RIETI Discussion Paper Series 07-E-020

Male-Female Wage and Productivity Differentials: A Structural Approach Using Japanese Firm-level Panel Data

ASANO Hirokatsu
Asia University

KAWAGUCHI Daiji
RIETI

Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 07-E -020

Male-Female Wage and Productivity Differentials: A Structural Approach Using Japanese Firm-level Panel Data

Hirokatsu Asano[1]
Faculty of Economics, Asia University
and
Daiji Kawaguchi[2]
Faculty of Economics, Hitotsubashi University
Research Institute of Economy, Trade and Industry

March 30, 2007

[1]Address: 5-24-10 Sakai, Musashino, Tokyo 180-8629, Japan, Tel: (81)-422-36-5206, E-Mail: asano@asia-u.ac.jp

[2]Address: Naka 2-1, Kunitachi, Tokyo 186-8601, Japan, Tel: (81)-42-580-8851, Fax: (81)-42-580-8182, E-Mail: kawaguch@econ.hit-u.ac.jp

Abstract

In an attempt to explain the male-female wage differential, we estimated the relative marginal productivity and relative wage of female workers compared to those of male workers, using panel data from Japanese firms. The estimation results indicate that firms hiring 10 percentage points more women produce 0.8 percent more given the total wage bill and other inputs. Cross-sectional estimates that neglect firm fixed effects indicate that female workers' marginal productivity is 45 percent of male workers', while female wage is 30 percent of male wage. These estimates indicate that part of the wage differential cannot be explained by the productivity differential. The estimation that allows for the correlated productivity/demand shocks suggests the robustness of the results. The IV estimator that allows for firm-level fixed effects seems to suffer from the bias due to the positive correlation between productivity/demand shocks and female employee proportion. Evidence found in this study rejects the null hypothesis that the male-female wage differential reflects the male-female productivity differentials.

Keywords: Gender Wage Differential, Productivity, Panel Data, Fixed Effects, Proxy Variable Estimation

JEL Classification Code: J31

1 Introduction

Wage differentials between the sexes are observed worldwide, and these differentials have persisted for a long time. Figure 1 shows the hourly wages of male and female workers in Japan between 1988 and 2002. Female workers' hourly wage has been persistently around 60 percent of male workers' hourly wage. Labor economists, as well as the general public, have argued about why this sex-wage differential exists.

Several theories attempt to explain the sex wage differentials and they can be largely classified into two categories. The first category claims that the wage differential reflects the productivity differential between the two sexes. Men and women may have different productivity due to biological reason (Ichino and Moretti (2006)), the difference in the job matching quality (Bowlus (1997)) or the difference in the effort level put in the market production (Becker (1985)). Second, the differential may be due to employer sex discrimination. If there are sufficiently large number of discriminatory employers in the labor market so that the marginal female worker is employed by a discriminatory employer, then a male-female wage differential emerges in the market equilibrium (Becker (1971)).

Labor economists have been using multiple regression to distinguish these two hypotheses. They regress the log of wage on independent variables that presumably capture workers' productivity, and the residual male-female wage differential is attributed to sex discrimination. However, drawing a definitive conclusion from this regression is difficult because the productivity differential between the sexes that cannot be observed may be included in the error term. In particular, if the productivity differential raises due to the difference in biological characteristics, job match quality or the effort level, these differences are not captured by labor force survey that are typically used for estimating a wage equation.

In an attempt to overcome the difficulties mentioned above, this paper suggests a new reduced form approach of identifying the discrepancy between relative productivity and wage of female compared with male. This method utilize wage bill as the measure of labor input because if wages are paid according to workers' marginal product, the wage bill captures quality adjusted labor input. Once the wage bill is conditioned on, under the null of no discrimination against women, female proportion should not explain the firm' output. The estimation results based on Japanese firm's panel data strongly rejects this null hypothesis.

Given the rejection of no discrimination against female labor, we directly estimate the relative productivity of male and female workers by estimating the production function using the same data. This estimated productivity differential is compared with the wage differential estimated from individual firms' accounting data. This exercise reveals whether the wage differential is due to the productivity differential. This approach has been employed by Hellerstein and Neumark (1999) to analyze Israeli data, and Hellerstein et al. (1999) and Hellerstein and Neumark (2007) to analyze US data. These researchers found supportive evidence for sex discrimination in the US, but not in Israel; they found a larger sex-wage gap than productivity gap in the US, but not in Israel. Following the similar methods, Crepon et al. (2002) analyzed French data and Ilmakunnas et al. (2004) analyzed Finnish data.

A related approach examines the empirical implication of employers' discrimination theory. If the workers' sex composition in each firm is determined to maximize its profit, the workers' sex composition should not affect firms' profits after conditioning on output and input prices because of the envelope theorem (Hotelling's lemma). This result no longer holds once the firms' objectives include satisfying the employers' preference for sex discrimination. If employers' objectives are heterogeneous and the equilibrium male-female wage differential reflects discrimination against women, then those employers without discriminatory preferences against women should hire more women and earn higher profits than discriminatory employers. This hypothesis was tested by Hellerstein et al. (2002) using US data and by Kawaguchi (2007) using the Japanese firm-level panel data that are used in this paper. Both papers found evidence that is consistent with the existence of sex discrimination.[1]

This paper attempts to complement Kawaguchi (2007) by obtaining structural parameters, and consequently, it sheds light on the mechanism of the male-female wage differential more directly. As Hellerstein and Neumark

[1]Kodama et al. (2005) and Sano (2005) also found that firms with a higher female employee proportion earn higher profits.

(1999) and Crepon et al. (2002), we utilize panel data to estimate production function and wage equation. Using panel data, we can allow for the heterogeneity in the individual firms' productivity that may be correlated with the sex composition of their workers. Controlling for unobserved technological heterogeneity across firms is important because those firms with productive technology may accommodate female workers, while at the same time earning higher profits due to their productive technology.

The results obtained in this paper are summarized as follows. Cross-sectional estimates that neglect firm fixed effects indicate that the marginal productivity of female workers is 44 percent of that of male workers, while female wage is 31 percent of male wage. These estimates indicate that a part of the wage differential cannot be explained by the productivity differential. On the other hand, the IV estimates that allow for firm-level fixed effects indicate that both female marginal productivity and wage are about 50 percent of those of male workers. However, given speedier labor adjustment among female, the temporary shock to a firm may have strong correlation between female proportion and this correlation may severely bias the fixed effects estimator as recent studies of production function estimation indicates (Ackerberg et al. (2007)). Thus we eventually rely on the results from the proxy variable estimation developed by Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg et al. (2006) and Wooldridge (2005).

The rest of this paper is organized as follows. Section 2 explains the data used in this study and implement descriptive and reduced from analysis.

Section 3 introduces the structural model of the production function and wage equation. This section also discusses the method of estimation. Section 4 reports the results, and Section 5 concludes.

2 Data and Descriptive Analysis

We used the basic survey of firms' activity collected by the Ministry of Economy, Trade, and Industry (METI) of the Japanese government to examine female's relative productivity to male's. The survey is a firm-level census survey that covers all firms that hire more than 50 employees and hold more than 30 million yen in capital. The available data cover 7 years, 1992 and every year between 1995 and 2000; and the sample size is about 25,000 firms for each year.[2]

From the data sets, we extracted each firm's total sales, sales cost, overhead cost, data on the firm's employees, such as the number of employees with sex breakdown, the book value of its fixed assets, the year of the firm's origin, and a three-digit code indicating the industry in which the firm operates. There were originally 180,838 firm-year observations in the 7 years of data, but after excluding observations with missing sales information or inconsistent employee records, there remained 177,868 firm-year observations. The survey record unfortunately does not distinguish missing values and zeros, except when a firm did not answer the entire survey. Since replying to

[2]The survey was not conducted in 1993 and 1994. The survey still continues but it stopped collecting the sex decomposition of workers in 2001 and afterward.

the survey is compulsory due to the Statistics Law and because the METI exerts its best effort to fill in the missing values with a follow-up phone survey, missing values are presumably rare. Thus, all values of zero in the record are treated as actual zeros.

The descriptive statistics of the analysis sample are reported in Table 1. The table shows that firms with more of women tend to be smaller in its total sales, total employment, fixed assets, costs of materials than the firms with less of women. Also per capita wage payment is lower in firms with more women. Firms with more of women are concentrated in Light Manufacturing and Wholesale and Retail industries.

Figure 2 illustrates the distribution of average female employment composition during the sample period by firms. The mode of the distribution is around 0.2 and the distribution is skewed to the left. This figure shows that there are sufficient amount of cross-sectional variation in female proportion across firms.

As a first step reduced form analysis, we estimate the following Cobb-Douglas production function:

$$\ln y_{it} = \beta_{0i} + \beta_1 \ln ql_{it} + \beta_2 \ln k_{it} + \beta_3 \ln m_{it} + ind + year + u_{it}, \qquad (1)$$

where i and t are the subscripts for firm and time, respectively, y_{it} is total sales, $\tilde{a}_i$ is the firm-specific, time-constant technology, ql_{it} is the labor input that is measured in efficiency units, k_{it} is capital input, m_{it} is the intermediate input, ind is industry dummy variables and $year$ is year dummy variables, and u_{it} is the unobserved idiosyncratic shock to production.

The quality adjusted labor input is $ql_{it} = \sum_{j=1}^{J} q_{itj} \times l_{itj}$ where q_{itj} is marginal productivity and l_{itj} is man-hour labor input of worker type j in firm i at time t. If workers are paid according to their marginal productivity, $q_{itj} = w_{itj}$ where w is hourly rate of pay and the wage bill (i.e. $\sum_{j=1}^{J} w_{itj} \times l_{itj}$) captures the quality adjusted labor input.

Under the null hypothesis that all workers are paid according to their marginal products, the error term u_{it} should not be correlated with labor force composition. However, if the female relative wage to male is lower than the female relative productivity to male, higher female proportion results in the higher amount of sales given wage bill and other inputs.

Table 2 reports the results of this reduced form regression. Column (1) reports the result of fitting Cobb-Douglas production function. Firm level average of the residual from this regression is regressed upon the firm-level average female proportion. The result is reported in Figure 3 and this result indicates more female proportion results in higher total factor productivity. Table 2 Column (2) reports the specification with female proportion. The coefficient for female ratio indicates that 10 percentage points more women results in 0.8 percent higher level of production given other inputs. The results of these reduced form regression indicates that relative pay for women is smaller than the relative productivity of women compared with men.

3 Structural Model

To estimate the marginal product of the labor of male and female workers, we need to specify the functional form of the production function. we assume the Cobb-Douglas production function, as in Hellerstein and Neumark (1999), Hellerstein et al. (1999) and Hellerstein and Neumark (2007). The production function is specified as:

$$y_{it} = (\exp a_i)(ql_{it})^{\alpha} k_{it}^{\beta} m_{it}^{\gamma} (\exp u_{it}), \tag{2}$$

where i and t are the subscripts for firm and time, respectively, y_{it} is total sales, a_i is the firm-specific, time-constant technology, ql_{it} is the labor input that is measured in efficiency units, k_{it} is capital input, m_{it} is the intermediate input, and u_{it} is the unobserved idiosyncratic shock to production.

The labor input in the efficiency unit is the weighted sum of the numbers of male employees and female employees as follows:

$$ql_{it} = l_{m,it} + \psi l_{f,it} = l_{it}(1 + (\psi - 1)(\frac{l_f}{l})_{it}), \tag{3}$$

where l_m stands for the number of male employees, l_f stands for the number of female employees, and l stands for the total number of employees. The parameter ψ indicates the relative productivity of female workers to male workers. By taking a logarithm of (2) and substituting (3) with ql, we obtain:

$$\ln y_{it} = a_i + \alpha \ln(l_{it}(1 + (\psi - 1)(\frac{l_f}{l})_{it})) + \beta \ln k_{it} + \gamma \ln m_{it} + ind\delta + year\tau + u_{it}. \tag{4}$$

we included one-digit industry dummies (in nine categories) to allow for the differences in a across industries. Time dummies capture the effect of macro-economic shocks and inflation.

The estimated, relative productivity of women is compared to their relative wage. The data used in this study contain the total wage bill, but do not contain its sex breakdown. Thus, we estimate the relative wage of women under the assumption that all firms behave as price takers. The total wage bill is defined as:

$$wb_{it} = w_{m,t}l_{m,it} + w_{f,t}l_{f,it} = w_{m,t}(l_{it} - l_{f,it} + \frac{w_f}{w_m}l_{f,it}) = w_m l_{it}(1 + (\lambda - 1)(\frac{l_f}{l})_{it}), \tag{5}$$

where wb_{it} is the wage bill, w_m is male wage, w_f is female wage, and λ is the female wage relative to male wage ($= w_f/w_m$). This equation is a definitional equation, rather than a behavioral one.

we estimate this equation by taking a natural logarithm and allowing for unobserved factors. The estimation equation becomes:

$$\ln(wb_{it}/l_{it}) = \ln w_m + \ln(1 + (\lambda - 1)(\frac{l_f}{l})_{it}) + ind\delta + year\tau + e_{it}, \tag{6}$$

where e_{it} is the error term. We included industry dummies, assuming that the inter-industry wage differential could persist because of friction in the labor movement across industries. Year dummies capture the effect of inflation or macro-economic shock. This estimated, relative wage of women λ is compared to relative female productivity ψ. A consistent estimation of parameters is possible via GMM under certain exogeneity assumption on u_{it}

and e_{it}.

4 Econometric Estimation

4.1 Pooled Cross-Sectional Estimation

This section lays out the methods to estimate the parameters in the structural equations. To simplify the notation, the production function is expressed as:

$$y_{1it} = f_1(x_{1it}; \theta_1) + u_{1it}, \tag{7}$$

where y_{1it} is log of output and x_{1it} is a $1 \times k_1$ vector of inputs, θ_1 is a $k_1 \times 1$ vector of production function parameters, and u_{1it} is the error term. The index i is for firms and t is for time. The wage bill equation is expressed as:

$$y_{2it} = f_2(x_{2it}; \theta_2) + u_{2it}, \tag{8}$$

where y_{2it} stands for log of wage bill and x_{2it} is a $1 \times k_2$ vector of labor composition, θ_2 is a $k_2 \times 1$ vector of wage bill equation parameters, and u_{2it} is the error term.

Stacking time-series of observations for each individual, we define each individuals' explanatory variable matrices x_{1i} and x_{2i} as:

$$x_{1i} = \begin{pmatrix} x_{1i1} \\ \vdots \\ x_{1iT} \end{pmatrix}, \; x_{2i} = \begin{pmatrix} x_{2i1} \\ \vdots \\ x_{2iT} \end{pmatrix}. \tag{9}$$

These are the $k_1 \times T$ and $k_2 \times T$ matrices, respectively.

Two equations are stacked together to form the explanatory variable matrix

$$x_i = \begin{pmatrix} x_{1i} & 0 \\ 0 & x_{2i} \end{pmatrix}, \tag{10}$$

which is a $2T \times (k_1 + k_2)$ matrix.

The stacked error vector is also defines as following:

$$u_i = \begin{pmatrix} u_{1i} \\ u_{2i} \end{pmatrix}. \tag{11}$$

Under the assumption that the explanatory variables and the error term in each equation are contemporaneously not correlated (i.e. $Ex'_{jit}u_{jit} = 0$ for $j = 1, 2$ and all t), the moment condition for the identification is stated as following:

$$Ex'_i u_i = 0. \tag{12}$$

Using this moment condition, the GMM estimator of $\theta \equiv [\theta'_1 \; \theta'_2]'$ is defined as the solution to the following problem:

$$\min_{\{\theta\}} (\sum_{i=1}^{N} x'_i u_i)' \Xi^{-1} (\sum_{i=1}^{N} x'_i u_i), \tag{13}$$

where Ξ is an arbitrary $2T \times 2T$ symmetric positive semi-definite matrix. The minimum variance estimator is obtained by setting $\Xi = Var(x'_i u_i) = E(x'_i u_i u'_i x_i)$ (Hansen (1982)).

We attempt to capture intertemporal correlation (i.e. $Cov(u_{jit}, u_{ji\tau}) = 0$ for $j = 1, 2$, $t \neq \tau$ and all i) because the shock to the production or wage bill could be persistent for some period. A way to allow for time dependence of the error term is not to impose assumption on the error structure and directly estimate $E(x'_i u_i u'_i x_i)$. The sample correspondence of this flexible error structure is given as:

$$\hat{\Xi} = \frac{1}{N} \sum_{i=1}^{N} \begin{pmatrix} x'_{1i}\hat{u}_{1i}\hat{u}'_{1i}x_{1i} & x'_{1i}\hat{u}_{1i}\hat{u}'_{2i}x_{2i} \\ x'_{2i}\hat{u}_{2i}\hat{u}'_{1i}x_{1i} & x'_{2i}\hat{u}_{2i}\hat{u}'_{2i}x_{2i} \end{pmatrix}, \tag{14}$$

where $\hat{u}_{1i}$ and $\hat{u}_{2i}$ are the residuals from the GMM estimation that uses identity matrix as the weighting matrix. The GMM estimator that uses this matrix as weighting matrix results in the efficient estimator under very flexible error structure assumption.

The variance of this estimator is estimated as:

$$(\frac{1}{N}\sum_{i=1}^{N} x_i' \nabla_\theta u(y_i, x_i; \hat{\theta}))' \hat{\Xi}^{-1} (\frac{1}{N}\sum_{i=1}^{N} x_i' \nabla_\theta u(y_i, x_i; \hat{\theta})), \tag{15}$$

where $\nabla_\theta u(y_i, x_i; \hat{\theta})$ is the $2T \times (k_1 + k_2)$ matrix defined as following:

$$\begin{pmatrix} \frac{\partial u_1}{\partial \theta_1}(y_{1i}, x_{1i}; \hat{\theta}_1) & 0 \\ 0 & \frac{\partial u_2}{\partial \theta_2}(y_{2i}, x_{2i}; \hat{\theta}_2) \end{pmatrix}. \tag{16}$$

4.2 Fixed Effects Estimation

Thus far we assumed the exogeneity of u_{jit} from x_{jit} for $j = 1$ (production function) and $j = 2$ (wage bill equation). However, if the error component captures the firm heterogeneity in the productivity, the input level may vary across firms corresponding to the productivity heterogeneity. In addition the unobserved wage bill determinant may be correlated with workers' composition as Kawaguchi (2007) found the correlation between time-invariant high profit factor and workers' sex composition.

To address the possible endogeneity of explanatory variables caused by firm-level time-invariant heterogeneity, we implement fixed effects estimation. We assume that the error term is decomposed into time-invariant fixed effects and idiosyncratic error (i.e. $u_{jit} = c_{ji} + e_{jit}$ for $j = 1, 2$. We allow for the correlation between c_{ji} and x_{jit} but assume the strict erogeneity of idiosyncratic error term. The strict exogeneity implies that the idiosyncratic shock of the current period is not correlated with explanatory variable in the other periods. This is to say, $Ex'_{ji\tau}e_{jit} = 0$ for $\tau = 1, ..., T$ for given t. Under this strict exogeneity assumption, the composite error term u_{jit} is not correlated with mean-deviation form of the explanatory variable, x_{jit}. This is because the mean deviation form of the explanatory variable is not correlated with the fixed effects; $E[(x'_{jit} - \bar{x}'_{ji})c_{ji}] = E[(x'_{jit} - (1/T)\sum_{t=1}^{T} x'_{jit})c_{ji}] = E[x'_{jit}c_{ji}] - (1/T)\sum_{t=1}^{T} E[x'_{jit}c_{ji}] = 0$ for $j = 1, 2$. In addition, strict exogeneity implies $E[(x'_{jit} - \bar{x}'_{ji})e_{jit}] = 0$.

Using the fact that the mean deviation form of the explanatory variables does not correlate with the composite error terms, we define the following instrumental variable matrix for production function:

$$z_{ji} = \begin{pmatrix} x_{j1i1} - \bar{x}_{j1i} & x_{j2i1} - \bar{x}_{j2i} & \cdots & x_{jk_1i1} - \bar{x}_{jk_1i} \\ \vdots & \vdots & \ddots & \vdots \\ x_{j1iT} - \bar{x}_{j1i} & x_{j2iT} - \bar{x}_{j2i} & \cdots & x_{jk_1iT} - \bar{x}_{jk_1i} \end{pmatrix}, \tag{17}$$

for $j = 1, 2$. Combining z_{1i} and z_{2i}, we define system instrumental variable matrix:

$$z_i = \begin{pmatrix} z_{1i} & 0 \\ 0 & z_{2i} \end{pmatrix}. \tag{18}$$

The moment condition that allows for the correlation between fixed effects and explanatory variable but assumes the strict exogeneity of idiosyncratic error term becomes:

$$Ez'_i u_i = 0. \tag{19}$$

Using this moment condition, the GMM estimator of $\theta \equiv [\theta'_1 \theta'_2]'$ using z_i

as instrument is defined as the solution to the following problem:

$$\min_{\{\theta\}} (\sum_{i=1}^{N} z_i' u_i)' \Xi^{-1} (\sum_{i=1}^{N} z_i' u_i), \tag{20}$$

where Ξ is an arbitrary $2T \times 2T$ symmetric positive semi-definite matrix. The choice of weighting matrix Ξ depends on the assumption on the covariance structure of error term as in the cross-sectional estimator. However u_{jit} necessarily has serial correlation within a firm due to the existence of firm fixed effects. Thus we use the robust formula of the variance covariance matrix as in (14). The variance of the estimator is obtained by replacing x_i with z_i in the formula (16).

4.3 Proxy Variable Estimation

Recent studies cast doubt on the consistency of the fixed effects estimator because within variation of input at the firm level can be correlated with the idiosyncratic shock to the firm. To overcome the limitation of the fixed effects estimation, Olley and Pakes (1996) and Levinsohn and Petrin (2003) suggested to utilize investment amount or intermediate input amount as proxy variables for the temporal shock the firm is experiencing. We now employ the method originally suggested by Levinsohn and Petrin (2003). The actual implementation is based on Wooldridge (2005) because his discussion on identification is transparent and the estimator is more efficient than the original estimator.

Now, assume that the production function is given as:

$$y_{1it} = f_1(ql_{it}, k_{it}, m_{it}; \theta_1) + v_{1it} + e_{1it}, \tag{21}$$

where ql_{it} is quality adjusted labor input, k_{it} is capital input, m_{it} is intermediate input. The first part of the error term v_{1it} is the productivity shock that can be correlated with ql_{it} and m_{it}. The other part of error term e_{1it} is assumed to be sequentially exogenous (i.e. $E[e_{it}|ql_{it}, k_{it}, m_{it}, ql_{it-1}, k_{it-1}, m_{it-1}, ..., ql_{i1}, k_{i1}, m_{i1}] = 0$).

Following Levinsohn and Petrin (2003), we assume that this productivity shock is recovered from the firm's choice of intermediate input quantity, given the level of capital that is a state variable (i.e. $v_{1it} = g(m_{it}, k_{it})$). As in Levinsohn and Petrin (2003), we assume that the productivity shock follows Markov process that is:

$$E(v_{1it}|k_{it}, l_{it-1}, k_{it-1}, m_{it-1}, ..., l_{i1}, k_{i1}, m_{i1}) \tag{22}$$

$$= E(v_{1it}|v_{1it-1}) = f(g(m_{it-1}, k_{it-1})). \tag{23}$$

Thus, the current productivity shock can be denoted as: $v_{1it} = E(v_{1it}|v_{1it-1}) + b_{it}$, where b_{it} is the productivity shock that can be correlated with ql_{it} and m_{it}. For the simplicity of the analysis, we assume the productivity shock follows a specific process that is $E(v_{1it}|v_{1it-1}) = \nu v_{1it-1}$ or equivalently, $f(\cdot) = \nu \cdot (\cdot)$.

By substituting v_{1it} into the production function, we obtain

$$y_{1it} = f_1(ql_{it}, k_{it}, m_{it}; \theta_1) + g(m_{it}, k_{it}) + e_{1it}. \tag{24}$$

All the explanatory variables are exogenous but problem is that a part of $f_1(ql_{it}, k_{it}, m_{it}; \theta_1)$ cannot be identified due to the multicolinearity with $g(m_{it}, k_{it})$. To overcome this difficulty in the identification, we exploit the Markov property of the productivity shock.

$$y_{1it} = f_1(ql_{it}, k_{it}, m_{it}; \theta_1) + \nu g(m_{it-1}, k_{it-1}) + b_{it} + e_{1it}. \tag{25}$$

Notice that the current productivity shock b_{it} is allowed to be correlated with the labor input ql_{it} and intermediate input m_{it}. These two variables are accordingly instrumented by its lagged values ql_{it-1} and m_{it-1}. Because m_{it-1} is already used as an explanatory variable, the system becomes under identified. We use m_{it-2} as well as ql_{it-2} and m_{it-2} as an additional instrumental variable to obtain identification.

All of the parameters in f_1 is identified by equation (25) with a moment condition $E(b_{it} + e_{1it}|k_{it}, ql_{it-1}, k_{it-1}, m_{it-1}, ..., ql_{i1}, k_{i1}, m_{i1}) = 0$. In addition, a part of f_1 in the equation (24) is identified with a moment condition $E(e_{1it}|ql_{it}, k_{it}, m_{it}, ql_{it-1}, k_{it-1}, m_{it-1}, ..., ql_{i1}, k_{i1}, m_{i1}) = 0$. Wooldridge (2005) recommends estimating these two equations jointly due to the efficiency reason. We estimate three equations including the wage equation (6) jointly.

In the actual implementation, the production function is given as: $f_1(\cdot) = \alpha \ln(l_{it}(1 + (\psi - 1)(\frac{l_f}{l})_{it})) + \beta \ln k_{it} + \gamma \ln m_{it} + ind\delta + year\tau$ and $g(\cdot)$ is approximated by the third order polynomial of k and m. The wage equation is given as: $\ln(wb_{it}/l_{it}) = \ln w_m + \ln(1 + (\lambda - 1)(\frac{l_f}{l})_{it}) + ind\delta + year\tau + e_{it}$.

The estimation is analogous to the pooled cross-sectional estimation of two equations.

The system of equations is

$$
\begin{aligned}
y_{1it} &= f_1(ql_{it}, k_{it}, m_{it}; \theta_1) + g(m_{it}, k_{it}; \theta_g) + e_{1it} && (26)\\
y_{1it} &= f_1(ql_{it}, k_{it}, m_{it}; \theta_1) + \nu g(m_{it-1}, k_{it-1}; \theta_g) + b_{it} + e_{1it} && (27)\\
y_{2it} &= f_2(ql_{it}; \theta_2) + u_{3it}. && (28)
\end{aligned}
$$

The vectors of instruments for each equation are $z_{1it} = [l_{it}, (\frac{l_f}{l})_{it}, k_{it}, m_{it}]$, $z_{2it} = [k_{it}, l_{it-1}, (\frac{l_f}{l})_{it-1}, m_{it-1}, k_{it-1}, l_{it-2}, (\frac{l_f}{l})_{it-2}, m_{it-2}, k_{it-2}]$, and $z_{3it} = [l_{it}, \frac{l_f}{l})_{it}]$. In addition, each instrument vector includes higher order polynomial and industry/year dummy variables. Each vector is $k_1 \times 1$, $k_2 \times 1$ and $k_3 \times 1$, respectively. By stacking up time series of each firm vertically, we form the vectors z_{1i}, z_{2i} and z_{3i}. Each vector is $(T-2) \times 1$. The instrument matrix is formed as:

$$
z_i = \begin{pmatrix} z_{1i} & 0 & 0 \\ 0 & z_{2i} & 0 \\ 0 & 0 & z_{3i} \end{pmatrix}, \tag{29}
$$

which is a $3(T-2) \times (k_1 + k_2 + k_3)$ matrix. Given the error vector:

$$
u_i = \begin{pmatrix} u_{1i} \\ u_{2i} \\ u_{3i} \end{pmatrix}, \tag{30}
$$

where $u_{1i} = e_{1i}$ and $u_{2i} = b_i + e_{1i}$. The moment condition $Ez'v = 0$ is used to identify the parameters.

The GMM estimator is defined as the solution to the following problem:

$$
\min_{\{\theta\}} (\sum_{i=1}^{N} z_i' u_i)' \hat{\Xi}^{-1} (\sum_{i=1}^{N} x_i' u_i), \tag{31}
$$

where Ξ is given as:

$$\hat{\Xi} = \frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix} x'_{1i}\hat{u}_{1i}\hat{u}'_{1i}x_{1i} & x'_{1i}\hat{u}_{1i}\hat{u}'_{2i}x_{2i} & x'_{1i}\hat{u}_{1i}\hat{u}'_{3i}x_{3i} \\ x'_{2i}\hat{u}_{2i}\hat{u}'_{1i}x_{1i} & x'_{2i}\hat{u}_{2i}\hat{u}'_{2i}x_{2i} & x'_{2i}\hat{u}_{2i}\hat{u}'_{3i}x_{3i} \\ x'_{3i}\hat{u}_{3i}\hat{u}'_{1i}x_{1i} & x'_{3i}\hat{u}_{3i}\hat{u}'_{2i}x_{2i} & x'_{3i}\hat{u}_{3i}\hat{u}'_{3i}x_{3i} \end{pmatrix}. \tag{32}$$

where $\hat{u}_{1i}$, $\hat{u}_{2i}$ and $\hat{u}_{3i}$ are the residuals from the GMM estimation that uses identity matrix as the weighting matrix.

Analogous to the previous exercise, the variance of this estimator is estimated as:

$$(\frac{1}{N}\sum_{i=1}^{N} z'_i \nabla_\theta u(y_i, x_i; \hat{\theta}))'\hat{\Xi}^{-1}(\frac{1}{N}\sum_{i=1}^{N} z'_i \nabla_\theta u(y_i, x_i; \hat{\theta})), \tag{33}$$

where $\nabla_\theta u(y_i, x_i; \hat{\theta})$ is the $3(T-2)\times(k_1+k_2+k_3)$ matrix defined as following:

$$\begin{pmatrix} \frac{\partial u_1}{\partial \theta_1}(y_{1i}, z_{1i}; \hat{\theta}_1, \hat{\theta}_g) & \frac{\partial u_1}{\partial \theta_g}(y_{1i}, z_{1i}; \hat{\theta}_1, \hat{\theta}_g) & 0 & 0 \\ \frac{\partial u_2}{\partial \theta_1}(y_{1i}, z_{2i}; \hat{\theta}_1, \hat{\theta}_g, \hat{\nu}) & \frac{\partial u_2}{\partial \theta_g}(y_{1i}, z_{2i}; \hat{\theta}_1, \hat{\theta}_g, \hat{\nu}) & \frac{\partial u_2}{\partial \nu}(y_{1i}, z_{2i}; \hat{\theta}_1, \hat{\theta}_g, \hat{\nu}) & 0 \\ 0 & 0 & 0 & \frac{\partial u_2}{\partial \theta_2}(y_{2i}, z_{3i}; \hat{\theta}_2) \end{pmatrix}. \tag{34}$$

5 Results

5.1 Cross-sectional and fixed effects estimation

The results of joint estimations of the production function (4) and the wage equation (6) appear in Table 3. The estimate of ψ indicates that the marginal product of female workers was 45 percent of that of male workers. However, the estimate of λ indicates that female workers earned 30 percent of what male workers earned. This joint estimation allows us to estimate $\psi - \lambda$ and the associated standard error. The estimate of the difference is 0.15, with a standard error of 0.01. Thus, the null hypothesis of no discrimination against women, which is $H_0 : \psi - \lambda = 0$, is strongly rejected. If we take the point estimates seriously, of the 70 percentage points of the wage differentials observed in data, 15 percentage points cannot be explained in the productivity difference between men and women. Thus 20 percent of the male-female wage differential (=0.15/0.70) arguably can be attributed to employers' discrimination.

The pooled least-squares estimator is a consistent estimator when each firm's time-constant heterogeneity is not correlated with the firm's input mix. This is a rather restrictive assumption because if there is firm-specific heterogeneity in production technology, then the optimal input mix is likely to be heterogenous. If each firm's technological heterogeneity is correlated with inputs, the pooled least-squares estimator is inconsistent. To work around this potential endogeneity issue, the production function and wage equations are estimated via an instrumental variable estimation using the mean deviation of the explanatory variables from each firm's mean. To ensure that the idiosyncratic error term of (4) is exogenous from each firm's mean of independent variables, the strict exogeneity of the error term, stated as (**??**), is required.

I also allow for the firm's time-constant heterogeneity in the wage equation (6) and estimate the following wage equation:

$$\ln((wb/l)_{it}) = \ln w_m + \ln(1 + (\lambda - 1)(\frac{l_f}{l})_{it}) + ind\delta + year\tau + d_i + e_{it}, \quad (35)$$

where d_i is time-constant, firm-level, unobserved heterogeneity that affects the per capita labor cost. This firm-level heterogeneity could arise due to the heterogeneity of workers' quality across firms. Even if firms operate in a perfectly competitive labor market and pay the same wage for an efficiency unit of labor, those firms that hire eligible workers pay a higher wage per capita. If the quality of workers in a specific firm is time-constant, then the effect of the heterogeneity of workers' quality is captured by d_i. If male workers are more skilled on average, then d_i and female worker proportion are negatively correlated. Accordingly, the pooled least-squares estimator of λ is downward inconsistent. However, an IV estimation that uses the mean deviation of independent variables from each firm's mean renders a consistent estimator.

Table 3 Columns (4) - (6) report the results of the joint IV estimation applied to production function and wage bill equation. The result in Column (4) shows that female workers' productivity relative to male workers' is 54 percent. Compared with the cross-sectional estimate reported in Column (1) of Table 3, this number is 10 percentage points higher, which implies that a_i and $\frac{l_{fit}}{l_{it}}$ are negatively correlated.

We can roughly test whether these two estimates are significantly different in a statistical sense by using Hausman statistics. Under the homoscedasticity assumption for the idiosyncratic error term, the non-linear, least-squares estimator is an efficient estimator under the null of no correlation between a_i and $\frac{l_{fit}}{l_{it}}$. Accordingly, Hausman statistics can be constructed for the difference of these two estimators as

$$H = (\hat{\psi}_{NLIV} - \hat{\psi}_{NL})^2 / (Var(\hat{\psi}_{NLIV}) - Var(\hat{\psi}_{NL})) \sim \chi(1). \tag{36}$$

for the two estimates, $\hat{H} = 30.01$ ($p < 0.000$), and thus we can conclude that the two estimates are different in a statistical sense. The implied positive correlation between a_i and $\frac{l_{fit}}{l_{it}}$ is consistent with the hypothesis that firms with a technological advantage hire fewer women because employers face less pressure of market competition and have room to indulge their preference for discrimination against females. This finding is consistent with the finding in Kawaguchi (2007). This earlier study found that firms with a persistent, high-profit factor tend to hire fewer women.

The parameter $\phi - \lambda$ is precisely estimated to be 0.016 ($s.e = 0.018$). This difference is economically negligible and statistically insignificant. At the first glance, these results seem to suggest that the relative wage of female workers compared to that of male workers reflects their relative productivity if they work for the same company. However, we should note that the fixed effects estimator is tenuous when the temporal shock and female proportion is highly correlated because the identification is completely based on within variation of female proportion. As Houseman and Abraham (1993) reports Japanese firms tend to adjust female labor more rapidly than male labor in response to the demand or technological shocks. If female proportion increases in response to the positive demand shock due to rather fast labor adjustment, the fixed effects estimator may be severely upward biased.

5.2 Proxy variable estimation

Considering the possible bias that fixed effects estimator might have suffered, we rather rely on the proxy variable estimation results reported in Table 4 to derive our conclusions. Because the proxy variable estimation utilizes two-period lagged variables as instrumental variables, the first two years of the sample period drops from the analysis sample. To confirm the change of the estimation result because of the change in the analysis sample, we estimate the model without proxy variable and report the results in Columns (1) - (3) in Table 4. The results with this restricted sample is very close to the results reported in Columns (1) -(3) in Table 3 and this implies the sample restriction significantly modifies the results.

Columns (4) - (9) in Table 4 report the results of proxy variable GMM estimation. The results in Columns (4) - (6) reports the results based on the two moments (27) and (28). The estimates of the relative productivity/wage are almost identical to the results without the proxy variables. The coefficients for labor and material inputs decreased and the coefficient for capital increases. These are the changes expected a priori because the demand/productivty shock is likely to have positive correlation with variable inputs (labor and material) and this results in the upward bias for these coefficients. This positive bias for labor and material coefficients are transmitted to the negative bias for the invariable input (capital) coefficient because variable inputs and the invariable input are positively correlated. The columns (7) - (9) report the results of joint estimation of (26), (27) and (28). Remember that the additional moment comes from the assumption that current labor input is exogenous from current productivity/demand shock. The estimate of the relative productivity of female workers to male workers declines slightly but the relative wage does not change in a meaningful way. The coefficients for labor and material increase and that for capital decreases; the results get closer to the results without proxy variables. These changes seem to suggest that the current labor input is not exogenous from the current productivity/demand shock. The the results reported in Columns (7) - (9) could be biased.

Overall, we judge the results based on the lagged moment for the production function and the current moment of the wage equation are most preferable. These estimates indicate that a typical female worker is as 47 percent productive as a typical male worker while a female worker receives about 31 percent of a male worker. The 16 percentage points gap between relative productivity and relative pay is consistent with the existence of the discriminatory employers and the marginal female worker is employed by the discriminatory employer.

6 Conclusion

In an attempt to explain the male-female wage differential, we estimated the relative marginal productivity and relative wage of female workers to male workers using panel data from Japanese firms. As a reduced form approach, we estimated the production using total wage bill as a measure of labor input.

Under the null hypothesis that male-female wage differential represents the productivity differential, female proportion in a firm should not affect the output because total wage bill fully captures the quality adjusted labor input. This null hypothesis was strongly rejected; the gender wage gap is larger than the gender productivity gap.

Given the results from the reduced form approach, we proceeded to structurally estimate the relative productivity and pay between male and female workers by jointly estimating the production function and the wage equation. Cross-sectional estimates showed that the marginal productivity of female workers was 44 percent of that of male workers, while the female wage was 31 percent of male workers' wage. These estimates were consistent with employers' discrimination against women. However, the IV estimates, which allowed for firm-level fixed effects indicated that both female workers' marginal productivity and wage were around 50 percent of the male workers'. Although this fixed effects results seem to be consistent with the absence of the discrimination, the female relative productivity may be severely upward biased due to the temporal correlation between female proportion and demand/productivity shocks in both production function and wage equation. Eventually, we rely on the recently developed methods to control for the demand/productivity shocks using proxy variables for the shocks. The most preferable estimates basically based on Levinsohn and Petrin (2003) indicate that the level estimates were not severely biased due to the demand/productivity shock and the relative female productivity is 47 percent while the relative female pay is 31 percent.

The results consistent with the employer taste discrimination against women is in line with the US evidence provided by Hellerstein et al. (1999). In the Japanese context, the implied employer taste discrimination against women is consistent with the robust findings that firms with higher proportion of female workers earn higher profit (Kodama et al. (2005), Sano (2005) and Kawaguchi (2007)). The reason why employer's taste discrimination does not disappear despite it lowers firm's profit level still remain as a puzzle.

Acknowledgement

This research was conducted as a part of the project of Research Institute of Economy, Trade and Industry. We thank the Ministry of Economy, Trade, and Industry and the Ministry of Public Management, Home Affairs, Posts, and Telecommunications of the Japanese government for releasing the micro data used in this study. We thank Hiroyuki Chuma, Gigi Foster, Hidehiko Ichimura, Hisako Ishii, Sadao Nagaoka, Kazuhiko Odaki, Kei Sakata, Tetsushi Sonobe, Kotaro Tsuru, Futoshi Yamuchi, Jeffrey Wooldridge, and the seminar participants at Hitotsubashi University, the University of South Australia, the Japan Labor Economics Conference, the FASID Hakone Conference, the International Conference on Panel Data, the Econometric Society Australasian Meeting, and Research Institute of Economy Trade Industr for their comments. Errors and omissions remain the author's responsibility.

References

Ackerberg, D., L. Benkard, S. Berry, and A. Pakes (2007). *Handbook of Econometrics, Volume 6*, Chapter Econometric Tools for Analyzing Market Outcomes. Elsevier. forthcoming.

Ackerberg, D., K. Caves, and G. Frazer (2006). Structural identification of production functions. Mimeo, UCLA.

Becker, G. S. (1971). *The Economics of Discrimination 2nd Edition*. University of Chicago Press.

Becker, G. S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics 3*(1), S33–S58.

Bowlus, A. J. (1997). A search interpretation of male-female wage differentials. *Journal of Labor Economics 15*(4), 625–657.

Crepon, B., N. Denison, and S. Perez-Duarty (2002). Wages, productivity and worker characteristics: A french perspective. mimeo, CREST-INSEE, Paris, France.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*(4), 1029–1054.

Hellerstein, J. and D. Neumark (2007). Production function and wage equation estimation with heterogeneous labor: Evidence from a new matched employer-employee data set. In E. R. Berndt and C. R. Hulten (Eds.), *Hard-To-Measure Goods Services: Essays In Honor Of Zvi Griliches*. NBER and University of Chicago Press.

Hellerstein, J. K. and D. Neumark (1999). Sex, wages and productivity: An empirical analysis of israeli firm-level data. *International Economic Review 40*(1), 95–123.

Hellerstein, J. K., D. Neumark, and K. R. Troske (1999). Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of Labor Economics 37*(2), 409–446.

Hellerstein, J. K., D. Neumark, and K. R. Troske (2002). Market forces and sex discrimination. *Journal of Human Resources 37*(2), 896–914.

Houseman, S. N. and K. G. Abraham (1993). Female workers as a buffer in the Japanese economy. *American Economic Review 83*(2), 45–51.

Ichino, A. and E. Moretti (2006). Biological gender differences, absenteeism and the earning gap. NBER Working Paper No. 12369.

Ilmakunnas, P., M. Mailranta, and J. Vainiomaki (2004). The roles of employer and employee characteristics for plant productivity. *Journal of Productivity Analysis 21*, 249–276.

Kawaguchi, D. (2007). A market test for sex discrimination: Evidence from Japanese firm-level panel data. *International Journal of Industrial Organization*.

Kodama, N., K. Odaki, and Y. Takahashi (2005). Female employment and firm performance (in Japanese). *JCER Journal 52*.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies 70*(2), 317–341.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Sano, S. (2005). Is male-female wage gap due to the taste discrimination (in Japanese). *Japanese Journal of Labor Studies 540*.

Wooldridge, J. M. (2005). On estimating firm-level production functions using proxy variables to control for unobservables. Mimeo, Michigan State University.

Table 1: Descriptive Statistics

Sample Period: 1992, 1995-2000

Variable Name	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Sample	All		Female Proportion > Median		Female Proportion <= Median	
Total Sales (Million Yen)	23399	233046	17064	183394	29735	273698
Wage Bill/Total Employment (Million Yen)	4.69	1.74	4.06	1.57	5.32	1.67
Total Employment	387.1	1597	344.3	1211	429.9	1906
Female Proportion	0.32	0.20	0.47	0.17	0.17	0.06
Fixed Assets (Million Yen)	6081	47442	3655	21845	8509	63346
Cost of Materials (Million Yen)	17516	217045	12770	175855	22262	251494
Industry						
Agriculture, Forestry, and Fishery	0.0004		0.0006		0.0004	
Mining and Construction	0.02		0.004		0.04	
Light Manufacturing	0.16		0.23		0.09	
Material Manufacturing	0.17		0.12		0.21	
Electronics and Machinery	0.22		0.19		0.25	
Public Utilities and Transportation	0.003		0.001		0.005	
Wholesale and Retail	0.41		0.44		0.39	
Finance, Insurance, and Real Estate	0.001		0.001		0.0006	
Service	0.02		0.014		0.02	
Number of Observations	177868		88939		88929	
Number of Firms	37131		19427		17704	

Table 2: Reduced Form Production Function Estimation

Dependent Variable: Log (Total Sales)

	(1)	(2)
Log (Wage Bill)	0.371	0.370
	(0.001)	(0.001)
Log (Fixed Asset)	0.032	0.032
	(0.001)	(0.001)
Log (Material Cost)	0.604	0.606
	(0.001)	(0.001)
Female Ratio	-	0.079
		(0.003)
Constant	1.050	1.006
	(0.025)	(0.025)
Observations	177868	177868
R-squared	0.967	0.967

Note: Standard errors are reported in the parenthesis. One digit industry dummy variables and year dummy variables are included in the regression but the coefficients are suppressed.

Table 3: Non-linear Pooled Cross-Sectional and Fixed Effects GMM Estimation of the Production Function and Wage Equation

Year: 1992, 1995-2000 Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
	Production Function	Wage Equation	φ - λ	Production Function	Wage Equation	φ - λ
Relative Productivity (φ), Relative Wage (λ)	0.446	0.297	0.149	0.544	0.528	0.016
	(0.011)	(0.004)	(0.010)	(0.021)	(0.013)	(0.018)
Log (Employment)	0.379	-	-	0.362	-	-
	(0.007)			(0.009)		
Log (Fixed Asset)	0.028	-	-	0.013	-	-
	(0.002)			(0.002)		
Log (Material Cost)	0.627	-	-	0.502	-	-
	(0.004)			(0.007)		
Instruments		Level			Within Transformation	

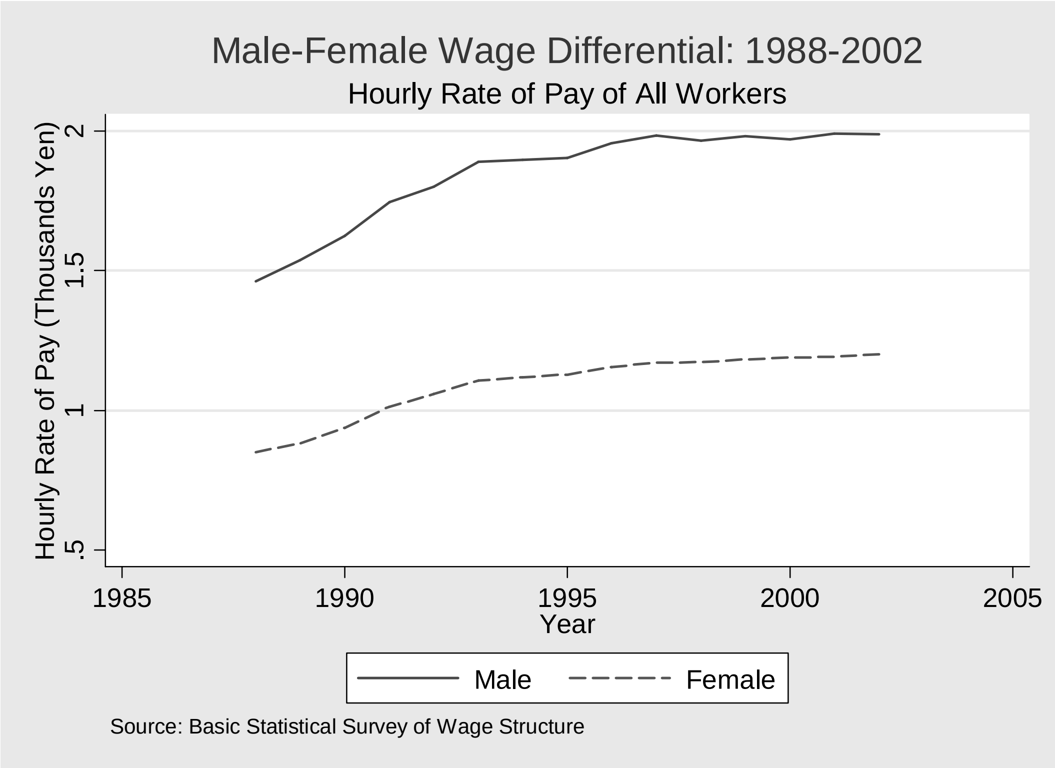
Note: Weighting matrix allows for panel clustering, heteroskedasticity and autocorrelation within a panel unit. Standard errors are in parentheses. Year and industry dummies are included but the coefficients are suppressed. N=177,868.

Table 4: Non-linear Pooled Cross-Sectional GMM Estimation of the Production Function and Wage Equation with Proxy Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Production Function	Wage Equation	φ - λ	Production Function	Wage Equation	φ - λ	Production Function	Wage Equation	φ - λ
Relative Productivity(φ)	0.472	0.307	0.165	0.470	0.307	0.163	0.447	0.304	0.143
Relative Wage (λ)	(0.014)	(0.004)	(0.012)	(0.017)	(0.004)	(0.016)	(0.013)	(0.004)	(0.012)
Log (Employment)	0.368	-	-	0.271	-	-	0.298	-	-
	(0.008)			(0.011)			(0.006)		
Log (Fixed Asset)	0.034	-	-	0.052	-	-	0.035	-	-
	(0.003)			(0.003)			(0.003)		
Log (Material Cost)	0.634	-	-	0.565	-	-	0.665	-	-
	(0.004)			(0.024)			(0.011)		
Instruments		Level			Level			Level	
Levinshon-Petrin-Wooldridge		No			Lagged Moment Only			Current + Lagged Moments Together	

Note: Standard errors are in parentheses. Year and industry dummies are included. N=108,749. The third order polynomials of material cost and fixed asset are used as a proxy variable for the demand/productivity shock.

Figure 1: The Male-Female Wage Differential in Japan



Note: The hourly rate of pay is the weighted average of the hourly rate of pay of full-time workers (*ippan rodosha,* in Japanese), which is calculated as the fixed monthly salary (*Shoteinai Kyuyo Gaku* in Japanese) divided by the fixed monthly hours of work (*Shoteinai Jitsu Rodo Jikan* in Japanese), and the hourly rate of pay of part-time workers using the number of workers as a weight.

Figure 2: Distribution of Firm-Level Average Female Proportion

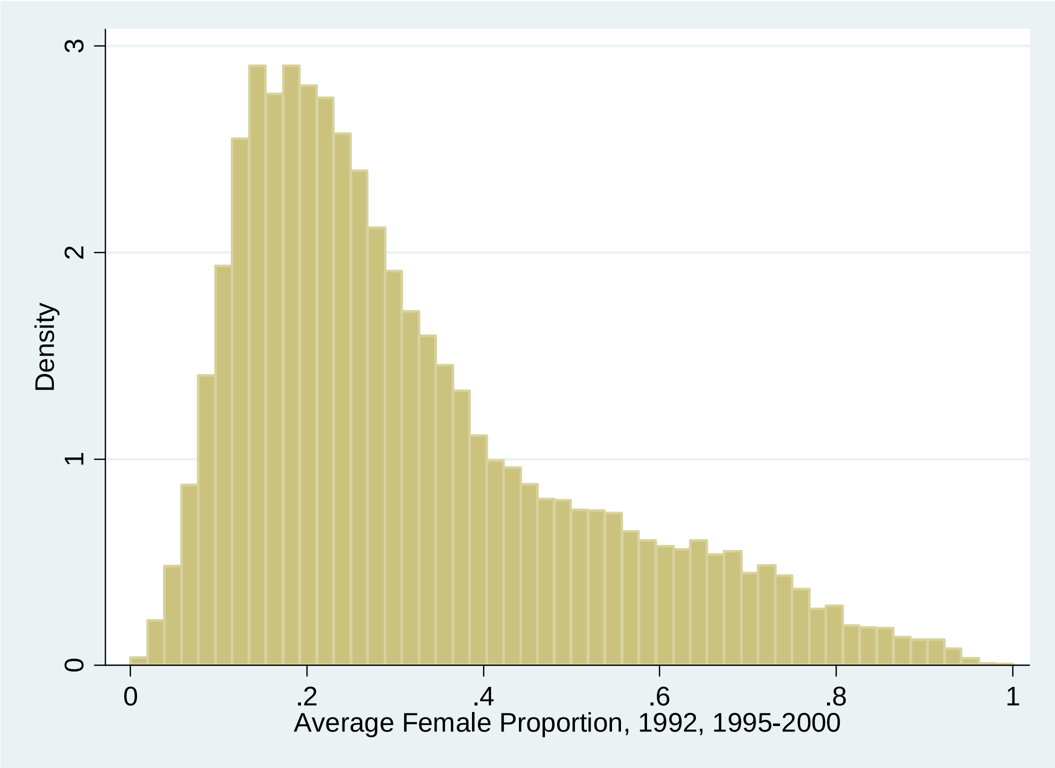
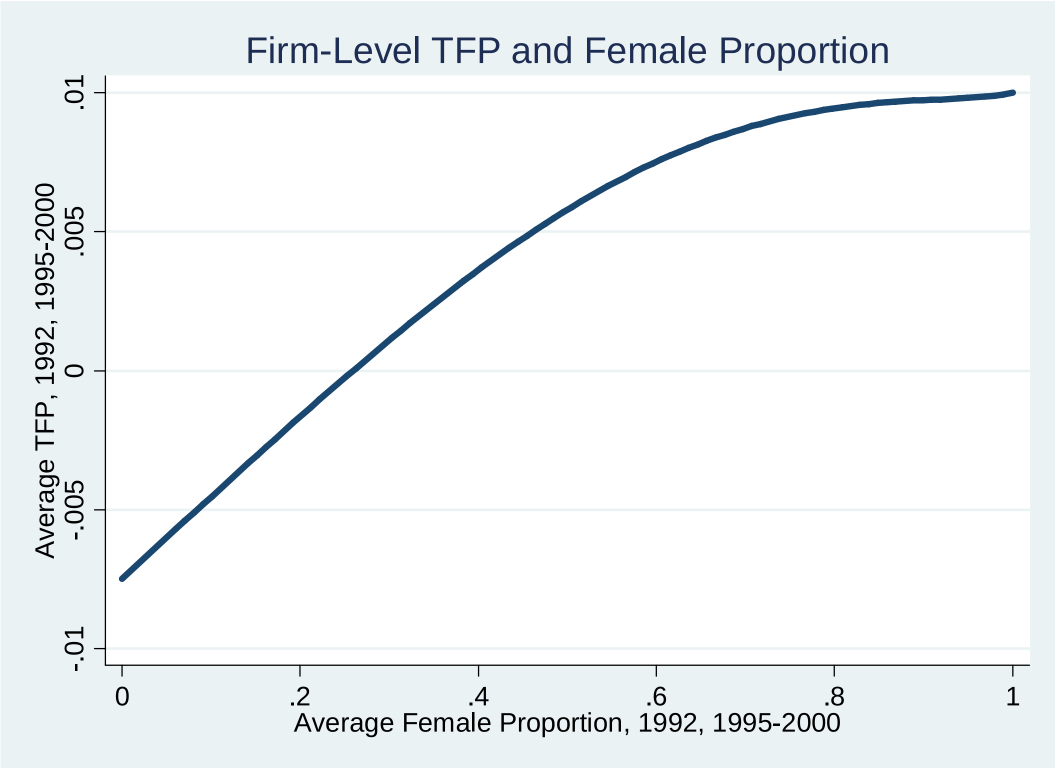


Figure 3: Relationship between Firm-Level Average TFP and Female Proportion



Note: Firm-level average of residuals from Cobb-Douglas production function (Table 2 Column 1) is regressed upon firm-level average of female proportion. Kernel regression with bandwidth of 0.3 and Gaussian kernel is applied.