Monopsonistic discrimination, worker turnover, and the gender wage gap

Erling Barth ¹, Harald Dale-Olsen ²

Institute for Social Research, P.O. Box 3233 Elisenberg, N-0208 Oslo, Norway

A R T I C L E  I N F O

Article history:
Received 9 August 2007
Received in revised form 9 January 2009
Accepted 10 February 2009
Available online 28 February 2009

JEL classification:
J16
J31
J42
J63
J71

Keywords:
Gender wage gap
Monopsony

A B S T R A C T

Motivated by models of worker flows, we argue in this paper that monopsonistic discrimination may be a substantial factor behind the overall gender wage gap. Using matched employer–employee data from Norway, we estimate establishment-specific wage premiums separately for men and women, conditioning on fixed individual effects. Regressions of worker turnover on the wage premium identify less wage elastic labour supply facing each establishment of women than that of men. Workforce gender composition is strongly related to employers’ wage policies. The results suggest that 70–90% of the gender wage gap for low-educated workers may be attributed to differences in labour market frictions between men and women, while the similar figures for high-educated workers ranges from 20 to 70%.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Modern labour economics does not give the theory of monopsonistic discrimination much credit as an explanation of gender differences in pay. In this paper, however, we identify gender differences in worker turnover patterns that provide employers with incentives to pay men and women differently. We show that the labour supply of men facing each establishment is more elastic than the labour supply of women.

Joan Robinson (1933) developed the idea of monopsonistic discrimination in the labour market. The idea is simple: a single buyer, a monopsonist, sets wages below marginal revenue product. The more inelastic the labour supply, the lower are wages relative to productivity. By differentiating wages between groups with different elasticities of labour supply, the monopsonist may obtain higher profits. Robinson suggests gender as one of the dimensions along which the employer may discriminate. If female labour supply is more inelastic than male labour supply, women will earn less than men relative to their productivity, and thus face a higher level of exploitation in the labour market.

While some works argue in favour of monopsonistic discrimination, the general consensus now seems to be that this model does not add much to the understanding of the overall gender wage gap. This is true on both sides of the Atlantic: Jane Humphries (1995) writes “But this classic case (pure monopsony) seems to have little empirical purchase” ¹, in the theoretical chapter of The Economics of Equal Opportunities, edited by herself and Jill Rubery. Blau, Ferber and Winkler (1998) write in a footnote “It seems likely.... that the monopsony explanation is more applicable to specific occupations and specific labour markets than to the aggregate gender pay differential.” The model is refuted because single buyer situations are rare, but most

² An earlier version of this paper has been circulated as NBER working paper # 7197. The empirical part of the paper is, however, completely redone with panel data and more comprehensive methods. We thank Monica Galizzi, Claudia Goldin, the editor and the two anonymous referees for helpful and constructive comments and suggestions. We acknowledge financial support from the Norwegian Research Council, grants No. 124582/510 and No. 173951/S20. Erling Barth thanks Harvard University and NBER for their hospitality during parts of this work. This work has also benefited from the research environment at ESOP, Department of Economics, University of Oslo.

3 She does, however, add, “women are more constrained than men in choice of occupation and job opportunities of married women react negatively to spatial monopsony”. For both, see Madden (1973). Also, several studies report evidence consistent with such behaviour in particular labour markets (Ferber et al., 1978; Rooton and Lane, 1985; Ransom, 1993; Bratsberg et al., 2003). Winkler (1998) finds that wages and job opportunities of married women react negatively to spatial monopsony indicators.

4 Winter-Ebmer (1995) finds that wages of high-educated workers ranges from 20 to 70%.
More recent theoretical developments have revitalised the concept of monopsony in the labour market. Among the theoretical works, the analyses of job-to-job flows within a search theoretic framework by Burdett and Mortensen (1998) and Manning (2003) have established the idea that each single firm or establishment faces its own individual labour supply curve. The point is that workers quit endogenously, and have to be replaced by new hires. The higher the wage, the fewer the quits and also the easier it is to attract replacement hires. We analyse Robinson’s idea of monopsonistic discrimination within a modern model framework based on the dynamics of labour supply to each firm.

In the dynamic monopsony model, differences in the exogenous quit rate or in the probability of receiving a job offer produce incentives for monopsonistic discrimination.6

Several conditions have to be met in order for the model of monopsonistic discrimination to work. One is that employers should be able to distinguish between men and women in the wage setting process. We argue that even in the absence of pure wage discrimination – unequal wages for equal work – employers may distinguish between jobs with uneven gender composition.5 Lucifora and Reilly (1990) show that female-dominated occupations pay less than male-dominated occupations. Meyerson-Milgrom et al. (2001) conclude that there are very small wage differences between men and women within finely defined job-cells in the same establishment. Gender differences arise across jobs or occupations and establishments.

Next, the labour supply curve of women has to be less elastic than the labour supply curve of men. This is the very point on which the model of monopsonistic discrimination has been scrapped. It seems that female labour supply is equally, or more, wage sensitive than men’s labour supply. However, this observation is made at the margin of the labour force; i.e. on the participation decision of men and women. An important point for our study is that even if the aggregate labour supply of women is more wage sensitive than the aggregate labour supply of men, the labour supply of women facing each establishment may be less wage sensitive than the labour supply of men facing each establishment. The main reason is that the labour supply facing each establishment also depends crucially on job-to-job search by employees in own and other establishments. The burden of proof then shifts from participation decisions to turnover behaviour.

However, several studies reveal that women’s turnover is actually similar to that of men’s, once appropriate control is included (e.g. Blau and Kahn, 1981; Galizzi, 2001; Viscusi, 1979). Even in Manning’s book on monopsony, he writes about gender discrimination in the labour market and gender differences in the elasticity of worker turnover with respect to wages that “the gender differences that we have identified in previous sections do not show up in these estimated elasticities. Whether this is because this approach to estimating elasticities is not very informative or because the total effect of the gender differences in constraints and motivation is small, is an issue that deserves further consideration” (Manning, 2003:208).

The main empirical contribution of our paper is to show that once we rinse the wage measure used in turnover regressions for worker differences in qualification and outside options, the estimated labour supply of women facing each employer is less wage elastic than the supply of men.7 Thus we provide strong support for the idea that employers have an incentive to apply monopsonistic discrimination against women in their wage policies. We use establishment fixed effects as measures of the establishments’ wage policies, and distinguish between demand and supply effects by using instruments for the establishments’ wage policy. Next, we show that the establishments’ gender composition is closely related to the establishments’ wage policies. Finally, we estimate the amount of the observed gender wage gap that may be attributed to frictions and monopsonistic discrimination.

Our study is not, however, the only one providing evidence supporting the model of monopsonistic discrimination. Using data on high school and college graduates, Bowlius (1997) identifies higher labour market frictions for women than for men. Her study was the first to apply an equilibrium search model on gender wage differentials. Bowlius finds that the differences in search parameters explain 20–30% of overall male–female wage differentials of high school and college graduates.

Related evidence has also been presented by Green, Machin and Manning (1996) and Manning (1996). Green et al. (1996) identify higher size-wage effects for women than for men, an observation which is consistent with a model of monopsonistic discrimination in the labour market. Manning (1996) analyses relative female employment following from the large rise in the relative earnings of women in the UK after the Equal Pay act of 1970 was passed. He attributes the observation that female relative employment did not fall, to monopsony in the female labour market. Differences in turnover behaviour between men and women are identified in several studies. Loprest (1993) finds that young women have on average less than 50% of the wage growth of young men when changing jobs. Sicherman (1996) finds that, at low levels of tenure, women have higher rates of departures than men do, but as tenure rises, women were less likely than men to leave the firm. The evidences of Pissarides and Wadsworth (1994) (UK) and Keith and McWilliams (1999) (US) suggest that there are gender differences in search behaviour and job-to-job search intensity.

The paper is structured as follows. Section 2 presents a theoretical model of monopsonistic discrimination. Section 3 presents the empirical specification, while Section 4 describes the data. Section 5 reports results from separation and excess turnover regressions. In Section 6 we analyse the relationship between gender workforce composition and wage differentials between men and women, as well as establishment-specific wage premiums for each gender. Section 7 studies the importance of search frictions and monopsonistic discrimination for the gender wage gap. Section 8 concludes the paper.

2. A theory of monopsonistic discrimination

In this section we develop a model of monopsonistic discrimination based on the standard models of job-to-job search and equilibrium wage distribution of the Burdett and Mortensen (1998) and Manning (2003) type. We consider an economy consisting of two labour inputs, \( j = 1,2 \), and where the employers may freely set wages for each type of labour. We assume that the two types of labour operate in completely segregated labour markets, which means that they draw wage offers from separate distributions. Under this assumption, we may think of the two groups as men and women, or

---


5. Green, Machin and Manning (1996) show in a dynamic monopsony model that the elasticity of wages with respect to employer size is increasing in the ratio of exogenous quits to the arrival rate of jobs. Both Black (1995) and Bowlius and Eckstein (2002) develop equilibrium search models associating discrimination with the presence of a discrimination factor on the part of employers. Particularly, Bowlius and Eckstein use a similar model framework as ours, where in their model, differences in job arrival rates follow from employers’ discrimination factor.

6. Employers hardly employ a separate wage policy for each gender. It is not legal for a firm to pay men and women differently within a job. It is hardly the practice to do so either (see Meyerson-Milgrom et al., 2001). As suggested by a referee, employers may also discriminate by creating different job titles when they hire women or men. We expect such behaviour to be constrained by other determinants of the job structure within establishments as well as considerations related to the gender of future replacement hires.

7. We use the terms firm, employer and establishment interchangeably throughout. In the empirical analysis, an establishment is defined by an unique employer and location identification (see Section 3 for a discussion).
more realistically, as two types of occupational groups with no inter-occupational mobility in the short term.\footnote{This is obvious a simplification, but note that most workers do not reenter the educational system in Norway (in our data less than 10%) and OECD (1998) ranks Norway as number 2 as far as occupational segregation in the OECD countries is concerned, while Great Britain ranks 4.}

Assuming that there are some frictions in the labour market, the number of employees of type j the employer may hire in a given period of time, is an increasing function of the wage $H^j(w^j)$, $H^j(w^j) > 0$, while the fraction of the employer’s stock of employees that leaves the firm over the same period is a decreasing function of the wage $q^j(w^j)$, $q^j(w^j) < 0$. While quits are assumed proportional to the number of employees, the number of hires is, by an assumption of random matching (Burdett and Vishwanath, 1988) a function of w and independent of $L^j$ (the employer’s stock of employees of type j). In steady state, $L^j$ is constant and the labour supply to any one firm is given by $L^j(w^j) = H^j(w^j)/q^j(w^j)$.

Let $\lambda^j$ be the probability that an employee of type j receives a job offer. Let $F_i(w^j)$ be the endogenously determined wage offer cumulative distribution function (cdf). $\lambda^j$ is an exogenous separation rate. The probability that an employee of type j separates is then given by $q^j(w^j) = \lambda^j(1 - F_i(w^j))$. Consider next the hiring function. Let the probability that an unemployed worker receives a wage offer be $\lambda^j$ as well. We have $F_i(b) = 0$, where b is the common reservation wage of the unemployed workers. Unemployed workers who receive a wage offer accept. In addition, the firm hires from employed workers who earn less than w. The cdf of workers over wages is given by $G_i(w^j)$, and may deviate from the distribution of $f_i$ facing the employer due to the number of workers per firm may differ. Thus, $H^j(w^j) = \lambda^j \int_0^w \lambda^j \sum_{j^1 \neq j} q_j^j(w^j) G_j(w^j) + \lambda^j \sum_{j^1 \neq j} q_j^j(w^j) G_j(w^j)$, where $U^j$ is the number of unemployed workers, $N^j$ is the labour force and M is the number of firms in the economy. The steady state unemployment rate is given by $\delta/(\delta + \lambda^j)$.

Equating the flows in and out of the wage distribution, and using the condition of a steady state unemployment rate, we obtain the labour supply of group j facing each firm which is given by:

$$L^j(w^j) = \frac{\delta \lambda^j}{\delta^j + \lambda^j(1 - F_i(w^j))} \frac{N^j}{M}, \quad j = 1, 2. \tag{1}$$

The profit of an employer is given by $R[L^1(w^j), L^2(w^j)] = w^j L^1(w^j) - w^j L^2(w^j)$, where $R[.]$ is revenue. We assume that each labour input has the same marginal revenue product, $p$, and is independent in production. The first order conditions for the wage-setting employers are:

$$\left( p - w^j \right) \frac{\partial L^j}{\partial w^j} - b^j = 0, \quad j = 1, 2. \tag{2}$$

Eq. (2) readily translates into the standard expression of monopsonistic exploitation for each group of workers:

$$p - w^j = \frac{1}{\epsilon^j}, \quad j = 1, 2, \tag{3}$$

where $\epsilon^j$ is the elasticity of labour supply of type j facing a firm. Thus wages are set so that the marginal value product equals wages multiplied a mark-up, $\omega^j$, which depends on the elasticity of labour supply of type j facing a firm, i.e., as expressed by Eq. (4):

$$p = \frac{1 + \epsilon^j}{\epsilon^j} w^j = \omega^j w^j. \tag{4}$$

Transformation of Eq. (4) gives the wage gap between workers of the two types as:

$$\frac{w^2 - w^1}{w^1} = \frac{\epsilon^2 - \epsilon^1}{\epsilon^1 (\epsilon^2 + 1)} = \frac{\omega^2}{\omega^1} - 1, \tag{5}$$

which increases in the ratio of the elasticity of labour supply of group 2 relative to the elasticity of labour supply of group 1.

The crucial empirical prediction from the model of monopsonistic discrimination is that the labour supply of women facing each firm is less wage elastic than that of men. This is what provides the employer with an incentive to discriminate between the two groups, and this is exactly the feature that we examine and test in the empirical part of the paper.

Consider next the relationship between the labour supply facing each establishment and aggregate labour supply. The labour supply of one type of labour facing establishment f may be written $L^j_f = b^j(w^j) N^j$, where $b^j(w^j) = \frac{\delta^j + \lambda^j}{\delta^j + \lambda^j} \sum_{j^1 \neq j} \frac{q_j^j(w^j)}{G_j(w^j)}$ is establishment f’s share of aggregate labour supply (see Eq. (1)). Traditional labour supply models analyse the participation decision, or the relationship between aggregate labour supply, $N$, and the expected wage, $\omega^j$. They have found women’s labour market participation to be more wage elastic than that of men.\footnote{This is for example the implication found in a recent contribution studying the impact of taxation on the labour supply of men and women (husbands and wives) in Sweden, where exogenous tax variation cause variation in earnings (Gelman, 2008).}

Manning (2003) incorporates aggregate labour supply into the monopsony model. He does so by allowing for heterogeneity in reservation wages and shows that $L^j_f = b^j(w^j) \Gamma_f(w^j) N$, where $\Gamma_f(.)$ is the cdf of reservation wages.\footnote{See the appendix to Section 3.4 in Manning (2003).} The elasticity of labour supply facing each firm equals: $\epsilon^j = \epsilon^j + \omega^j \epsilon^j$, where $\epsilon^j$ is the elasticity of $\Gamma_f$ arising from the firms’ need to accommodate turnover in a market with frictions (i.e. from the standard monopsony model as analysed above), and $\omega^j \epsilon^j$ is the elasticity of $\Gamma_f$ arising from the participation decision (i.e. aggregate labour supply). It is thus an empirical question to what extent the labour supply of women facing each firm is more or less wage elastic than the labour supply of men facing each firm, even if the elasticity of participation is higher for women. What matters for the wage policy of the firm is the labour supply facing each establishment, and if the labour market frictions are sufficiently larger for women than for men, the monopsony model may still provide the explanation of the gender wage gap even if aggregate supply of women is more wage elastic. The purpose of the empirical part of this paper is thus to provide an answer to this question: Are there significant differences between the labour supply elasticities of men and women facing each establishment?

3. Empirical specification

From Eq. (1), the elasticity of labour supply of type j facing a firm may be written:

$$\epsilon^j = \frac{w^j}{U^j(w^j)} \frac{\partial L^j(w^j)}{\partial w^j} = -2 \frac{w^j}{q^j(w^j)} \frac{\partial q^j(w^j)}{\partial w^j}, \quad j = 1, 2, \tag{6}$$

where $q^j(w^j)$ denotes the separation rate defined previously in Section 2. Thus, examining the elasticity of labour supply with respect to wages is equivalent to examining the elasticity of the steady state separation rate with respect to wages. This implies that what we need in our econometric analysis, are measures of the employer’s wage policy and of the steady-state separation rate. These measures are described in the next sections, where we for expositional simplicity suppress the superscript for the type of worker.

3.1. The employer’s wage policy

Let the expected alternative wage of an employee i outside his current establishment f at time t, $w^i_{t, \Delta}$, be given by:

$$\ln w^i_{t, \Delta} = \gamma_i + x_{it} \beta + \theta_i, \tag{7}$$

where $x_{it}$ are covariates describing time-varying observable productivity characteristics such as experience and establishment-specific seniority,
and $\theta_i$ is an individual productivity factor (including both unobservable and observable characteristics). The establishment-specific wage premium or wage policy is modelled as follows. We assume the employer chooses a wage policy $c_f$ which produces the wage:

$$\ln w_{if} = \gamma + \ln w_{if}^* + c_f + u_f,$$

for individual $i$ at time $t$, $\gamma$ describes time dummies. $c_f$ expresses an establishment-specific fixed effect, which defines the establishment-specific wage premium of establishment $f$ relative to the alternative wage of its workers. $\ln w_{if}^*$ is an error term with standard properties. Our measure of the wage policy of the firm is completely rinsed of all wage effects associated with the individual’s level of human capital. As is standard with all fixed individual and firm-models, potential match-specific wage effects are disregarded.  

We estimate Eq. (8) as follows: First we introduce a vector of establishment dummies. Then we conduct the ordinary panel data within-individual transformation of all covariates (including the establishment dummies) as well as of $\log(\text{wage})$. Running OLS for each group of workers on this transformed equation then produces unbiased efficient estimates of the establishment dummies and $\beta$ (see Table A1 in the Appendix for the results). $c_f$ is then normalised by taking the deviation from the employment-weighted mean of the estimated establishment dummies across all establishments. This way of modelling the wage policy of the firm, produces the following relationship between the elasticity of labour supply facing a specific employer $f$ and the employer-specific separation rate:

$$c_f = -2 \left( \frac{1}{q_f} \frac{\partial q_f(w_f)}{\partial \ln w_f} \right) = -2 \frac{\partial \log q_f(w_f)}{\partial \ln w_f},$$

according to Eq. (6) and the semi-logarithmic specification of the wage equation. Our turnover regressions follow from Eq. (9).

### 3.2. Steady-state separation rate

In steady-state where establishment-size is kept constant, each separation has to be met by a replacement hire. In this case, the separation rate is equal to the turnover rate. However, in our data some establishments grow, while other establishments decline. We use two empirical measures of the steady-state separation rate—the empirical separation rate and an excess turnover rate. The separation rate, $s$, is constructed by first defining a dummy-variable at the individual level, $S = 1$ if an employee separates, otherwise $S = 0$, and then estimate the establishment-specific separation rate by taking the establishment-specific mean. The excess turnover rate corrects for the fact that growth (decline) of an establishment also induces hires (separations). Excess turnover (et) is defined as $et = \min[h,s]$, where $h$ and $s$ denote starts ($h$) and separations ($s$) during the period. Excess turnover is thus the part of worker turnover needed in a stationary environment for each firm to keep its stock of workers constant. The excess turnover rate is then defined by dividing the excess turnover by the average stock of workers during the period. Both the separation rate and excess turnover rate are calculated separately for four groups of workers (gender/level of educational attainment) (see Section 4).

We use establishments as our unit of observation. This implies that we cannot disentangle the movement of workers between establishments within the same firm from movements of workers between firms. If the personnel and wage policies of the firms are establishment specific, this poses no problem for us. If the wage policies of firms are firm specific, on the other hand, the movements of workers across plants within a firm may create some noise in our data. However, a majority of the Norwegian firms consists of a single establishment only, and movements of workers between establishments within multi-establishment comprise a relatively small part of the aggregate separation rate.

Finally, one may argue that worker turnover and the wage policy of an employer are jointly determined. In some of the turnover regressions we therefore instrument the wage policy by capital (see Section 6 for discussion and details).

### 4. The data

Our data comprise all establishments and employees in the central south-east region of Norway, covering about 45% of all employees in Norway during the period 1989–97. They are based on information from the Current System of Social Data (CSSD). CSSD consists of several public administrative registers linked together into an integrated data system, which is managed by Statistics Norway. However, we restrict our analysis to individuals 20–60 years of age working in what may be considered the same geographical labour market. We focus on four groups of individuals, depending on their gender and their level of educational attainment, whom are working in the central south-east area of Norway (comprising e.g., the capital Oslo).  

Our four educational groups are: I) Women – low education, II) Men – low education, III) Women – high education (college or university) and IV) Men – high education (college or university). Unskilled labour, vocational training and high school (highest educational attainment less than five years education in excess of 9-years of compulsory schooling) are denoted as low education.

Wages are measured as daily earnings. Our earnings measure comprises ordinary wages and all taxable fringe benefits reported to (and evaluated by) the Norwegian tax authorities. To avoid observations of outliers, we have estimated a simple OLS log wage regression controlling for standard human capital variables. We then discard observations where the residual from the regression is outside $+/-5$ standard errors.

The identification of fixed worker and establishments effects rests on observations of movers (workers that change employers during our period of observation). The fixed worker and establishments

---

11 Potential match specific components will be distributed on the individual- or firm-specific effects according to the variations in the data set. For instance, for stayers a match specific component over and above the firm effect will be attributed to the individual, while for movers the distribution will depend on the other movers that identify any given firm effect.  

12 The transformation implies that a variable is subtracted the individual-specific average (over the observation period) of that variable.  

13 We use the term “separations”, since we do not know whether the employee quits voluntarily or if the employee was fired. This is not ideal, but common in many linked employer–employee data studies. However, voluntary quits are likely to dominate involuntary quits in Norway, at least partly due to the rather strict employment protection legislation (EPL). Furthermore, according to our model, displacements/ involuntary quits are not related to wages, so when we study empirically the relationship between separations and wages, we still measure the desired relationship. On the other hand, if involuntary quits and displacements are influenced by wages, for this to matter for our key results, the relationship between involuntary quits and wages for men and women have to be contrary to the relationship between voluntary quits and wages. Since EPL does not provide women with different rights or protection than men, there is little reason to think that the relationship between involuntary quits and wages differ between men and women.  

14 The excess turnover rate is equivalent to what Albæk and Sørensen (1998) define as the replacement hiring rate, or to half of the churning rate as defined by Burgess et al. (2000).  

15 One referee pointed out this potential shortcoming. Using similar Norwegian register data for the growth years of 1996 and 1997 comprising the complete population of private sector jobs, establishments and firms, we explored this issue. Of 118,405 private sector firms in 1996, 113,800 firms (96.1%) comprised one single establishment only. The average aggregate separation rate in 1996-97 was 27%, whereas 3.7 percentage points are caused by within-firm mobility, we thus regard this potential shortcoming of our data to be of little worry.  

16 All our analyses are conducted by only including observations with no missing variables.
Table 1
The elasticity of worker separation with respect to wages.

<table>
<thead>
<tr>
<th>MODEL:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low educated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All establishments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>-0.812**</td>
<td>-0.743**</td>
<td>-0.872**</td>
<td>-0.836**</td>
<td>-0.534**</td>
<td>-0.780**</td>
<td>-0.574**</td>
<td>-0.862**</td>
</tr>
<tr>
<td>Women</td>
<td>(0.032)</td>
<td>(0.028)</td>
<td>(0.031)</td>
<td>(0.029)</td>
<td>(0.043)</td>
<td>(0.038)</td>
<td>(0.046)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>High educated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>-0.069</td>
<td>-0.036</td>
<td>0.246**</td>
<td>0.288***</td>
<td>(0.043)</td>
<td>(0.057)</td>
<td>(0.064)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Women</td>
<td>(0.043)</td>
<td>(0.044)</td>
<td>(0.047)</td>
<td>(0.057)</td>
<td>(0.043)</td>
<td>(0.057)</td>
<td>(0.064)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Controls:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>County, industry</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Education</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>4709</td>
<td>4267</td>
<td>4709</td>
<td>4267</td>
<td>1867</td>
<td>1971</td>
<td>1867</td>
<td>1971</td>
</tr>
<tr>
<td>R²-adj.</td>
<td>0.319</td>
<td>0.335</td>
<td>0.349</td>
<td>0.366</td>
<td>0.264</td>
<td>0.319</td>
<td>0.362</td>
<td>0.353</td>
</tr>
</tbody>
</table>

Note: Unit of observation is establishments, where each observation expresses average across the period of observation. In the WLS-regressions, each observation is weighted by the number of years the establishment is observed. All regressions also include an intercept. Controls for 2-digit NACE industry and county are in the form of dummies. Workforce educational qualification is controlled for in the form of variables measuring proportion of workers having educational qualification within 8 specific fields and by variables expressing years of education in excess of compulsory schooling and years of education in excess of compulsory schooling squared. Standard errors reported in parentheses, ***, ** and * denote significant at 1, 5 and 10% level, respectively. Dependent variable: log separation rate.

Effects cannot be identified in establishments with no movers. Thus we discard all observations of establishments with less than five movers. The wage estimations were conducted separately for each group. The groups vary in size, from just below 100000 workers to more than 200000 workers. Since our wage equations are estimated separately for each group, they identify fixed group-specific establishment effects.

Establishments’ separation rates and excess turnover rates are constructed by using information about all employees of the given gender/educational group, who is employed on the 16th of May each of the years 1989–97. This date is chosen to maximise the quality of the register information, since an administrative correction procedure is conducted by the public authorities late winter each year. In order to avoid integer problems and small sample uncertainty, we discard in the group-specific turnover regressions all observations of establishments with less than 10 employees belonging to any specific group.

Since our analyses of the groups are mainly conducted separately for each group, the variables related to an establishment’s employees are usually constructed for employees belonging to one group only. The main reason for doing this is to focus on the group differences. In general, our establishments may very well employ a mixture of employees belonging to different groups.

5. Gender differences in the elasticity of worker turnover with respect to wages

A prerequisite for the theory of monopsonistic discrimination to provide a reasonable explanation for the gender wage gap is that worker turnover of women is less sensitive to wages than that of men’s worker turnover. Previous studies, such as for example Manning (2003), have not been able to find strong evidence for this. As a backdrop to our empirical analysis, we first show that given Manning’s empirical approach this appears true also in our data. We will later show that these results are basically caused by lack of appropriate control for individual productivity differences.

In Table 1 we repeat Manning’s analysis on the elasticity of worker turnover with respect to wages (Table 7, Manning, 2003). We have split our sample of establishments into four categories depending on the workforce’s level of education (low or high education) and gender. Then we run, separately for each group, weighted least squares regressions (WLS) of establishment-specific quit rates on the establishment-specific average wage, including controls for industry (2-digit SIC-code) and county.

Table 1 shows that for low educated workers, the estimated elasticity of worker turnover w.r.t. wages for both men and women is around −0.9. For high educated workers, however, the estimated elasticity of worker turnover w.r.t. wages for women is around −0.5. Thus for the high-educated workers, we even observe that women’s quits are significantly more sensitive to wages than that of men’s. This confirms Manning’s findings that “now the female elasticity is larger than the male in three of the data sets and the gap is actually significant.”

Evidently, this evidence appears to go against the theory of monopsonistic discrimination as an explanation of the gender wage gap. However, this exercise fails to distinguish between the parts of the wage that are specific to the establishment and the parts of the wage, like human capital, that also influence the outside option of the workers. Thus, following the description of Section 3, we now proceed by reporting results using the establishments wage policy, measured by the establishment specific wage premium relative to workers’ alternative wages, in the turnover regressions. Tables 2 and 3 reports the impact of the establishment specific wage premium on our two measures of quit rates: log separation rate and log excess turnover rate, respectively. According to Eq. (9) this yields directly empirical evidence on the elasticity of labour supply with respect to wages.

As before, we split our sample of establishments into four categories depending on the workforce’s level of education (low or high education) and gender, and run separately for each group, weighted least squares log quit regressions (WLS) on the establishment-specific wage premium and controls for workforce average years of education (and squared), proportion of workforce’s fields of educational qualification (we differentiate between 8 categories which basically expresses occupational categories), industry, and county. The models of Tables 2 and 3 are comparable to the models of Table 1. Models 1–4 of Table 4 repeat the analyses for low educated workers employed by establishments in the Manufacturing Statistics only.

Changing the focus from establishment average wages to establishment wage premium has a major impact on our conclusions regarding gender differences in the elasticity of labour supply. Regardless of how we measure quits and regardless of level of educational attainment, men’s turnover is more sensitive to wages than women’s turnover.

---

17 Information on the group composition (observations/wage earners/establishments): group I (114377/227927/6291), group II (1040826/216155/5434), group III (471345/95144/2380) and group IV (465769/98508/2423).

18 Each observation is weighted according to the number of years the establishment is observed.
While this difference is not significant when estimated for high educated workers and using separations as turnover measure, it is strongly significant for all groups using excess turnover as our measure. Since individual productivity differences are controlled for, our results are not caused by gender differences in human capital.

Since wages are set by profit maximizing employers, the wage premium may be considered endogenous. In the regressions of Table 4, we take this into account by introducing instruments for the wage premium. Unfortunately, our analyses are in this case limited to workplaces included in the Statistics Norway’s Manufacturing Statistics and then for low-educated workers only.19

Table 4 presents four sets of the result. Models 1–2 and 3–4 of Table 4 present the WLS results for the separation rate and the excess turnover rate, respectively, on this sample for comparison. Compared to Tables 2 and 3 we see that focusing on these establishments employing low-educated workforces implies more wage-sensitive turnover for men and less wage-sensitive turnover for women. Models 5–6 and 7–8 of Table 4 then present the similar regressions using weighted IV-regressions. We use log capital related to machinery and inventory and capital related to buildings (in millions) as instruments, since there is a rich literature arguing and identifying that more capital intensive firms pay higher wages.20 Due to the quasi-fixed nature of capital, these capital measures may be considered appropriate instruments for wages. As seen in Table 4 our instruments are strong and they also satisfy the Sargan-test.

When we instrument the wage premiums using these instruments, Models 5–8, our results are even stronger, i.e. men’s turnover turns strong and they also satisfy the Sargan-test.

19 Information on our instruments is limited to workplaces participating in Statistics Norway’s Manufacturing Statistics. This comprises mainly the manufacturing industries, but also mining, construction, service and trade workplaces are found. However, the number of establishments in our sample with a sufficient number of high educated workers of each gender moving between establishments in our panel is too small to identify effects for the high educated workers.

20 Within the equilibrium search framework this relationship is for example described by Robin and Roux (2002). Note that we have estimated our models using log capital related to machinery and inventory, capital related to machinery and inventory (in millions), log capital related to buildings and capital related to buildings (in millions) as instruments. The first stage of these regressions reveal that capital related to machinery and inventory (in millions) and log capital related to buildings are highly insignificant (P-values around 0.9) in the first-stage regressions and the inclusion of these as instruments in the regressions for women yield test results implying weak instruments (F-test of the excluded instruments yield F-values around 5–6). The second stage results still imply significantly higher turnover elasticity of men than that of women. However, to avoid weak instruments we exclude capital related to machinery and inventory (in millions) and log capital related to buildings from our final regressions.
The elasticity of worker turnover with respect to establishments’ wage premiums.

<table>
<thead>
<tr>
<th>MODEL:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV estimation:</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Dep. variable:</td>
<td>log separation rate</td>
<td>log turn. rate</td>
<td>log separation rate</td>
<td>log turn. rate</td>
<td>log separation rate</td>
<td>log turn. rate</td>
<td>log separation rate</td>
<td>log turn. rate</td>
</tr>
<tr>
<td>Wage premium</td>
<td>−0.877***</td>
<td>−0.495***</td>
<td>−0.943***</td>
<td>−0.382***</td>
<td>−2.451***</td>
<td>−0.871**</td>
<td>−2.292***</td>
<td>−0.218</td>
</tr>
<tr>
<td>Difference</td>
<td>−0.378***</td>
<td>−0.556***</td>
<td>(Men−women)</td>
<td>(1.311)</td>
<td>(0.133)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(pf)</td>
<td>0.525</td>
<td>0.381</td>
<td>0.278</td>
<td>0.398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Lm)</td>
<td>2.074</td>
<td>1.580</td>
<td>2.451</td>
<td>1.580</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Lk)</td>
<td>2.451</td>
<td>1.580</td>
<td>2.451</td>
<td>1.580</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(ωf)</td>
<td>0.649</td>
<td>0.649</td>
<td>0.649</td>
<td>0.649</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The manufacturing statistics cover the manufacturing industries, but also mining, oil, construction and some trade industries. Unit of observation is establishments, where each observation expresses average across the period of observation. In the regressions, each observation is weighted by the number of years the establishment is observed. Controls for 2-digit NACE industry and for county are in the form of dummies. Workforce educational qualification is controlled for in the form of variables measuring proportion of workers having educational qualification within 8 specific fields and by variables expressing years of education in excess of compulsory schooling and years of education in excess of compulsory schooling squared. The instrument vector consists of capital related to buildings (in 1000000 Nok) and log capital related to machinery and inventory. Standard errors reported in parentheses. ** denote significant at 5% level.

Table 5
The elasticity of labour supply facing each establishment.

<table>
<thead>
<tr>
<th>Excess turnover</th>
<th>Low educated</th>
<th>High educated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>Women</td>
<td>Men</td>
</tr>
<tr>
<td>Quits</td>
<td>1.492</td>
<td>1.142</td>
</tr>
<tr>
<td>Excess turnover</td>
<td>1.710</td>
<td>1.170</td>
</tr>
</tbody>
</table>

Note: Cells express the elasticities of labour supply facing each establishment, which are calculated using the estimates of models 3, 4, 7, and 8 of Table 2 (separations) and Table 3 (excess turnover) for all establishments.
across establishments that would arise from more wage elastic job-to-job moves on part of men.

7. The gender wage gap and labour market frictions

In this section we finally turn to the gender wage gap and ask how much of this gap can be related to frictions and thus potentially monopsonistic discrimination. In Table 7 we report from an estimation of log hourly wages from the pooled Level of Living Surveys 1991 and 1995. We have, for each educational group, estimated separately a simple log wage human capital regression model including years of education, potential experience and its squared value, a part time dummy, and a dummy for being considered high education. Robust standard errors are reported in parentheses. ***, ** and * denote significant at 1, 5 and 10% level of significance, respectively.

### Table 6

<table>
<thead>
<tr>
<th></th>
<th>Average log observed wage</th>
<th>Establishment’s wage premium</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Ratio of women in establishment</td>
<td>−0.7240** (0.0084)</td>
<td>−0.4723*** (0.0103)</td>
</tr>
<tr>
<td>Ratio of women in establishment squared</td>
<td>−0.2874*** (0.0581)</td>
<td></td>
</tr>
</tbody>
</table>

Control for: Intercept, education, year, county, industry, sector

R²(adjusted) 0.4940 0.4878 0.4934 0.6015 0.6009

N 10675 10519 10519 10519 10519

Estimation WLS WLS WLS WLS WLS 2WLS

Note: Unit of observation is establishment, where each observation is establishment-specific averages. Dependent variable in Model 1 is establishment-specific average of log daily wages. Dependent variable in models 2–5 is establishment-specific total wage premium. In the regression of Model 1 each observation is weighted according the number of observations of each establishment. In the regressions of models 2–5, each observation is weighted by the inverse variance of the wage premium. In Model 5 the ratio of women in establishment is considered endogenous, and is instrumented by the predicted proportion of women in establishment given that the establishment employs workers following the occupational gender distribution in the economy. Education denotes establishment-specific average years of education and dummy for being considered high education. Robust standard errors are reported in parentheses. ***, ** and * denote significant at 1, 5 and 10% level of significance, respectively.

As pointed out by Manning (2003) the elasticity of labour supply facing each establishment is composed of two parts: one part arising from the effects on aggregate labour supply and the other part arising from the effects on turnover associated with job-to-job transitions in the competition between employers. When competition between employers is less than perfect, each establishment may experience an upward sloping labour supply curve. The part arising from competition between employers may well dominate the part arising from effects on labour supply. Together with the previously reported results on aggregate supply, our results thus strongly indicate that the wage elasticity of job-to-job transitions of women is considerably smaller than that of men. In fact the difference in wage elasticity of job-to-job transitions has to be larger in absolute value than the difference in wage elasticity of aggregate labour supply, since in sum they add up to a negative number.

On average, women tend to have at least as high turnover as men. An insightful tradition arose from observations of this fact: Building on Mincer and Polacheck (1974) and Polacheck (1975) a host of work has demonstrated the importance of differences in turnover for differences in investment in human capital, and used this mechanism to explain gender differences in labour market outcomes. Our observation does not run counter to this mechanism. Quite on the contrary, we demonstrate that the endogenous part of women’s turnover; that which is sensitive to the firm’s wage policy, is small. This means that the exogenous part of women’s turnover is relatively larger, compared to men’s turnover behaviour. Gender differences in labour market outcomes may thus arise from both dimensions of turnover: less wage sensitivity provides employers with incentives for monopsonistic discrimination and more exogenous quits provides both employers and employees with fewer incentives for human capital accumulation and the provision of career opportunities.

It is a well known fact that during the recent decades, gender wage differences did not arise as a result of differences in wages between men and women in the same detailed job category within the same establishment (see for example Meyerson-Milgrom et al., 2001). However, gender segregation across jobs and establishments implies gender based differences in the elasticity of labour supply facing each establishment across jobs and occupations. These differences may explain the fact that female dominated jobs and occupations pay less, and that female dominated occupations tend to have flatter wage

### Table 7

<table>
<thead>
<tr>
<th>Gender-wage gap</th>
<th>Low educated</th>
<th>High educated</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Observed&quot; wage gap</td>
<td>−0.1600</td>
<td>−0.191</td>
</tr>
<tr>
<td>Estimated wage gap (quits)</td>
<td>−0.1095</td>
<td>−0.0381</td>
</tr>
<tr>
<td>Estimated wage gap (excess turnover)</td>
<td>−0.1455</td>
<td>−0.1277</td>
</tr>
</tbody>
</table>

Note: Cells express the gender wage gap, i.e., the amount lower wages women receive compared to that of men relative to men’s wages. The “Observed” wage gap follows from an estimation of a standard human capital wage regression using pooled Level of Living Surveys 1991 and 1995 data. The estimated wage gaps are calculated using figures from Table 5.

labour market frictions may potentially be an important source of the gender wage differentials in Norway.

8. Concluding remarks

Monopsonistic discrimination was refuted as an explanation of the male–female wage gap mainly because female labour supply has been found to be at least as wage-elastic as male labour supply. A recent study by Gelber (2008) confirms this finding utilizing IV methods and family responses to the Swedish 1991 tax reform. He finds that women’s labour supply is more responsive to changes in the marginal tax rate than men’s. How do we reconcile our findings of a less wage elastic labour supply facing each establishment with these observations?

If one allows the marginal revenue product to vary between men and women, Eq. (5) can be expressed as

$$\frac{\partial w_{\text{wm}}}{\partial w_{\text{em}}} = \frac{\partial \text{wage gap}}{\partial \text{wage gap}}.$$ 

The observed wage differential between men and women may then be completely explained by the model, assuming that the marginal revenue product of men is 2–6% (low educated) and 7–15% (high educated) higher than that of women.

---

21 If one allows the marginal revenue product to vary between men and women, Eq. (5) can be expressed as

$$\frac{\partial w_{\text{wm}}}{\partial w_{\text{em}}} = \frac{\partial \text{wage gap}}{\partial \text{wage gap}}.$$ 

The observed wage differential between men and women may then be completely explained by the model, assuming that the marginal revenue product of men is 2–6% (low educated) and 7–15% (high educated) higher than that of women.
profiles. Even if employers do not apply a policy of monopsonistic discrimination towards women as individuals or a group, the level of segregation in the labour market ensures that the same mechanisms may be at work, but rather at the level of jobs and occupations.

This study does not try to sort out potential reasons for the differences in job-offer arrivals or exogenous separations between the sexes that may underlie the differences in the elasticity of labour supply.22 Thus most of the usual explanations are still relevant (see Manning, 2003:200–204), such as potential gender differences in time constraints or in the evaluation of non-pecuniary versus pecuniary aspects of a job. Allowing for endogenous search intensities for both employer and employees is also likely to highlight self-enforcing mechanisms that may add to the explanation of the magnitude of the differences between men and women.

Although Norway is known for gender equalising policies, for example with respect to public child-care, women still have the main domestic responsibility for children. Thus when Manning finds that family commitment prevent both job search and taking jobs to a higher degree for women than men in the UK (Manning, 2003: Table 7.2 on p.201), this should be valid for Norway as well. Since the public child-care is geographically linked within the home municipality, and there exists an excess demand for public day-care, job lock-in issues may arise at important stages of the career for women.

Our main empirical result is that excess turnover of workers in the establishments is significantly more sensitive to the wage premium of men than to that of women, which again implies a higher elasticity of labour supply facing each establishment for men than for women. This difference provides employers with an incentive to employ the policy of monopsonistic discrimination. Even if employers cannot discriminate between men and women within the same jobs, a practice which is illegal in most countries, they may differentiate their wage policy towards different types of jobs or occupations with different gender compositions. Our results show that as much as 70–90% of the gender wage gap for low-educated workers may be attributed to differences in labour market frictions between men and women, while the similar figures for high-educated workers ranges from 20 to 70%. Search frictions may thus be one of the major sources of gender wage inequality.

Appendix A

Table A1

The estimation of establishments' wage policy: parameter estimates from the within-individual log wage regressions.

<table>
<thead>
<tr>
<th></th>
<th>Low educated</th>
<th>High educated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men  Women</td>
<td>Men  Women</td>
</tr>
<tr>
<td>Seniority-w</td>
<td>0.0079</td>
<td>0.0127</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>Seniority squared-w</td>
<td>-0.0002</td>
<td>-0.0005</td>
</tr>
<tr>
<td></td>
<td>(6.745E-6)</td>
<td>(0.000001)</td>
</tr>
<tr>
<td>Experience squared-w</td>
<td>-0.0005</td>
<td>-0.0002</td>
</tr>
<tr>
<td></td>
<td>(4.434E-6)</td>
<td>(6.779E-6)</td>
</tr>
<tr>
<td>Short part-time</td>
<td>-0.8559</td>
<td>-0.5244</td>
</tr>
<tr>
<td>(4-19 h)</td>
<td>(0.0024)</td>
<td>(0.0017)</td>
</tr>
<tr>
<td>Long part-time</td>
<td>-0.3573</td>
<td>-0.2257</td>
</tr>
<tr>
<td>(20-29 h)</td>
<td>(0.0031)</td>
<td>(0.0016)</td>
</tr>
<tr>
<td>= transformed year and establishment dummies</td>
<td>root mean square error</td>
<td>0.2047</td>
</tr>
</tbody>
</table>

Note: The results from OLS-regressions on the within-individual transformed covariates (each covariate subtracted its individual-specific mean), estimated separately for high and low-educated women and men.

References


22 In addition to potential differences between male versus female search behaviour, the theoretical analysis of Bowlius and Eckstein (2002) shows that even if only a fraction of employers have discriminating tastes in employment, there will be effects in the whole labour market via differences in job-arrival rates between groups.