimental Studies*

In the absence of a good set of experimental or quasi-experimental studies of the labour supply curve to an individual employer, we need to also discuss non-experimental estimates. In the literature there have been tow main methodologies used.

As the labour supply curve is a positive relationship between wages and employment one might simply look at this relationship. That this is positive is a well-known and robust empirical finding - the employer size-wage effect (Brown and Medoff, 1989, Oi and Idson, 1999). But there are problems with using this as an estimate of the elasticity of the labour supply curve to an employer. Manning (2003, chapter 4) discusses the problems and nobody has seriously tried this approach.

The approach that has been used more commonly is to attempt to estimate the elasticity of the separation rate with respect to the wage. Why this might be thought useful can be explained very simply. Suppose that the flow of recruits rate to a firm is

R(w), dependent on the wage and the separation rate is s(w) also dependent on the wage. In a steady-state, recruits must equal separations which leads to:

$$N(w) = \frac{R(w)}{s(w)} \tag{19}$$

As pointed out by Card and Krueger (1995), this implies that:

$$\varepsilon_{Nw} = \varepsilon_{Rw} - \varepsilon_{sw} \tag{20}$$

so that knowledge of the elasticities of recruitment and quits with respect to the wage can be used to estimate the elasticity of labour supply facing the firm. The elasticity of separations with respect to the wage is important here but so is the elasticity of recruits with respect to the wage. However, as discussed below there are arguments for linking the two. But, before discussing that argument, let us discuss how the sensitivity of quits with respect to the wage can be estimated.

There is a long tradition of being interested in the sensitivity of labour turnover to the wage, quite apart from any insight these studies might have for the extent of imperfect competition in the labour market. These studies are not confined to economics e.g. see Griffeth, Hom and Gaertner (2000) for a meta-analysis from the management literature. The bottom line is that, as predicted by models of imperfect competition a robust negative correlation between the wages paid and labour turnover is generally found. However, these elasticities are not generally found to be high – Table x indicates estimated elasticities in the range.

As these elasticities are not high it is worth reflecting on the way in which they may be biased. The basic equations regress some measure of labour turnover on the wage and other covariates [expand].

There may be omitted variables, correlated with the wage. The most serious problem in estimating the wage elasticities is, as always, going to be the result of a failure to control adequately for other relevant factors. One potential source of problems in estimating the separation elasticity is a failure to control adequately for the average level of wages in the individual's labour market. Separations are likely to depend on the wage relative to this alternative wage so that a failure to control for the alternative wage is likely to lead to a downward bias on the wage elasticities. On the other hand, we would expect separations to be more sensitive to the permanent component of wages than to the

part of wages that is a transitory shock or measurement error. In this case, the inclusion of controls correlated with the permanent wage is likely to reduce the estimated wage elasticity. Manning (2003, chapter 4) investigates this and finds that, for a number of US and UK data sets, the inclusion of standard human capital controls does not make much difference to the estimated wage elasticities.

However, one variable whose inclusion or exclusion makes a lot of difference to the apparent estimated wage elasticity is job tenure[7]. The inclusion of job tenure always drastically reduces the estimated wage elasticity as high-tenure workers are less likely to leave the firm and are more likely to have high wages. There are arguments both for and against the inclusion of job tenure. One of the benefits of paying high wages is that tenure will be higher so that one needs to take account of this indirect effect if one wants the overall wage elasticity when including tenure controls: in this situation, excluding tenure may give better estimates. On the other hand, if there are seniority wage scales, the apparent relationship between separations and wages may be spurious. One study that attempts to deal with this problem is Ransom (2010), which investigates how turnover is related to wages among Missouri school teachers. They use as their wage measure, the base wage in the school district.

The bottom line from these studies is that while wages do affect quit rates, worker mobility does not appear to be very sensitive to the wage. While it is possible that various biases mean that the estimates we have are too low, one would have to believe these biases are extremely large to make the quit elasticity very large. For example, the Clotfelter et al (2008) experiment did not find an elasticity above 4. These studies generally have a naïve view of worker decision-making but the study of Fox (2010) who estimates a dynamic forward-looking model of worker turnover also finds low elasticities.

4.1.3 *The link between separation and recruitment elasticities*

The studies that have used the separations elasticity to estimate the elasticity of labour supply to the individual employer have all equated the recruitment elasticity to the

7 The word 'apparent' is appropriate here because the dependence of job tenure on the wage needs to be taken account of here when estimating the full wage elasticity.

separation elasticity, essentially using the formula in (20) to double the separation elasticity to get an estimate of the elasticity of labour supply to an individual employer. Equating the quit and recruitment elasticities was first proposed in Manning (2003) and attracts a certain amount of suspicion, some suspecting it something of a sleight of hand. In fact, there are good reasons to believe it a reasonable approximation for separations to other jobs and recruits from other jobs. The reason is that when a worker leaves employer A for employer B because B offers a higher wage, this is a worker who is recruited to B because it is paying a higher wage than A.

To illustrate the robustness of the result a more general result is shown here, using the generalised model of monopsony in which employers can also influence their supply of labour by spending more resources on recruitment. We will only model job-to-job transitions and consider moves between employment and non-employment at the end. Assume that job offers arrive at a rate λ and that the distribution of wages in those job offers is $g(x)$. Furthermore assume that a worker who is currently paid w and who receives a job offer of x will leave with a probability $\phi\left(\frac{x}{w}\right)$. If the wage is the only factor in job mobility decision this will be one if x is above w and zero if it is below but it is probably more realistic to think of it as a differentiable function. The assumption that it is only the relative wage that matters is the critically important assumption for what follows but it is not an unreasonable assumption. If this condition was not satisfied, one would expect that, as average wages rise, separations to trend up or down. No clear trend [data - reference] suggests this is a reasonable approximation. Define $\varepsilon_{\phi}\left(\frac{x}{w}\right)$ to be the elasticity of $\phi\left(\frac{x}{w}\right)$ with respect to its argument – we will call this the wage-specific quit elasticties.

Consider a firm that pays wage, w. The overall separation rate will be given by:

$$s(w) = \lambda \int g(x)\phi\left(\frac{x}{w}\right)dx \tag{21}$$

The appendix then proves the following result:

Result 1: The elasticity of the separation rate with respect to the wage is given by:

$$\varepsilon_s(w) = \frac{ws'(w)}{s(w)} = \int g_s(x;w)\varepsilon_\phi\left(\frac{x}{w}\right)dx \tag{22}$$

Where $g_s(x;w)$ is the share of separations in a firm that pays w that go to a firm that pays x i.e.

$$g_s(x;w) = \frac{g(x)\phi\left(\frac{x}{w}\right)}{\int g(x')\phi\left(\frac{x'}{w}\right)dx'} \tag{23}$$

Proof: See Appendix

(22) says that the overall separation elasticity can be thought of as a weighted average of the wage-specific elasticities where the weights are the shares of quits to different wages.

To derive the elasticity of recruits with respect to the wage we need to think about the distribution of wage offers, g(w). This will be influence by the distribution of wages across firms – which we will denote by f(w) and, we will assume, the hiring activity of firms. If $H(w)$ is the amount spent by a firm that pays w on hiring, then we will assume that the distribution of wage offers is given by:

$$g(w) = \frac{H(w)^\beta f(w)}{\int H(x)^\beta f(x)dx} = \left(\frac{H(w)}{H}\right)^\beta f(w) \tag{24}$$

Where:

$$H = \left[\int H(x)^\beta f(x)dx\right]^{\frac{1}{\beta}} \tag{25}$$

Is an index of aggregate hiring activity. It is natural to assume that λ the job offer arrival rate depends on the aggregate hiring rate as well as other factors (e.g. the intensity of worker job search). It is natural to assume that λ, the job offer arrival rate depends on the aggregate hiring rate as well as other factors (like the level of job search by workers). As we shall see, the parameter β is of critical importance – it measures whether marginal costs of recruitment are increasing or decreasing in the level of recruitment.

Now, consider recruitment. The flow of recruits to a firm that pays w and recruits at intensity h can be written as:

$$R(w,h)=\left(\frac{h}{H}\right)^{\beta}\lambda\int f(x)N(x)\phi\left(\frac{w}{x}\right)dx=\left(\frac{h}{H}\right)^{\beta}R(w) \tag{26}$$

Where $N(x)$ is employment in a firm that pays x. From this we have that:

Result 2: The elasticity of the recruitment rate with respect to the wage is given by:

$$\varepsilon_R(w)=\frac{wR'(w)}{R(w)}=\int g_R(x,w)\varepsilon_\phi\left(\frac{w}{x}\right)dx \tag{27}$$

Where:

$$g_R(x,w)=\frac{f(x)N(x)\phi\left(\frac{w}{x}\right)}{\int f(x')N(x')\phi\left(\frac{w}{x'}\right)dx'} \tag{28}$$

Is the density of recruits to a firm that pays w from firms that pay x.

Proof: See Appendix

Comparing (23) and (27) one can see the inevitable link between the quit elasticity and the recruitment elasticity – they are both averages of the individual wage elasticities. The quit elasticity for a firm that pays w is a weighted average of the elasticity of quits to firms that pay other wages with the weights being the share of quits that go to these firms. The recruitment elasticity for a firm that pays w is a weighted average of the elasticity of quits from firms that pay other wages to firms that pay w with the weights being the share of recruits that come from these firms. If this function was iso-elastic then quit and separation elasticities have to be equal though this is impossible as ϕ has to be between zero and one. However, a further result shows how they must be linked.

For an individual firm the quit and recruitment elasticity will not generally be the same but, averaging across the economy as a whole they must be.

Result 3: the recruit-weighted recruitment elasticity must be equal to the recruit-weighted quit elasticity i.e.:

$$\int f(w) R\left(w, H(w)\right) \varepsilon_R(w)\, dw = \int f(w) R\left(w, H(w)\right) \varepsilon_s(w)\, dw \tag{29}$$

Proof: See Appendix.

This is all about moves between employers. One cannot apply the same approach for the elasticity of separations to non-employment and recruits from non-employment as there is no need for one to be the mirror image of the other. However, Manning (2003) discusses how one can deal with this problem.

The bottom line from this is that while separations are negatively related to the wage, the elasticity is not especially high. Table 4 reports some estimates [to do]. Many workers seem reluctant to move for sizeable wage differentials. This also implies that the flow of recruits to an employer is not very sensitive to the wage.

This conclusion is robust to using the generalized rather than the simple model of monopsony, and the wage elasticity referred to is the elasticity holding constant the recruitment activity of the firm. So, using (26) one would have:

$$N(w,h) = \frac{R(w,h)}{s(w)} = \left(\frac{h}{H}\right)^{\beta} \frac{R(w)}{s(w)} = \left(\frac{h}{H}\right)^{\beta} n(w) \tag{30}$$

And the elasticity of this with respect to the wage is approximately twice the quit elasticity.

However, the way in which one interprets and uses this elasticity does need to be modified. Using a simple-minded model of monopsony, one would be inclined to conclude that there is an incredible amount of monopsony power in labour markets and conclude there is a massive amount of exploitation in the labour market that could, for example, be reduced by a very large increase in the minimum wage. In a later section we make clear that this is not the correct conclusion.

4.1.4 The *Marginal and Average Costs of Hiring revisited*

Earlier, we discussed how important is whether there are increasing marginal costs to hiring but also emphasized how hard it is to get good estimates of this parameter. Here, we show how an estimate can be backed-out from the model described above.

Consider a firm choosing the wage and recruitment intensity to maximize steady-state profits:

$$\pi = F(N) - wN - h \tag{31}$$

Subject to the constraint that labour supply is given by:

$$N = \frac{R(w,h)}{s(w)} = \frac{R(w)}{s(w)}\left(\frac{h}{H}\right)^{\beta} = n(w)\left(\frac{h}{H}\right)^{\beta} \tag{32}$$

The first-order condition for the wage is going to be:

$$\pi = \left[F'(N) - w\right]\frac{\partial N}{\partial w} - N = 0 \tag{33}$$

Which can be re-written as the following familiar condition:

$$w = \frac{\varepsilon_n}{1+\varepsilon_n}F'(N) \tag{34}$$

So that the relationship between the wage and the marginal product is the familiar one. If, as the estimates discussed above suggest, the elasticity is low there will be a big gap between the marginal product and the wage. This then implies that employers make considerable rents from the employment relationship so should be prepared to spend quite large amounts of money to hire workers. But, as we saw in the previous section, the estimates of the average hiring cost are, while not trivial, not enormous. What we show here is that these two facts can only be reconciled if there is a big difference between the marginal and average costs of hiring which implies strongly diminishing returns to hiring expenditure.

To see this, consider the choice of hiring rate. From (31) and (32) this will be given by:

$$\left[F'(N) - w\right]\frac{\partial N}{\partial h} - 1 = 0 \tag{35}$$

Which can be written as:

$$\left[F'(N) - w\right]\frac{\beta N}{h} = 1 \tag{36}$$

So that the optimal hiring cost per worker is given by:

$$\frac{h}{N} = \beta\left[F'(N) - w\right] \tag{37}$$

Using (34) this can be re-arranged to give:

$$\frac{h}{wN} = \frac{\beta}{\varepsilon} \tag{38}$$

The left-hand side is the ratio of total expenditure on hiring to the total wage bill. We have already discussed data on this. We have also discussed how one can get an estimate of ε from the separation elasticities. This can then be used to give us an estimate of β the sensitivity of recruits to hiring expenditure. The implied value is small – for example, if the elasticity is 5 and hiring costs are 5% of the total wage bill, this implies that $\beta = 0.25$.

4.2 *Measuring Labour Market Frictions.*

A simple yet plausible idea is that the higher the degree of competition among employers for workers, the greater will be workers' share of the surplus. In the important and influential strand of work that sees rents in the labour market as deriving primarily from labour market frictions, the fact that it takes time for workers and employers to find each other, a natural way to capture this idea is to seek some measure of transition rates between employment and non-employment and from one employer to another.

The particular measure that has been used is the ratio of the arrival rate of job offers for an employed worker (denote this by λ_e) to the rate at which workers leave employment for non-employment (denote this by δ). A higher value of (λ_e / δ) is more competition among employers for workers which would be expected to raise wages. In many canonical search models e.g. Burdett and Mortensen (1998), the share of rents going to be workers can be shown to be some function of (λ_e / δ). It can be interpreted as the expected number of job offers a worker will receive in a spell of employment (Ridder and van den Berg, 2003).

There are a lot of measures of (λ_e / δ) in the literature, with a large degree of variation. Often these estimates come from the estimation of structural models in which

it is not entirely clear which features of the data play the most important role in influencing the estimates. Here, we will simply describe ways in which (λ_e / δ) can be estimated directly using data on labour market transition rates.

δ can be estimated very simply using data on the rate at which the employed leave for non-employment. λ_e is more complicated as the theoretical concept is the rate at which job opportunities arrive to the employed. One might think about simply using the job-to-job transition rate but as the employed only move jobs when the new offer is better than the current one, this is an under-estimate of the rate at which new job opportunities arise. However, in simple search models there is a mapping between the two. The reason is that if all workers always prefer high-wage to low-wage jobs and always move whenever they get a higher wage offer (however small the wage gain), then there is a simple expression for the fraction of workers G(f) who are in jobs at or below position f in the wage offer distribution. Equating inflows and outflows we have that:

$$\left[\delta + \lambda_e (1-f)\right] G(f)(1-u) = f \lambda_u u \tag{39}$$

As, in steady-state we must have that:

$$u = \frac{\delta}{\delta + \lambda_u} \tag{40}$$

This can be written as:

$$G(f) = \frac{\delta f}{\left[\delta + \lambda_e (1-f)\right]} \tag{41}$$

Now the transition rate to unemployment rate is δ and the transition rate to other jobs is:

$$\begin{aligned} \lambda_e \int (1-f)\, g(f)\, df = \lambda_e \int G(f)\, df &= \int \frac{\lambda_e \delta f}{\left[\delta + \lambda_e (1-f)\right]} df \\ &= \delta \left[\frac{\delta + \lambda_e}{\lambda_e} \ln\left(\frac{\delta + \lambda_e}{\delta} \right) - 1 \right] \end{aligned} \tag{42}$$

Which means that the ratio of transition rates to employment relative to transition rates to non-employment is given by:

$$\left[\frac{1+k}{k} \ln(1+k) - 1 \right] \tag{43}$$

Which is monotonically increasing in k. In a steady-state this can be shown to be equal to the fraction of recruits who come from unemployment, a measure proposed by Manning (2003).

One might wonder about the relationship between (λ_e / δ) and estimates of the labour supply elasticity discussed earlier in this section. In many search models there is a simple connection between the two because one can always write the profit-maximizing choice of the wage as being related to the elasticity of the labour supply curve to the firm so that (λ_e / δ) must be related to this. However, if, for example, one relaxed the assumption that it is only current or future wages that motivate wage changes, then (λ_e / δ) would not seem to be a good measure of the market power of employers while an estimate of the wage elasticity still gets to the heart of the issue.

How do estimates of the balance of power between workers and employers based on this methodology compare to those based on the wage elasticity of the labour supply curve (or separations). The advantage is perhaps that they are relatively easy to compute but the disadvantage is that they are indirect and may rely for their validity on assumptions that do not hold. For example, in these models perfect competition is the case where there is massive churning of workers, where the employer you work for one day (or hour?) has no bearing on who you work for the next. In some sense, that is a correct characterization of a perfectly competitive equilibrium as that determines the market wage but not who of the large number of identical employers a worker works for which is indeterminate. But, the inclusion of even a small fixed cost of changing jobs would change the prediction to one of very little turnover in an equilibrium close to perfect competition. Secondly, there is good reason to believe that not all turnover is for wage gains which is what is relevant for employers deciding on the wage to pay. The one empirical application (Hirsch and Schumacher, 2005) does not find this measure works well in explaining variation in nurse pay across US cities.

4.3 *Conclusions*

This section has reviewed estimates we have of the distribution of rents in the typical employment relationship, While this might be regarded as intrinsically interesting, one

still has to deal with the 'so what?' question, what difference does this make to how one thinks about labour markets.

5. So What?

So why is an imperfect competition perspective not pervasive in labour economics? There are two sorts of answers. First that it has little value-added above the perfectly competitive model – it adds more complication than insight8. This might be because perfect competition is seen as a tolerable approximation to reality so that the mistakes one makes by assuming the labour market is perfectly competitive are small. Or it might be because, as is definitely the case, the comparative statics of models of imperfect and perfect competition are in many cases the same so both perspectives give the same answer. For example, shifts in the demand curve and supply curve of labour will be predicted to have the same effects in perfect and imperfect competition.

The second reason why labour economists do not do their work while adopting the perspective that the labour market is imperfectly competitive is that they do not adopt any conceptual framework at all. A well-designed and executed randomized experiment tells us about the effect of an intervention without the need for any theory or conceptual framework at all. A generation of labour economists have grown up who are not accustomed to thinking in terms of economic models at all, seeking instead good research designs. But, while estimates from randomized experiments have internal validity, their external validity is more problematic. The results tell us what happened but not why. And without at least some understanding of 'why' it is difficult to draw conclusions from such studies that are of general use and enable us to make a forecast of will happen with a similar but not identical treatment in another time and place. We want to use evidence not just to understand the past but to improve the future. In practice, people do assume estimates have external validity all the time – they implicitly generalize. But perhaps it would be better if this as more explicit and we had a theory of why and this is where an overall perspective on the workings of the labour market might help.

8 Although, there is a part of economics that sees complication as a virtue and there does seem to be a part of research on imperfect competition in labour markets that is attracted to that.

At the other extreme are structural models and there is a small industry of structural modelling of the labour market based on imperfect competition (see, for example, [expand] and the survey in Eckstein and van den Berg, 2006). Structural models have the advantage is that they can be used to make a prediction about anything. However, the problem is that one can estimate any model, however crazy (just write down its likelihood function and maximize it) so it is not clear that the predictions of these models are any good. It is often disturbing how uninterested structural modellers are in whether their model is the right model and how obviously poorly many of these models could do in dimensions other than that which is sought to be fitted to the data. The discussion of identification often leaves a lot to be desired.

6. Applications

As argued in the previous section, labour economists will probably only be convinced of the merits of thinking about labour markets through the lens of imperfect competition if they can be convinced that it makes a difference to perspectives on certain issues. In this section we review several areas in which it has been argued to make a difference though we make no claims that this is exhaustive and we try to list others at the end.

6.1 *Labour Market Regulation*

If labour markets are perfectly competitive then we know that the equilibrium will be Pareto efficient and that regulation can only be justified on distributive and not efficiency grounds. If labour markets are imperfectly competitive there is no such presumption that the market is efficient and there is at least the potential for some regulation to improve efficiency.

The labour market regulation that has received the most attention is the minimum wage. If the labour market is perfectly competitive then a minimum wage must reduce employment as it raises the cost of labour. However, this is not necessarily the case if the labour market is imperfectly competitive. To illustrate this, we will consider the case of monopsony though one could do the same with a matching-style model.

In the simplest model of monopsony in which there is a single employer and the wage is the only available instrument for influencing its labour supply, there is a very

simple formula relating the minimum wage to the elasticity of the labour supply to an individual employer. As we have emphasized that the labour supply to individual firms is not very sensitive to the wage, this would suggest very large potential rises in employment could be obtained from an artfully chosen minimum wage.

However, there are at least two important reasons for why such a conclusion is likely to be misleading. First, we have emphasized how the simple model of monopsony is not the best way to think about the labour market. Secondly, the model of market power we have used is a model of a single employer that ignores interactions between employers so is only a partial equilibrium analysis.

Lets consider the first point first. Take the model of the previous section in which the labour supply curve is given by (30) and can be influenced not just by the wage paid but also by the level of recruitment activity. To keep things simple assume the marginal revenue product of labour is constant and equal to p. First, consider the optimal employment level given the wage paid . This satisfies the first-order condition:

$$(p-w)=\frac{\gamma}{N}\left[\frac{N}{n(w)}\right]^{\gamma} \tag{44}$$

Where $\gamma=\left(\frac{1}{\beta}\right)$. Profit-maximization leads to the following level of employment:

$$N=n(w)^{\frac{\gamma}{\gamma-1}}\left[\frac{(p-w)}{\gamma}\right]^{\frac{1}{\gamma-1}} \tag{45}$$

Assume that $n(w)$ is iso-elastic with elasticity ε. If the employer has a free choice of the wage we know they will choose a wage like(34). Now, consider the effect of a binding minimum wage. First, consider the minimum wage that will maximize employment i.e. the wage that maximizes (45). It is easy to show that this is given by:

$$w^{*}=\frac{\gamma\varepsilon}{1+\gamma\varepsilon}p \tag{46}$$

The important point is that this is bigger than the wage that the employer will choose for itself which will be given by:

$$w^{m}=\frac{\varepsilon}{1+\varepsilon}p \tag{47}$$

Where the 'm' superscript denotes the choice of a monopsonist. The log difference between the free market wage and the employment-maximizing wage is hence given by:

$$\ln w^* - \ln w^m = \ln\left(\frac{\gamma\varepsilon}{1+\gamma\varepsilon}\right) - \ln\left(\frac{\varepsilon}{1+\varepsilon}\right) = \ln\left(\frac{\gamma+\gamma\varepsilon}{1+\gamma\varepsilon}\right) \quad (48)$$

Now consider the gain in employment from an artfully chosen minimum wage. Using (45) and the wage equations (46) and (47), one can show that this is given by:

$$\ln N^* - \ln N^m = \frac{1}{\gamma-1}\ln\left(\frac{1+\varepsilon}{1+\gamma\varepsilon}\right) + \frac{\gamma\varepsilon}{\gamma-1}\ln\left(\frac{\gamma+\gamma\varepsilon}{1+\gamma\varepsilon}\right) \quad (49)$$

The standard monopsony case corresponds to the case where $\gamma = \infty$. This leads to the prediction of very large potential employment gains from an artfully-chosen minimum wage e.g. even a high wage elasticity of 10 leads to a predicted employment gain of 95 log points from a wage rise of 9.5 log points. But if $\gamma = 2$ this is much lower.

The important point to note is that, unlike the simple model of monopsony, the potential gains from the minimum wage are not just influenced by the wage elasticity ε but also the parameter γ which is the relationship between average and marginal costs of hiring.

This is a partial equilibrium conclusion and not a reliable guide for policy. There are two important distinctions between partial equilibrium models of monopsony and general equilibrium models of oligopsony. First, in general equilibrium there is an important distinction between the elasticity of labour supply to the market as a whole and to individual employers. While the gap between marginal product and the wage is determined by the elasticity of the labour supply curve facing an individual employer, any employment effect will be determined by the elasticity of the labour supply curve to the labour market as a whole. There is no reason why these should be the same but it is exactly that assumption that is made by the model of a single monopsonist.

Secondly, it is important to take account of heterogeneity. There is no doubt that the minimum wage is a blunt instrument, applied across whole labour markets on employers who would otherwise choose very different wages. This means that it is almost certainly the case that the minimum wage will have different effects on employment in different employers and any measure of the impact on aggregate employment must take account of this heterogeneity. Manning (2003, chapter 12) takes

account of both these affects showing that even in a labour market in which all employers have some market power, a minimum wage, however low, may always reduce employment.

However, models of imperfect competition are different from models of perfect competition in not making a clear-cut prediction about the employment consequences of raising the minimum wage. It is empirical studies that are important and, though this is a long debate which will not be surveyed here (see Brown, 1999 for an earlier survey), recent studies with good research designs typically fail to find any negative effects on employment for the moderate levels of minimum wages set in the US (Dube, Lester and Reich, 2009, Giuliano, 2009).

Although the employment effect of minimum wages has become the canonical issue in wider debates about the pros and cons of regulating labour markets, one should also recognise that models of imperfect competition in the labour market often have different predictions from competitive models about many interventions. For example, one can show that regulation to restrict aspects of labour contracts like hours or holidays can improve employment (Manning, 2003, chapter 8). However, although imperfect competition can be used as a justification for some regulation on efficiency grounds, it always predicts some limits to regulation with quite what those limits are left to empirical research to decide.

6.2 The Law of One Wage

In a perfectly competitive market, the elasticity of labour supply to a single firm is perfectly elastic at the market wage for that type of worker9. Any attempt to pay a lower wage will result in a complete inability to recruit any workers at all while any higher wage simply serves to reduce profits. As a result, all employers who employ this type of worker will pay them the same wage – the law of one wage holds. And all workers of that quality will be paid the same wage, irrespective of their reservation wage.

Those who have studied actual labour markets have often observed that the law of one wage seems to be violated, that there is, to use the jargon, equilibrium wage

9 Abstracting from compensating differentials.

dispersion. The existence of equilibrium wage dispersion requires some degree of imperfect competition in labour markets.

In models of imperfect competition that are based on ex post wage bargaining, it is simple to explain the existence of equilibrium wage dispersion. Refer back to the wage equation (10) – this has wages depending on the specific productivity of that employer and the specific reservation wage of the worker, something that should not happen in a perfectly competitive labour market10.

In wage-posting models the most celebrated paper is Burdett and Mortensen (1998). They present a model with homogeneous workers and employers in which the only possible equilibrium is a wage distribution with no mass points. While that is an elegant and striking result, there is a very good reason for thinking it is deficient as an account of the origin of equilibrium wage dispersion. The reason is that one can track the result to an assumption of the model which is very unappealing as an assumption about the real world and, if this assumption is made more realistic, the result collapses. That assumption is that all workers will move for the smallest gain in wages. How this delivers equilibrium wage dispersion as the only possible equilibrium can be explained with a simple diagram. Think about the labour supply curve facing an individual employer in which there is a mass of firms paying some wage w_0. The labour supply curve will be discontinuous at this point so looks something like that drawn in Figure 2. No profit-maximizing employer would then want to pay the wage w_0 – they would rather pay something infinitesimally higher and get a lot more workers. The mass point will unravel.

But the assumption that all workers move for the smallest gain in wages is totally implausible so this is not a credible account of the origin of equilibrium wage dispersion. Furthermore, we do observe mass points of wages at, for example, the minimum wage and round numbers. Does this mean this type of model has no credible explanation of equilibrium wage dispersion? Far from it – the simplest and most plausible explanation is that, faced with the same labour supply curve that is always continuous in the wage, heterogeneous employers will choose to locate at different points on that supply curve.

10 Though a statement like this should not be confused with the fact that the level of reservation wages and marginal products will affect the equilibrium wage in a perfectly competitive market.

As put succinctly by Mortensen (2003, p6) "wage dispersion is largely the consequence of search friction and cross-firm differences in factor productivity".

The failure of the law of one wage in labour markets has important consequences, some of which we will discuss below. It means that achieving a higher level of earnings is, in part, the result of working oneself into the best jobs. One possible use for that is discussed in the next section on the gender pay gap.

6.3 The Gender Pay Gap

When Joan Robinson invented the term monopsony she used it as a potential explanation of the gender pay gap. If the labour supply of women to a firm is less elastic than that of men, then a profit-maximizing employer will choose to pay lower wages to women than men even if they have the same productivity.

A recent literature essentially builds on that observation to explain at least part of the gender pay gap. The main approach has been to see whether the separation elasticity of women is lower than that of men and then apply the logic outlined in section x.x to argue that this can explain some of the gender pay gap. Manning (2003, chapter 6) estimated quit elasticities for large-scale US and UK data sets but failed to find any difference in quit elasticities . more recently, Hirsch, Schrank and Schnabel (2010) and Oaxaca and Ransom (2010) have found such a difference. Some estimates of the different separation elasticities to be found in the literature are contained in Table x.

It is important to realize that a difference in quit elasticity is not necessary for models of imperfect competition to be able to explain the gender pay gap. Nor is actual wage discrimination by employers. It could simply be that women are more likely to interrupt their careers with spells of non-employment, primarily to look after young children. In a labour market where the law of one wage does not hold, this will reduce the ability of women to work themselves into and remain in the best-paying jobs. Several recent studies of the gender pay gap find that career interruptions can explain a sizeable proportion (Bertrand, Goldin and Katz, 2009). While the most common explanation for this is that those with career interruptions accumulate less human capital, the size of the pay penalty for even small interruptions seem very large. It is not surprising that career interruptions reduce wages, but is the penalty proportionate?

6.4 Economic Geography

Much of economic geography is about explaining the distribution of economic activity over space – in particular, why it is so uneven, the phenomenon of agglomeration. There are many theories of agglomeration which are not reviewed here but some of these involve the labour market. In his classic discussion of agglomeration, Marshall (1920) about possible labour market explanations e.g. "a localized industry gains a great advantage from the fact that it offers a constant market for skill. Employers are apt to resort to any place where they are likely to find a good choice of workers with the special skill which they require; while men seeking employment naturally go to places where there are many employers who need such skill as theirs and where therefore it is likely to find a good market. The owner of an isolated factory, even if he has access to a plentiful supply of general labour, is often put to great shifts for want of some special skilled labour; and a skilled workman, when thrown out of employment in it, has no easy refuge".

The important point is these arguments make little sense if the labour market is perfectly competitive. In such a market the prevailing wage conveys all the information a firm or workers needs to know about the labour market11. fn. In a perfectly competitive labour market, an employer who is small in relation to the whole market will not care about the total supply of labour to the market except insofar as it affects the prevailing level of wages.

Hence, to make any sense of Marshall's arguments, one would seem to require some degree of imperfect competition in labour markets. This is the case – for example the formalization in Krugman (1991) rests explicitly on there being a small number of employers in the labour market.

Once the labour market is monopsonistic one can begin to make sense of some of Marshall's arguments for agglomeration. If the labour supply curve to an individual employer is upward-sloping it makes sense to talk about a labour supply curve being 'further out' because of a generally high supply of labour. One might think that

11 Although, it may be that, when making a relatively long-term location decision, it is not just the level but also the variability in wages that affects choices.

monopsony models would struggle to explain agglomeration because it might be thought that an employer would like to be the only employer in an area because they would then have enormous monopsony power over the workers in that area. But, that is based on a misunderstanding. Although the degree of monopsony power over the workers in an area will be high, there will be few of them and this is not to the adavantage of an employer. Figure 3 conveys this very simply. It draws two labour markets one in which there are very few workers but over whom the employer has a lot of monopsony power so the labour supply curve is very inelastic. In the other, there are more workers but less monopsony power. In which labour market will the employer choose to locate? They will choose the market where the level of employment they desire can be obtained most cheaply. So, if the desired level of employment is very low, they will choose market A while if it is higher they will choose market B. Manning (2009) uses this idea to explain the existence of aggloermation where employers who desire to be small locating in rural areas where they have more monopsony power and large employers locating in urban areas. And Overman and Puga (2009) investigate the implication that firms with more volatile employment will want to locate where the labour supply curve is more elastic.

The current literature on agglomeration tends to focus on the product market more than the labour market – there is considerable useful research that could be done on labour market explanations.

Another area that has received some attention is commuting.

6.5 Human Capital Accumulation and Training

Imperfection in labour markets has important implications for the incentives to acquire human capital and make investments to raise productivity. As shown by Acemoglu (1997) part of the returns to investments by workers in general human capital can be expected to accrue to future employers of the worker as the wage will be below the marginal product. This could be used to provide a justification for the massive level of public subsidy to education.

Imperfect labour markets can also offer an explanation for why firms often seem to pay for the acquisition of general training by their workers – this has been exploted in

a series of papers by Acemoglu and Pischke (1999, 2001). A recent paper that seeks ot provide evidence for this is Benson (2009) who investigates the reason why many hospitals sponsor students to train as nurses in local nursing schools. In a perfectly competitive labour market, this behavior would not make sense as it is a subsidy to general training. But, in a monopsonistic labour market one can explain it as a desire of a local employer to increase its supply of labour if, as seems plausible and can be verified from the data, nurses are likely to remain in the area in which they trained. But the incentives for hospitals to subsidize nurse-training are higher where the hospital represents a higher share of nurse employment. In labour markets where there are several hospitals one might expect them to subsidize joint programs as they have a collective interest in increasing nurse supply. Benson (2009) confirms these predictions.

6.6 Macroeconomics

A separate chapter covers this but it should be mentioned.

6.7 A Miscellany

There are many other labour market phenomena where imperfect competition might be thought to offer plausible explanations. Examples include the growth in wages over the life-cycle [search capital], the earnings assimilation of immigrants. Hotchkiss and Quispe-Agnoli (2009) argue that monopsony can be used to explain why undocumented workers earn lower wages while the firms that employ them seem to make more profits.

7. Conclusion

[to come]

Appendix

A Model with Heterogeneous Worker Ability

Here we present a model to explain the difference in the apparent labour supply elasticity from a mandated wage increase and a mandated employment increase.

For simplicity, let us assume that the labour supply of workers of quality a to a firm that pays wage w, $L(w,a)$ is given by:

$$L(w,a) = L(w)\,f(a) \tag{50}$$

Where we assume f(a) is a density function. A firm has to make two decisions – the wage to pay and the minimum quality worker, a*, to employ. Profits will be given by:

$$\begin{aligned}\pi(w,a) &= pL(w)\int_{a^*} af(a)\,da - wL(w)\int_{a^*} f(a)\,da \\ &= \left(p\bar{a}(a^*) - w\right)N(w,a^*)\end{aligned} \tag{51}$$

Where:

$$\bar{a}(a^*) = \frac{\int_{a^*} af(a)\,da}{\int_{a^*} f(a)\,da} \tag{52}$$

And:

$$N(w,a^*) = L(w)\int_{a^*} f(a)\,da = L(w)\left[1 - F(a^*)\right] \tag{53}$$

Now let us consider the two types of policy intervention. First, the Matsudaira type intervention. The firm is required to increase the amount of employment it has. It needs to choose (w,a*) to solve:

$$\max\left(p\bar{a}(a^*) - w\right) \quad s.t. \quad L(w)\left[1 - F(a^*)\right] = N \tag{54}$$

The first-order conditions for this can be written as:

$$-1 + \mu L'(w)\left[1 - F(a^*)\right] = 0 \tag{55}$$

$$p\bar{a}'(a^*) - \mu L(w)\,f(a^*) = 0 \tag{56}$$

Collecting these leads to:

$$w = \varepsilon p\left[\bar{a}(a^*) - a^*\right] \tag{57}$$

Where ε is the elasticity of the labour supply curve which, to keep things simple we will assume is constant. (57) gives a relationship between w and a*.

Now consider a change in N. we will have:

$$\frac{L'(w)}{L(w)}\frac{\partial w}{\partial logN} - \frac{f(a^*)}{1 - F(a^*)}\frac{\partial a^*}{\partial w}\frac{\partial w}{\partial logN} = 1 \tag{58}$$

Which can be written as:

$$\frac{\partial \log w}{\partial logN} = \frac{1}{\varepsilon - \frac{f(a^*)}{1-F(a^*)} \frac{w}{\varepsilon p(\bar{a}'(a^*)-1)}} == \frac{1}{\varepsilon + \frac{\bar{a}'(a^*)}{(\bar{a}'(a^*)-1)}} \tag{59}$$

Note that in the case where a has an exponential distribution this implies that the wage w will not change as is found by Matsudaira. In this case:

$$\bar{a}(a^*) = a^* + \alpha \tag{60}$$

Now consider a forced change in the wage as examined by Staiger. The firm wants to maximize (51). This leads to the first-order condition for a* of:

$$p\bar{a}'(a^*)\left[1-F(a^*)\right] - f(a^*)\left(p\bar{a}(a^*) - w\right) = 0 \tag{61}$$

Which can be written as:

$$a^* = \frac{w}{p} \tag{62}$$

The first-order condition for w can be written as:

$$w = \frac{\varepsilon}{1+\varepsilon} p\bar{a}(a^*) \tag{63}$$

Now, consider a rise in the wage. We will have:

$$\frac{\partial \log N}{\partial logw} = \varepsilon - \frac{f(a^*)}{1-F(a^*)} \frac{\partial a^*}{\partial logw} = \varepsilon - \frac{a^* f(a^*)}{1-F(a^*)} \tag{64}$$

In the case with the exponenetial distribution and for a just-binding wage this becomes:

$$\frac{\partial \log N}{\partial logw} = 0 \tag{65}$$

Another alternative is an effort model then the profit can be written as:

$$(pa - w)N \tag{66}$$

And N=U(w)G(a) so iso-morphic to the quality model just described.

Results Equating Separation and Recruitment Elasticity

Proof of Result 1:
Simple differentiation of (21) leads to:

$$\varepsilon_s(w) = \frac{ws'(w)}{s(w)} = \frac{-\lambda \int g(x) \frac{x}{w} \phi'\left(\frac{x}{w}\right) dx}{\lambda \int g(x) \phi\left(\frac{x}{w}\right) dx} = \int g_s(x;w) \varepsilon_\phi\left(\frac{x}{w}\right) dx \tag{67}$$

Where $g_s(x;w)$ is given by:

$$g_s(x;w) = \frac{g(x)\phi\left(\frac{x}{w}\right)}{\int g(x')\phi\left(\frac{x'}{w}\right) dx'} \tag{68}$$

Proof of Result 2
Differentiation of (26) leads to:

$$\varepsilon_R(w) = \frac{wR'(w)}{R(w)} = \frac{\int f(x)N(x)\frac{w}{x}\phi'\left(\frac{w}{x}\right) dx}{\int f(x)N(x)\phi\left(\frac{w}{x}\right) dx} = \int g_R(x,w)\varepsilon_\phi\left(\frac{w}{x}\right) dx \tag{69}$$

Where:

$$g_R(x,w) = \frac{f(x)N(x)\phi\left(\frac{w}{x}\right)}{\int f(x')N(x')\phi\left(\frac{w}{x'}\right) dx'} \tag{70}$$

Proof of Result 3
Using (26) and the equilibrium condition that firms that pay w spend H(w) on recruitment (whatever that may be), one can write (28) as:

$$g_R(x,w) = \frac{f(x)N(x)\phi\left(\frac{w}{x}\right)\lambda\left(\frac{H(w)}{H}\right)^\beta}{R(w,H(w))} \tag{71}$$

Now use (23) and reverse the roles of x and w to give:

$$g_s(w;x) = \frac{g(w)\phi\left(\frac{w}{x}\right)}{\int g(x')\phi\left(\frac{x'}{x}\right) dx'} = \frac{\lambda f(w)\phi\left(\frac{w}{x}\right)\left(\frac{H(w)}{H}\right)^\beta}{s(x)} \tag{72}$$

Combining (71) and (72) one obtains:

$$g_R(x,w) = \frac{f(x)N(x)g_s(w,x)}{s(x)f(w)R(w,H(w))} = \frac{f(x)R(x,H(x))}{f(w)R(w,H(w))} g_s(w,x) \tag{73}$$

Or:

$$f(w)R\left(w,H(w)\right)g_R(x,w)=f(x)R\left(x,H(x)\right)g_s(w,x) \tag{74}$$

Now we have that:

$$\begin{aligned}\int f(w)R\left(w,H(w)\right)\varepsilon_R(w)\,dw &= \iint f(w)R\left(w,H(w)\right)g_R(x,w)\varepsilon_\phi\left(\frac{w}{x}\right)dxdw \\ &= \iint f(x)R\left(x,H(x)\right)g_s(w,x)\varepsilon_\phi\left(\frac{w}{x}\right)dxdw = \int f(x)R\left(x,H(x)\right)\varepsilon_s(x)\,dx\end{aligned} \tag{75}$$

So the recruit-weighted quit and recruitment elasticities must be equal.

Table 1: Self-reported Important Life Events in Past Year: UK Data

	All	Men	Women
Family	38	33	42
Employment	22	24	20
Nothing	20	22	18
Leisure	19	19	19
Education	13	11	15
Health	12	10	13
Consumption	9	9	8
Housing	8	7	9
Other	7	6	7
Financial	4	4	4

Source: British Houseold Panel Study

Table 2
Estimates of Hiring Costs

Study	Sample	Measure	Results
Oi (1962)	International Harvester, 1951		0.073 (all workers) 0.041 (common labourers)
Barron, and Bishop (1985)	US firms, 1982	Hours spent recruiting, screening and interviewing applicants for one hire	9.87 hours
Manning (2006)	British firms	Recruitment and Training Costs	2.4% (unskilled) 4.5% (others) 11.2% (sales)
Abowd and Kramarz (2003), Kramarz and Michaud (2009)	French firms, 2002	Includes training and external hiring costs; excludes internal hiring costs	0.028
Blatter, Muhlemann and Schenker (2009)	Skilled workers with vocational degree in Swiss firms, 2000, 2004	Costs of recruitment and initial training	0.033

Table 3
Estimates of Job search by unemployed workers

[to come]

Table 4
Quasi-Experimental estimates of Wage Elasticity of Supply to Individual Employer

Study	Sample	'Experiment'	Outcome Variable	Estimated Elasticity
Staiger, Spetz and Phibbs (2010)	Veteran Affairs Hospitals	Permanent Rise in Wages where recruitment difficulties	Employment Rise 1 year later	0.1
Falch (2010)	Norwegian schools	Wage Premium at schools with recruitment difficulties	Contemporaneous employment	1.0-1.9
Clotfelter, Glennie, Ladd and Vigdor (2008)	Maths.science, special education teachers in selected North Carolina schools	Annual bonus – meant to be permanent but perhaps perceived as temporary	Turnover Rates	
Matsudaira	Californian	Increase in required minimum staffing levels	Change in wages	0

Table 5
Estimates of Separation Elasticities

[to come]

Table 6
Estimates of Gender Differences in Separation Elasticities

[to come]

Figure 1
The Textbook Model of Monopsony

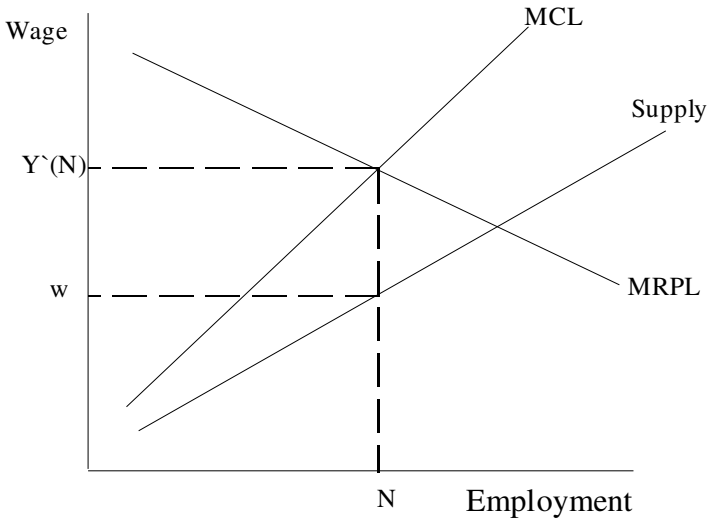


Figure 2

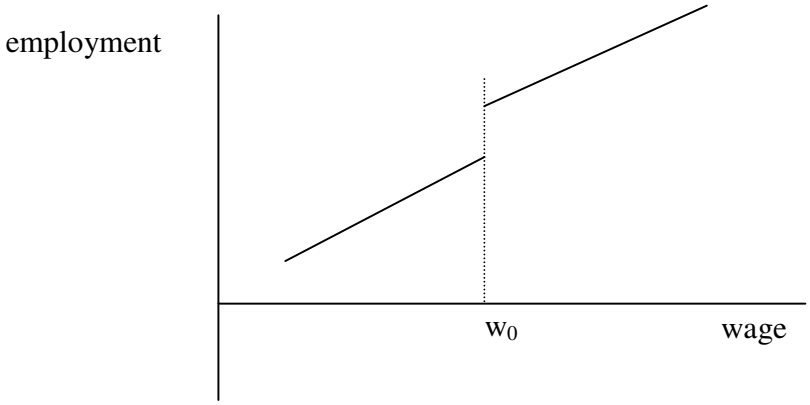
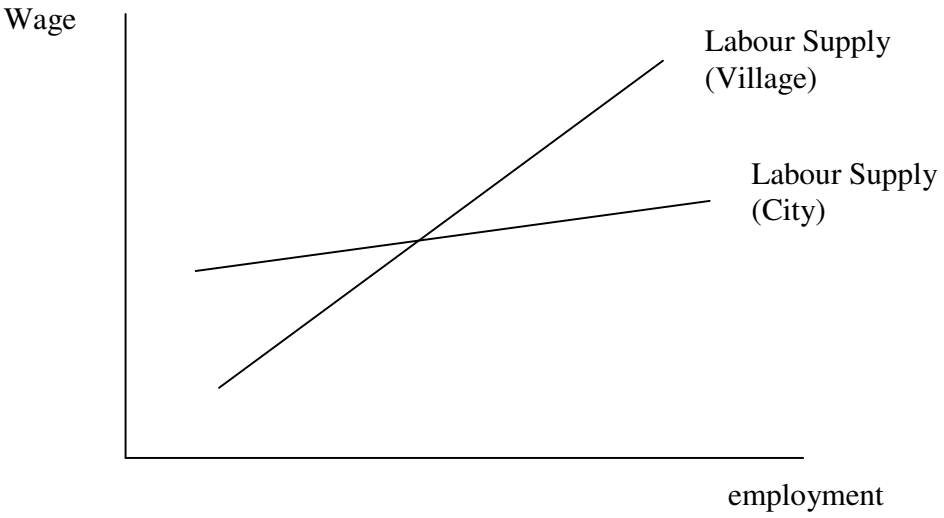


Figure 3
City and Village with A Monopsonistic Labour Market



Bibliography

Abowd, John, M. and Thomas Lemieux, 1993, "The Effects of Product Market Competition on Collective Bargaining Agreements: The Case of Foreign Competition in Canada", *Quarterly Journal of Economics*, 108(4), November, pages 983-1014.

Acemoglu, Daron, 1998, "Training and Innovation in an Imperfect Labour Market", *Review of Economic Studies*, 64, 445-464.

Acemoglu, Daron and Jorn-Steffen Pischke, 1998, "Why Do Firms Train? Theory and Evidence", *Quarterly Journal of Economics*; 113(1), February, pages 79-119.

Acemoglu, Daron. and Jorn-Steffen Pischke, 1999a, "Beyond Becker: Training in Imperfect Labour Markets", *Economic Journal*; 109(453), February pages 112-42.

Acemoglu, Daron and Jorn-Steffen Pischke, 1999b, "The Structure of Wages and Investment in General Training", *Journal of Political Economy*, 107(3), June, pages 539-72.

Acemoglu, Daron and Robert Shimer 2000 "Wage and Technology Dispersion", *Review-of-Economic-Studies*; Vol. 67, (October), pages 585-607.

Albrecht, Janes Pieter A. Gautier & Susan Vroman, 2006. "Equilibrium Directed Search with Multiple Applications," *Review of Economic Studies*, 73, 869-891.

Barth, Erling; and Harald Dale-Olsen, 1999, "Monopsonistic Discrimination and the Gender Wage Gap", National Bureau of Economic Research Working Paper 7197, June 1999.

Becker, Gary S., 1993, *Human Capital: a Theoretical and Empirical Analysis, with Special Reference to Education*, 3rd ed. University of Chicago Press: Chicago.

Benson, Alan (2009) "Firm-Sponsored general Education and Mobility Frictions: Evidence from Hospital Sponsorship of Nursing Schools and Faculty", MIT Sloan School.

Marianne Bertrand, Claudia Goldin, and Lawrence F. Katz, 2009, "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors" Aug. 2009.

Bewley, Truman F., 1999, *Why wages don't fall during a recession*, Cambridge, Mass.: Harvard University Press.

Bhaskar, V. and Ted To, 1999, "Minimum Wages for Ronald McDonald Monopsonies: A Theory of Monopsonistic Competition", *Economic Journal*; 109(455), April, pages 190-203.

Blatter, Marc, Samuel Muhlemann and Samuel Schenker, 2008 "The Costs of Hiring Skilled Workers"

Blau, David M. and Philip, K. Robins, 1990, "Job Search Outcomes for the Employed and Unemployed", *Journal of Political Economy;* 98(3), June, pages 637-55.

Boal, William M., 1995, "Testing for Employer Monopsony in Turn-of-the-Century Coal Mining", *Rand Journal of Economics*; 26(3), Autumn, pages 519-36.

Boal, William, M. and Michael, R. Ransom, 1997, "Monopsony in the Labor Market" *Journal of Economic Literature*; 35(1), March, pages 86-112.

Booth, Alison and Pamela Katic (2009) "Estimating the Wage Elasticity of Labour Supply to a Firm: Is there Monopsony Down-under?", unpublished, Australian National University.

Bronfenbrenner, Martin, 1956, "Potential Monopsony in Labor Markets", *Industrial and Labor Relations Review*, vol. 9, pp577-588.

Brown, Charles, 1999, "Minimum Wages, Employment, and the Distribution of Income", in Ashenfelter, Orley and David Card (eds) *Handbook of Labor Economics*. Volume 3. Amsterdam: North-Holland, pages 2101-63.

Brown, Charles, and James Medoff, 1989, "The Employer Size-Wage Effect", *Journal of Political Economy*; 97(5), October, pages 1027-59.

Brown, Charles, James Hamilton and James Medoff, 1990, *Employers large and small*, Cambridge, Mass.: Harvard University Press.

J. David Brown, Julie L. Hotchkiss, and Myriam Quispe-Agnoli. "Undocumented Worker Employment and Firm Survivability ." Federal Reserve Bank of Atlanta Working Paper 2008-28 (December 2008).

Burdett, Kenneth, and Dale T. Mortensen, 1998, "Wage Differentials, Employer Size, and Unemployment", *International Economic Review*; 39(2), May pages 257-73.

Cahuc, Pierre, Fabien Postel-Vinay and Jean-Marc Robin (2006) "Wage Bargaining with On-the-Job Search: Theory and Evidence", Econometrica, 74, 323-364.

Card, David, Raj Chetty & Andrea Weber, 2007. "Cash-On-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market," *Quarterly Journal of Economics*, 122, 1511-1560.

Card, David E. and Alan B. Krueger, 1995, *Myth and measurement: The new economics of the minimum wage*, Princeton: Princeton University Press, 1995.

Clotfelter, Charles, Elizabeth Glennie, Helen Ladd and Jacob Vigdor, 2006, "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina", NBER Working Paper No. 12285.

Cotton and Tuttle 1986 meta-analysis pay significant in 29 out of 32 studies. A Meta-Analysis of Antecedents and Correlates of Employee Turnover: Update, Moderator... Griffeth et al. *Journal of Management.*2000; 26: 463-488

Coles, Melvyn G. and Eric Smith, 1998, "Marketplaces and Matching" *International Economic Review*; 39(1), February, pages 239-54.

DellaVigna, Stefano and M. Daniele Paserman , 2005, "Job Search and Impatience", *Journal of Labor Economics*, 23, 527-588.

Dube, Arindrajit, T. William Lester and Michael Reich., 2010 **"Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties"**, *Review of Economics and Statistics*, forthcoming.

Dunlop, John, 1957, "The Tasks of Contemporary Wage Theory", in George Taylor and Frank Pierson (eds) *New Concepts in Wage Determination*, McGraw-Hill, New York.

Ebrahimy, Ehsan and Robert Shimer (2009) "Stock-Flow Matching", University of Chicago.

Eckstein, Zvi and Kenneth I. Wolpin, 1990, "Estimating a Market Equilibrium Search Model from Panel Data on Individuals", *Econometrica*; 58(4), July, pages 783-808.

Eckstein, Zvi and Gerard van den Berg, 2006, "Empirical labor search: A survey", *Journal of Econometrics*, 136, 531-564.

Edgeworth, Francis Ysidro, 1932, *Mathematical psychics : an essay on the application of mathematics to the moral sciences*, London: London School of Economics and Political Science.

Ellingsen, Tore and Asa Rosen,, 2002, "Fixed or Flesible? Wage Setting in Search Equilibrium", *Economica*, forthcoming.

Falch, Torberg, 2001, "Estimating the elasticity of labor supply utilizing a quasi-natural experiment", Working Paper, Norwegian University of Science and Technology, Trondheim, October 2001.

Fallick, Bruce C. and Charles A. Fleischman, 2001, "The Importance of Employer-to-Employer Flows in the U.S. Labor Market", Working Paper, Federal Reserve Board, April 2001.

Fox, Jeremy, 2010, "Estimating the Employer Switching Costs and Wage Responses of Forward-Looking Engineers", *Journal of Labor Economics*, forthcoming.

Giuliano, Laura, 2009, "Minimum Wage effects on employment, substitution, and the teenage labor supply: evidence from personnel data".

Gregg, Paul and Barbara Petrongolo, 1997, "Random or Non-Random matching? Implications for the use of the UV curve as a measure of matching effectiveness", LSE CEP Discussion Paper, No. 348, June 1997.

Hall, Robert E. and Edward P. Lazear, 1984, "The Excess Sensitivity of Layoffs and Quits to Demand" *Journal of Labor Econ*omics; 2(2), April, pages 233-57.

Hamermesh, Daniel.S., 1993, *Labor demand*, Princeton: Princeton University Press.

Barry T. Hirsch and Edward J. Schumacher, 1995, "Monopsony Power and Relative Wages in the Labor Market for Nurses," *Journal of Health Economics*, 14, 443-76.

Hirsch, Boris, Tomas Schrank and Claus Schnabel, 2010, "Differences in Labor Supply to Monopsonistic Firms and the Gender Pay Gap: An Empirical Analysis Using Linked Employer-Employee Data from Germany", *Journal of Labor Economics*, forthcoming.

Barry T. Hirsch and Edward J. Schumacher, 2005 "Classic or New Monopsony? Searching for Evidence in Nursing Labor Markets," *Journal of Health Economics*, 24, 969-89.

Holzer, Harry J., 1987, "Job Search by Employed and Unemployed Youth" *Industrial and Labor Relations Review*; 40(4), July, pages 601-11.

Holzer, Harry J., 1994, "Job Vacancy Rates in the Firm: An Empirical Analysis" *Economica*; 61(241), February, pages 17-36.

Julie L. Hotchkiss and Myriam Quispe-Agnoli. "Employer Monopsony Power in the Labor Market for Undocumented Workers." Federal Reserve Bank of Atlanta Working Paper 2009-14a (revised June 2009).

Jacobson, Louis S., Robert J. LaLonde and Daniel S. Sullivan, 1993b, "Earnings Losses of Displaced Workers", *American Economic Review*; 83(4), September, pages 685-709.

Klepinger, Daniel H., Terry R. Johnson and Jutta M. Joesch, 2002, "Effects of Unemployment Insurance Work-Search Requirements: The Maryland Experiment", *Industrial and Labor Relations Review*, 56, 3-22.

Kramarz, Francis and Marie-Laure Michaud, 2009, "The shape of hiring and separation costs in France", *Labour Economics*, forthcoming.

Krueger, Alan B. And Robert E. Hall, 2008, "Wage Formation between Newly Hired Workers and Employers: Survey Evidence", NBER Working Paper w14329,September 2008.

Krueger, Alan B. and Andreas Mueller, 2008. "Job Search and Unemployment Insurance: New Evidence from Time Use Data," Working Papers1070, Princeton University, Department of Economics, Industrial Relations Section.

Lazear, Edward P.,2003, "Firm-Specific Human Capital: A Skill-Weights Approach", NBER Working Paper No. W9679

Lester, Richard A., 1946, "Wage Diversity and its Theoretical Implications", *The Review of Economic Statistics,* vol.28, No.3, Aug., pp.152-159.

Lester, Richard A., 1948, *Company Wage Policies: A Survey of Patterns and Experience*, Princeton University Press: Princeton.

Lester, Richard A., 1952, "A Range Theory of Wage Differentials", *Industrial and Labor Relations Review*, Vol. 5, pp483-500.

Lewis, Michael, 1989, *Liar's Poker*, Coronet Books: London.

Machin, Stephen J. and Alan Manning, 2004, "A Test of Competitive Labor Market Theory: The Wage Structure among Care Assistants in the South of England**"**, *Industrial and Labor Relations Review* ,57,. 371-385

Manning, Alan, 1996, "The Equal Pay Act as an Experiment to Test Theories of the Labour Market ", *Economica*; 63(250), May, pages 191-212.

Manning, Alan, 2000, "Pretty Vacant: Recruitment in Low-Wage Labour Markets" *Oxford Bulletin of Economics and Statistics*; vol 62, Special Issue Dec., pages 747-70.

Manning, Alan, 2003, *Monopsony in Motion: Imperfect Competition in Labor Markets*, Princeton: Princeton University Press.

Manning, Alan, 2003, "The Real Thin Theory: Monopsony in Modern Labour Markets", *Labour Economics*, 10, 105-131.

Manning, Alan, 2006, "A Generalised Model of Monopsony", *Economic Journal.*

Manning, Alan, 2009, "The Plant Size-Place Effect: Monopsony and Agglomeration", forthcoming *Journal of Economic Geography.*

Manning, Alan, 2009 "A Little Collusion can go a long way"

Marshall, Alfred, 1920, *Principles of Economics: an Introductory Volume*, 8th ed. Macmillan: London.

Moen, Espen R., 1997, "Competitive Search Equilibrium", *Journal of Political Economy*; 105(2), April, pages 385-411.

Mortensen, Dale T., 1986, "Job Search and Labor Market Analysis" in Ashenfelter, Orley and Richard Layard (eds.) *Handbook of labor economics*, Volume 2, Amsterdam: North-Holland, pages 849-919.

Mortensen, D.T., 2003, *Wage Dispersion: Why Are Similar Workers Paid Differently?* MIT Press, Cambridge Mass..

Mortensen, Dale T. and Christopher A. Pissarides, 1994, "Job Creation and Job Destruction in the Theory of Unemployment", *Review of Economic Studies*; 61(3), July, pages 397-415.

Mortensen, Dale T. and Christopher A. Pissarides, 1999, "New Developments in Models of Search in the Labor Market", in Ashenfelter, Orley and David Card (eds) *Handbook of Labor Economics*. Volume 3B, Amsterdam: North-Holland, pages 2567-2627.

Nagypal, Eva (2005) "On the extent of job-to-job transitions",

Oi, Walter (1962), "Labor as a Quasi-Fixed Factor", *Journal of Political Economy*, 70, 538-555.

Oi, Walter Y. and Todd L. Idson, 1999, "Firm Size and Wages", in Ashenfelter, Orley and David Card (eds) *Handbook of Labor Economics*. Volume 3B, Amsterdam: North-Holland, pages 2165-2214.

Overman, Henry G. and Diego Puga, 2009, "Labor Pooling as a source of agglomeration: an empirical investigation", LSE.

Parsons, Donald O., 1972, "Specific Human Capital: An Application to Quit Rates and Layoff Rates", *Journal of Political Economy*; 80(6), Nov.-Dec., pages 1120-43.

Parsons, Donald O., 1973, "Quit Rates Over Time: A Search and Information Approach", *American Economic Review*; 63(3), June, pages 390-401.

Pissarides, Christopher.A., 1985, "Short-run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages", *American Economic Review*; 75(4), September, pages 676-90.

Postel-Vinay, Fabien and Jean-Marc Robin, 2002, "Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica* forthcoming.

Priest, George (2010) "Timing 'Disturbances' in Labor Market Contracting: Professor Roth and the effects of Labor Market Monopsony", Journal of Labor Economics, forthcoming.

Ransom, Michael and Ronald Oaxaca (2010) "Sex Differences in Pay in a 'New Monopsony' Model of the Labor Market", Journal of Labor Economics, forthcoming.

Ransom, Michael and David Sims (2010) "Estimating the Firm's Labor Supply Curve in a 'New Monopsony' Framework: school teachers in Missori", Journal of Labor Economics, forthcoming.

Reynolds, Lloyd G., 1946a, "The Supply of Labor to the Firm", *Quarterly Journal of Economics*, Vol. 60, No. 3., May, pp. 390-411.

Reynolds, Lloyd G., 1946b, "Wage Differences in Local Labor Markets" *American Economic Review*, Vol. 36, No. 3., June, pp. 366-375.

Reynolds, Lloyd G., 1951, *The Structure of Labor Markets: Wages and Labor Mobility in Theory and*

Ridder, Geert and Gerard van den Berg (2003) "Measuring Labor Market Frictions: A Cross-Country Comparison", Journal of the European Economic Association, 1, 224-244.

Robinson, Joan, 1933, *The Economics of Imperfect Competititon*, Macmillan: London.

Rogerson, Richard, Robert Shimer

Rogerson, Richard, Robert Shimer & Randall Wright, 2005. "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, 43, 959-988.

Roth, Alvin, 1990, "

Royalty, Anne Beeson, 1998, "Job-to-Job and Job-to-Nonemployment Turnover by Gender and Education Level", *Journal of Labor Economics*; 16(2), April, pages 392-443.
Sicherman, Nachum, 1996, "Gender Differences in Departures from a Large Firm" *Industrial and Labor Relations Review*; 49(3), April, pages 484-505.

Shimer, Robert (2006) "On-the-Job Search and strategic bargaining", European Economic Review, 50, 811-830.

Slichter, Sumner H., 1950, "Notes on the Structure of Wages", *Review of Economic Statistics*, Vol.32, No. 1, Feb., pp.80-91.

Smith, Adam, 1970, "*The Wealth of Nations*" London: Harmondsworth ,Penguin.

Staiger, Douglas, Joanne Spetz and Ciaran Phibbs, 2010, "Is There Monpsony in the Labor Market? Evidence from a Natural Experiment", *Journal of Labor* Economics, forthcoming.

Stigler, George J., 1961, “The Economics of Information”, *The Journal of Political Economy*, Vol. 69, No. 3. June, pp. 213-225.

Stigler, George J., 1962, “Information in the Labor Market”, *The Journal of Political Economy*, Vol. 70, No. 5, Oct., pp. 94-105.

Sullivan, Daniel, 1989, “Monopsony Power in the Market for Nurses”, *Journal of Law and Economics;* 32(2), Part 2, October, pages S135-78.

van den Berg, Gerard J. and Geert Ridder, 1998, “An Empirical Equilibrium Search Model of the Labor Market”, *Econometrica*; 66(5), September, pages 1183-1221.

Van Ommeren, Jos, Gerard van den Berg and Cees Gorter, 1999, “Estimating the Marginal Willingness to Pay for Commutinh”, unpublished, Free University of Amsterdam.

Van Ours, Jan C. and Geert Ridder, 1992, “Vacancies and the Recruitment of New Employees” *Journal of Labor Economics*; 10(2), April, pages 138-55.

Von Wachter, Till, Jae Song and Joyce Manchester, 2009, "Long-Term Earnings Losses due to Mass-Layoffs During the 1982 Recession: An Analysis Using Longitudinal Administrative Data from 1974 to 2004"

Viscusi, W. Kip, 1980, “Sex Differences in Worker Quitting”, *Review of Economics and Statistics;* 62(3), Aug., pages 388-98.

Weiss, Andrew W., 1984, “Determinants of Quit Behavior”, *Journal of Labor Economics*; 2(3), July, pages 371-87.

Weiss, Andrew W. and Henry, J. Landau, 1984, “Wages, Hiring Standards, and Firm Size” *Journal of Labor Economics*; 2(4), October, pages 477-99.