Compensating Differentials for Sexual Harassment

By Joni Hersch

Sexual harassment is prohibited under Title VII of the Civil Rights Act of 1964 as discrimination on the basis of sex. Although sexual harassment is illegal, some workers are sexually harassed on the job. For example, a 1994 survey of federal employees by the US Merit Systems Protection Board found that 44 percent of women and 19 percent of men had experienced unwanted sexual attention on the job in the preceding two years. Sexual harassment claims comprise a large component of charges filed with the Equal Employment Opportunity Commission (EEOC).

Sexual harassment is widely viewed as an instrument of power and intimidation rather than primarily as an expression of sexual desire, and such harassing behavior may cause victims as well as their coworkers to be less productive. Indeed, sexual harassment is a form of employment discrimination precisely because it alters the “terms, conditions, or privileges of employment” and interferes unreasonably with the ability of those in the protected classes to perform their jobs. Sexual harassment may lead to lower pay if harassment reduces worker productivity by, for instance, inducing inefficient turnover, increasing absenteeism, and generally wasting work time as workers attempt to avoid interaction with harassers.

An alternative hypothesis is that, similar to jobs in which workers face a high risk of death or disabling injury, workplace sexual harassment is an undesirable working condition that may generate a compensating pay differential. There are, of course, clear differences between job risks and sexual harassment. Job risks are usually a consequence of the technology in the industry, whereas sexual harassment arises from illegal personal behavior. Yet similar to job risks, sexual harassment is costly for firms to eliminate, which may result in sexually harassing behavior occurring in some workplace environments.

Thus, the direction of the relation between sexual harassment and wages is not predictable a priori. While pay differentials on the basis of sex and on job risks and other working conditions have been widely studied by economists, sexual harassment has received little attention within the economics literature.

This paper provides evidence of the relation between the risk of sexual harassment and wages. While one approach to detecting the effect on wages of sexual harassment would be to estimate wage equations controlling for whether an individual reports that he or she had been sexually harassed, sexual harassment on the job is unlikely to be exogenous with respect to wages, and it is difficult to identify appropriate variables that would allow instrumental variables estimation. In addition, there are almost no data reporting information on sexual harassment as well as wages and other determinants of wages.

To avoid these problems, I adopt the conventional hedonic wage methodology used to estimate compensating wage differentials for risk of injury or death. Specifically, using data on individual charges filed with the EEOC (which
I obtained under the Freedom of Information Act, I calculate gender-specific estimates of the risk of sexual harassment by industry and age group.

Matching these risk measures to data from the Current Population Survey (CPS), I estimate wage equations controlling for the risk of sexual harassment and for other determinants of wages, including occupation and the percent female in the worker’s industry. The wage equation estimates show that greater risk of sexual harassment is associated with a statistically significant wage premium. Women employed in jobs with an average probability of sexual harassment are paid a compensating differential of 25 cents per hour relative to comparable women employed in jobs with no risk of sexual harassment. Men employed in jobs with an average probability of sexual harassment are paid a compensating differential of 50 cents per hour relative to comparable men employed in jobs with no risk of sexual harassment.

I calculate gender-specific estimates of the risk of sexual harassment by industry and age group.2 The numerators in this risk measure are the number of sexual harassment charges by industry, age group, and sex. The denominators are the corresponding levels of industry employment by age group and sex from the CPS (excluding self-employed workers who would generally not be able to claim sexual harassment against an employer).

Because the sexual harassment rates are calculated at the industry level, the biggest problem in calculating sexual harassment risk is missing data on industry codes in the EEOC claims data. Industry is not a required field on the EEOC claims records. Industry code is missing in about 28 percent of the claims prior to 2006. According to an EEOC employee, missing data on industry became an even greater problem beginning in 2006 (for example, NAICS code is missing for 56 percent of the claims in 2008), in part because in 2006 the EEOC switched from SIC code to NAICS code and the drop-down menu for NAICS code was harder to use. For claims filed prior to 2006 with SIC code reported, the EEOC used a crosswalk to assign NAICS code. In addition, if the employer named in the claim could be linked to the EEO-1 database (e.g., private firms with 100 or more employees or private federal contractors with 50 or more employees), then the NAICS code reported in the EEO-1 form was transferred to the EEOC claims file. Thus, industry code is missing more frequently in smaller firms than in larger firms and is also missing far more frequently starting in 2006 than in earlier years.

2 This follows the methodology used to construct fatality rates by industry, age, and sex in W. Kip Viscusi and Joni Hersch (2008).
The number of claims by industry, age group, and sex are used as the numerators in the calculation of sexual harassment rates, so eliminating all observations with missing information, even if random, leads to lower sexual harassment rates for an industry/age group/sex than are true. If industry code is missing at random, then we have classical measurement error in an explanatory variable in the regression equation, and the coefficient on this variable will be biased toward zero. A regression of an indicator for missing industry on characteristics of claims shows that although there is some systematic variation in the probability that industry is recorded, observable characteristics explain fairly little of the variation in whether industry is reported. Controlling for age group, sex, whether the claim was reported to the EEOC or FEPA, indicators of strength of claim (from definitely litigate to dismiss), firm size, institution type (e.g., private employer, educational institution), filing year, and race, the adjusted $R^2$ is a very low 0.038 based on EEOC claims for the period FY 2000–FY 2004. Thus, while recognizing that the sexual harassment rate is measured with error, the low predictive power of observable characteristics suggests that assuming the measurement error is largely random is not unreasonable.

Because the number of missing industry codes increased substantially after 2006, I use data from FY 2000–FY 2004 to calculate the numerators in sexual harassment rate calculation. There are 48,741 individual claims that include sexual harassment as an issue. Of these claims, 42,065 are claims by women and 6,676 are claims by men. The denominators are based on employment data from the 2004 CPS excluding self-employed workers. The sexual harassment rates used in the wage equations are calculated by sex for two-digit industry (52 industries) and six age groups (15–24, 25–34, 35–44, 45–54, 55–64, and age 65 and older).

### II. Patterns of Sexual Harassment Risk

Table 1 reports sexual harassment rates per 100,000 workers by sex and major industry, as well as the percent female in the industry. The pattern across industries indicates that women are at a greater risk of sexual harassment in male-dominated industries, with the pairwise correlation between the female rate and percent female equal to $-0.68$ ($p = 0.01$). The male rate is not correlated with the female rate; nor is the male rate correlated with percent female. Figure 1 demonstrates the pattern of sexual harassment risk for women by age for four selected industries, which shows an inverted U-shaped pattern of risk of sexual harassment with age. The pattern is largely similar for the other industries. With the exception of mining, men also have an inverted U-shaped pattern of risk of sexual harassment with age. Men’s risk of sexual harassment is substantially below that of women at every age.
III. Wage Equations

I estimate conventional wage equations controlling for sexual harassment risk and for other standard determinants of wages using CPS data for 2005. The dependent variable is the log of the hourly wage, which is either reported directly or calculated as weekly earnings divided by usual hours worked per week. In addition to the sexual harassment rate, the explanatory variables are years of education and potential experience and its square, and indicator variables for occupation (management, business, financial; professional and related; healthcare support; protective service; food preparation and serving related; building and grounds cleaning and maintenance; personal care and service; sales and related; office and administrative support; natural resources, construction, maintenance; and production, transportation, material moving), race (white; black; American Indian; Asian or Pacific Islander; or more than one race reported), Hispanic ethnicity, married, employed by the government, union member or covered by union contract, full time employment, metropolitan location, and region.

Because there is evidence that sexual harassment is more prevalent in work settings with predominantly one sex, it is possible that any positive effect of sexual harassment for women reflects the higher pay associated with male-dominated jobs. I also therefore control for the percent female in the individual’s narrowly defined (four-digit) industry.

Table 2 summarizes the results with separate wage regressions estimated by sex, reporting only the coefficients on sexual harassment risk and percent female in the industry. The standard errors are clustered by industry and age group because all workers within the same industry and age group are assigned the same gender-specific value for sexual harassment. For ease of interpretation, this table also reports mean predicted log wages estimated at the sample means of all variables and at the sample means of all variables assuming a zero sexual harassment rate.

As expected, there is an inverse relation between the percent female in the industry and wages for both men and women. Both men and women receive a statistically significant wage premium for the risk of sexual harassment. The log wage difference between a job with zero sexual harassment risk and a job with the mean sexual harassment risk is 0.0155, or about 25 cents per hour for women, and 0.0252, or about 50 cents per hour for men. The large compensation for sexual harassment risk for men is surprising. One possible explanation is that since men infrequently file sexual harassment claims, those claims that are filed are particularly egregious, and exposure to such risk warrants a larger compensating differential than received by women.
IV. Concluding Remarks

Although illegal, sexual harassment occurs in the workplace. The risk of sexual harassment varies by sex, industry, and age. Women face far greater risk of sexual harassment than men in every industry and at every age, and women’s risk of sexual harassment is positively correlated with the percent male in the industry. The risk of sexual harassment is highest for workers between ages 25 and 44.

The central empirical issue addressed in this paper is whether sexual harassment lowers wages by reducing productivity or raises wages as workers require a compensating differential to incur this risk. Sexually harassing behaviors range from sexual looks or sexual jokes from coworkers to assault and rape, with surveys indicating that sexual looks and comments are by far the most common type of workplace sexually harassing behavior. Because sexual harassment is costly for firms to eliminate, some forms of sexual harassment such as looks and comments that are hard to monitor may occur in the workplace. This paper shows that, on balance, workers receive a wage premium for exposure to the risk of sexual harassment in much the same way that workers receive a wage premium for the risk of fatality or injury.

REFERENCES


<table>
<thead>
<tr>
<th>Coefficients</th>
<th>(Standard error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Sexual harassment rate by industry, age, and sex per 100,000 workers</td>
<td>0.0018*</td>
</tr>
<tr>
<td>Percent female in industry</td>
<td>-0.2000**</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.40</td>
</tr>
<tr>
<td>Number of observations</td>
<td>77,896</td>
</tr>
<tr>
<td>Average sexual harassment rate</td>
<td>8.6056</td>
</tr>
<tr>
<td>Predicted log wage:</td>
<td></td>
</tr>
<tr>
<td>At sample means</td>
<td>2.6403</td>
</tr>
<tr>
<td>Sexual harassment rate = 0</td>
<td>2.6248</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors clustered by two-digit industry and age group are reported in parentheses. Sample is comprised of respondents to the 2005 CPS who are employed, not self-employed, ages between 18 and 64, with wages between $1.50 and $100 per hour. Additional variables included in the regressions are a constant, potential work experience, potential experience squared, years of education, and indicator variables for occupation, race, Hispanic ethnicity, married, government employer, union or employee association, full-time employment, metropolitan location, and region. All values are weighted by CPS earnings weight.

**Significant at the 1 percent level.

*Significant at the 5 percent level.