Differences in Decline: Quantile Regression of Male-Female Earnings Differential in Malaysia

Goy Siew Ching
Univerisiti Malaysia Sarawak

Gender inequality in the labour market outcomes

- The economic structural transformation has led to unprecedented numbers of women participation in the economic activity.
- The labour market continues to display significant disparities in the opportunities presented to men and women in the workplace.

Representation Ratio in Eight Major Non-Agricultural Occupations by Gender in Malaysia, 2008

Occupational Category	2008
Legislators, senior officials and managers	0.58
Professionals	1.07
Technicians and associate professionals	0.97
Clerical workers	1.70
Services and shop and market sales workers	1.14
Craft and related trade workers	0.42
Plant and machine-operators and assemblers	0.87
Elementary occupations	1.01

- Household Income Survey Data 1995 - full-time employed male workers earn 1.3 times more than their female counterparts (Nagaraj *et al*., 2002).
- Malaysian Household Survey data 1984, 1989 and 1997 - Millanovic (2001) reported that the female-male real hourly wages ratio stood as 0.64 in 1984 and improved to 0.82 in 1989. The gap remained noticeable in 1997 at around 28%.
- Other studies in Malaysia also revealed that women consistently earn less than men (see Chua, 1984; Chapmen and Harding, 1986; Lee and Nagaraj, 1995; Ariffin et al., 1996; Mohamad-Nor, 1997;Low and Goy, 2006 and Fernandez ,2006).

- A substantial fraction of the persistent gender earnings gap is attributed to the unequal treatment of workers of equal productivity.

- In literature, when a wage gap remains after controlling for differences in productivity-link characteristics between sexes, it reflects discriminatory constraints in the labour market for women.
- This interpretation, however, is perceived as a less than wholly satisfactory measure.

Empirical measurement of wage discrimination

In analysing gender wage gap, research efforts tend to separate the gap into 2 components: differences in characteristics and differences in the rewards to those characteristics.

- **Parametric – OLS**

  Oaxaca (1973), Blinder (1973); Oaxaca and Ransom (1994), Juhn *et al*. (1991, 1993), Blau and Khan (1997, 2003)

- **Semiparametric – Quantile Regression**

  Gardeazabal and Ugidos (2005), Machado and Mata (2005), Felgueroso *et al*. (2007), Nicodemo (2009)

- **Nonparametric**

  DiNardo *et al*. (1996)

Motivations of this study

- The existing wage gap evidence in Malaysia focuses solely on OLS decomposition. It focuses the gap at the mean, it ignores an important part of the story, in that it heroically assumes the gap to be constant at other points of wage distribution.

- Attention has been shifted to quantile regression in developed countries, this line of research has yet been explored in the context of the Malaysian labour market.
- Official statistics of income distribution in Malaysia across the whole distribution display only according to ethnic groups and states.
- The mandatory minimum wages policy will be implemented for the first time in Malaysia in January 2013.

Data

- The data are drawn from the Malaysia Family and Population Survey (MFPS), a cross-section survey conducted separately at two points in time, 1994 and 2004.

Table 1: The Distribution of Respondent by Gender and Type of Survey, 1994-2004

Year of Survey and Type	Male	Female
MPFS 3 1994	4238 (4097)	4444 (2020)
Married women aged 15-49, 2004		3693 (1726)
Spouses of married women, 2004	2626 (2504)	
Youths aged 13-24, 2004	268 (199)	227 (164)
Citizens aged 50 and above, 2004	354 (219)	927 (316)
Unmarried aged 25-49, 2004	370 (318)	251 (205)

Note: The numbers in parentheses represent respondents who are in the 16-64 age groups and in paid employment.

Individual Variables	1994		Restricted Sample 2004		Unrestricted Sample 2004	
	Men	**Women**	**Men**	**Women**	**Men**	**Women**
Experience	24.29 (10.37)	21.16 (9.64)	24.58 (9.85)	20.45 (9.71)	23.75 (12.06)	21.80 (12.98)
Age	39.03 (8.54)	35.58 (7.30)	40.53 (8.33)	36.91 (7.58)	39.63 (10.34)	37.94 (10.25)
Secondary	0.495 (0.50)	0.450 (0.50)	0.596 (0.49)	0.526 (0.50)	0.586 (0.49)	0.489 (0.50)
Beyond secondary	0.114 (0.32)	0.124 (0.33)	0.176 (0.38)	0.255 (0.44)	0.174 (0.38)	0.252 (0.43)
Metropolitan	0.342 (0.47)	0.398 (0.49)	0.367 (0.48)	0.405 (0.42)	0.369 (0.48)	0.407 (0.49)
Urban	0.204 (0.403)	0.205 (0.404)	0.222 (0.49)	0.243 (0.43)	0.215 (0.41)	0.243 (0.43)
Bumiputra	0.609 (0.49)	0.575 (0.49)	0.726 (0.45)	0.708 (0.46)	0.703 (0.46)	0.671 (0.47)
Chinese	0.253 (0.43)	0.255 (0.44)	0.171 (0.38)	0.177 (0.38)	0.194 (0.40)	0.216 (0.41)
Married					0.846 (0.36)	0.850 (0.36)
Female		0.331 (0.47)		0.408 (0.49)		0.426 (0.50)
Monthly earnings (RM)	984.98 (1051.84)	616.82 (714.47)	1625.55 (1500.81)	1121.55 (1077.43)	1524.53 (1436.64)	1099.22 (1157.36)
Log monthly earnings	6.5736 (0.77)	6.0440 (0.87)	7.1049 (0.74)	6.6528 (0.89)	7.0379 (0.74)	6.6129 (0.90)

Gender Earnings Gap across Earnings Distribution

Figure 6.1: Unconditional Gender Earnings Gap, 1994

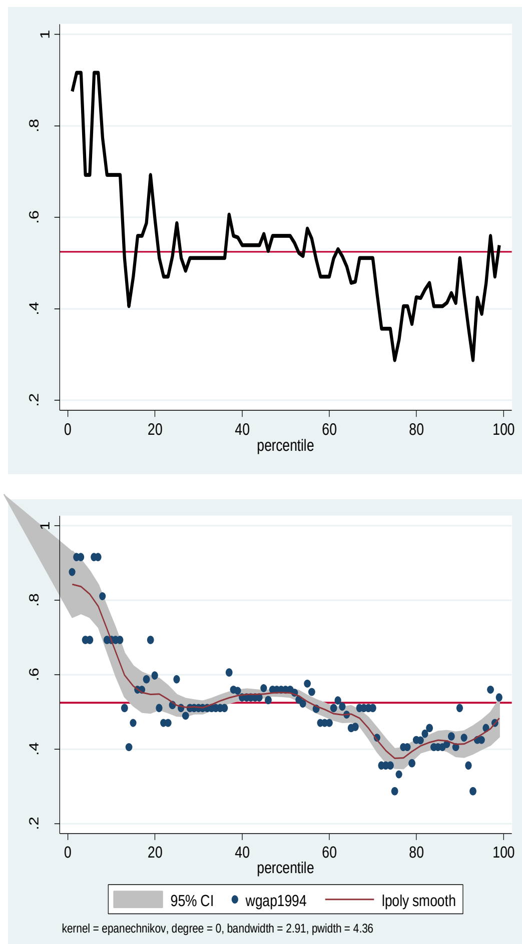


Figure 6.2: Unconditional Gender Wage Gap, Restricted Sample 2004

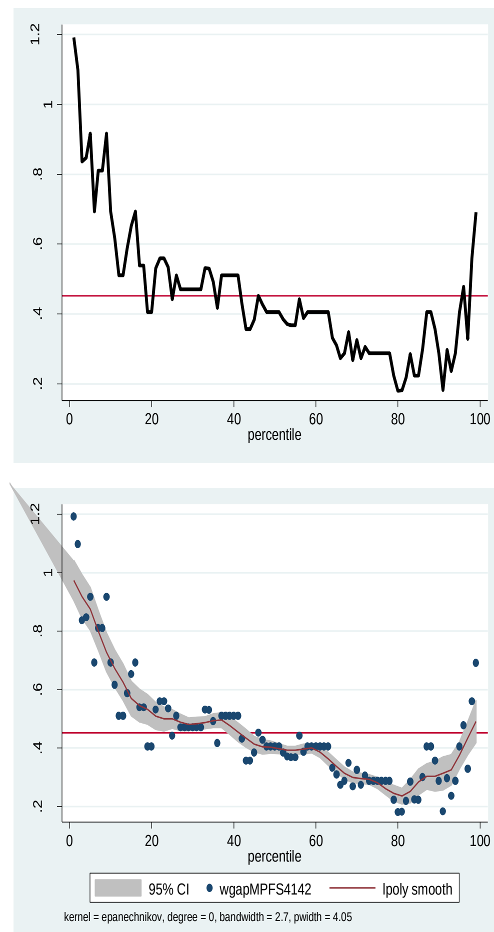
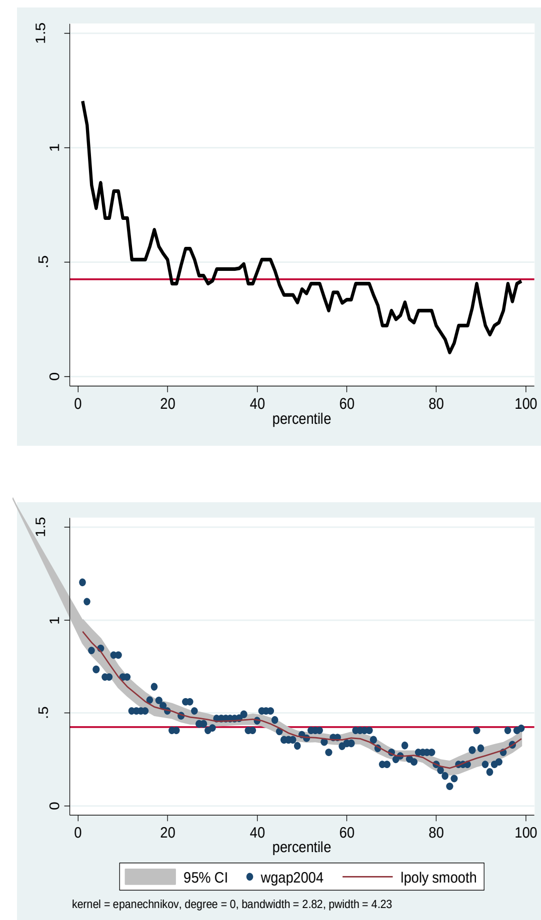


Figure 6.3: Unconditional Gender Wage Gap, Unrestricted Sample 2004



Quantile Regression and the Wage Gap Decomposition

- Following Koenker and Bassett (1978) and Buchinsky (1998), the wage equation at different quantiles of wage distribution is specified as $lnw_{ij} = X_{ij}\beta_\theta + \mu_{ij\theta}$ $\quad with\ Quant_\theta(lnw_{ij}|X_{ij}) = X_{ij}\beta_\theta$

  the distribution function of the error term for the θth quantile is left unspecified, but with the assumption that $Quant_\theta(\mu_{i\theta}|X_i) = 0$

- the vector of coefficients is estimated by minimising the sum of *absolute* value of the weighted residual, ranging from 0<θ<1 :

$$\min_{\beta_\theta}\Big(\sum_{i:lnw_i \geq X_i\beta_\theta} \theta\,|lnw_i - X_i\beta_\theta| + \sum_{i:lnw_i < X_i\beta_\theta} (1-\theta)|lnw_i - X_i\beta_\theta|\Big) \qquad (2)$$

Correction for Sample Selection based on Buchinsky (1998)

- In the presence of sample selection, the conditional quantile of the observed wage equation is specified as:

$$Quant_\theta(lnw_i|X_i, EMP = 1) = X_i\beta_\theta + Quant_\theta(\mu_{i\theta}|X_i, EMP = 1) \qquad (3)$$

where $Quant_\theta(\mu_{i\theta}|X_i, EMP = 1) \neq 0$.

- Assuming the quantity is a function of a known index g, the observed wage specification can be expressed as:

$$Quant_\theta(lnw_i|X_i) = X_i\beta_\theta + h_\theta(g) + v_\theta \qquad (4)$$ with $$Quant_\theta(v_{i\theta}|X_i, EMP = 1) = 0.$$

- Buchinsky suggests a series of estimators and we restrict to two selection bias terms (namely the linear bias term and its square). These are then used to compute the unknown form of sample selection in the quantile regression in the second stage.

Quantile wage decomposition

- The wage differential at various quantiles between men and women at the θth unconditional quantiles can be expressed as:

$$\begin{aligned}\overline{lnW}_{m\theta} - \overline{lnW_{f\theta}} &= \beta_{m\theta}\left[E(X_m | lnW_m = lnW_{m\theta}) - E\left(X_f \middle| lnW_{fm} = lnW_{f\theta}\right)\right] \\ &+ \left(\beta_{m\theta} - \beta_{f\theta}\right)E\left(X_f \middle| lnW_{fm} = lnW_{f\theta}\right) + E\left(u_{m\theta} | lnW_m = lnW_{m\theta}\right) \\ &- E\left(u_{f\theta} \middle| lnW_f = lnW_{f\theta}\right) \qquad (6)\end{aligned}$$

- Follow the procedure developed by Felgueroso *et al*. (2007) which based on Machado and Mata (2005) to determine the counterfactual female wage density.

  $\mathbf{lnw^c_{fi} = X_{if}\beta_{m\theta}}$ if women's labour market characteristics are rewarded by the same prices as are paid to men at each quantile; or $\mathbf{lnw^c_{fi} = X_{im}\beta_{f\theta}}$.

- The decomposition of the difference between men and women wage densities can be specified as:

$$Q_\theta\left(\overline{lnW_m}\right) - Q_\theta\left(\overline{lnW_f}\right) = \left[Q_\theta\left(X_{im}\hat{\beta}_{m\theta}\right) - Q^c_\theta\left(X_{if}\hat{\beta}_{m\theta}\right)\right] + \left[Q^c_\theta\left(X_{if}\hat{\beta}_{m\theta}\right) - Q_\theta\left(X_{if}\hat{\beta}_{f\theta}\right)\right] + \text{resd.} \quad (7)$$

- While the application of the Buchinsky method is instructive, it comes at a cost – estimation of the first stage of the model (using the Klein and Spady estimator) is highly computer-intensive.

- we simply follow the simpler approach of Manquilef-Bachler *et al*. (2009) – that is we estimate the selection terms under a probit model in the first stage.

Findings

Table 3: Quantile Regression With and Without Correction for Selection Bias, 1994-2004

Panel A:	Female 1994			Male 1994	
Variable	**Correction for Selection Bias[a]**	**Correction for Selection Bias[b]**	**Without Correction for Selection Bias**	**Correction for Selection Bias[b]**	**Without Correction for Selection Bias**
Metro	0.3452 (0.0383)	0.3389 (0.0418)	0.3378 (0.0364)	0.4030 (0.0267)	0.3923 (0.0181)
Urban	0.1189 (0.0446)	0.1235 (0.0445)	0.1141 (0.0425)	0.2040 (0.0394)	0.1855 (0.0201)
Bumiputra	-0.0361 (0.0549)	-0.0409 (0.0430)	-0.0507 (0.0439)	-0.1441 (0.0531)	-0.1291 (0.0225)
Chinese	0.1977 (0.0507)	0.2019 (0.0446)	0.1817 (0.0483)	0.3986 (0.0357)	0.4026 (0.0247)
Secondary	0.4965 (0.0459)	0.4942 (0.0491)	0.5144 (0.0437)	0.3766 (0.0289)	0.3747 (0.0195)
Beyond secondary	1.1369 (0.0688)	1.3617 (0.0692)	1.3876 (0.0636)	1.1098 (0.0587)	1.1065 (0.0303)
Experience	0.0292 (0.076)	0.0278 (0.0081)	0.0298 (0.0072)	0.0290 (0.0134)	0.0313 (0.0031)
Experience square	-0.0006 (0.0002)	-0.0005 (0.0002)	-0.0006 (0.0002)	-0.0005 (0.0003)	-0.0006 (0.0001)
Self-selection correction	0.0073 (0.0133)	-0.1670 (0.1521)		-1.2503 (2.0129)	
Self-selection correction square	-0.0001 (0.0004)	0.0819 (0.1033)		2.2789 (2.7004)	
Constant	5.1497 (0.1235)	5.2932 (0.1146)	5.2342 (0.0925)	5.7883 (0.2060)	5.7231 (0.0484)
Pseudo R^2	0.2207	0.2206	0.2202	0.2586	0.2582
N	2018			4077	

Panel B:	Female 2004 Restricted Sample		Male 2004 Restricted Sample	
	Correction for Selection Bias[b]	**Without Correction for Selection Bias**	**Correction for Selection Bias[b]**	**Without Correction for Selection Bias**
Metro	0.3341 (0.0479)	0.3390 (0.0417)	0.4894 (0.0349)	0.4694 (0.0270)
Urban	0.1874 (0.0566)	0.1790 (0.0461)	0.2090 (0.0397)	0.1983 (0.0298)
Bumiputra	-0.0734 (0.0696)	-0.0851 (0.0557)	-0.2222 (0.0526)	-0.1846 (0.0373)
Chinese	0.0718 (0.0834)	0.0605 (0.0652)	0.3247 (0.0548)	0.3155 (0.0435)
Secondary	0.5354 (0.0709)	0.5390 (0.0594)	0.5366 (0.0726)	0.4455 (0.0324)
Beyond secondary	1.4669 (0.0834)	1.4701 (0.0697)	1.2459 (0.0810)	1.1445 (0.0422)
Experience	0.0243 (0.0080)	0.0245 (0.0078)	0.0307 (0.0104)	0.0445 (0.0050)
Experience square	-0.0004 (0.0002)	-0.0004 (0.0002)	-0.0003 (0.0003)	-0.0008 (0.0001)
Self-selection correction	0.0190 (0.1888)	5.6577 (0.1007)	-1.2891 (0.8763)	
Self-selection correction square	0.0036 (0.1294)		0.5944 (0.6309)	
Constant	5.6413 (0.1317)		6.0100 (0.0916)	5.9514 (0.0769)
Pseudo R^2	0.2604	0.2602	0.2639	0.2631
N	1724		2503	

Panel C:	**Female 2004 Unrestricted Sample**		**Male 2004 Unrestricted Sample**	
	Correction for Selection Bias[b]	**Without Correction for Selection Bias**	**Correction for Selection Bias[b]**	**Without Correction for Selection Bias**
Metro	0.3320 (0.0422)	0.3252 (0.0348)	0.4686 (0.0288)	0.4653 (0.0325)
Urban	0.1511 (0.0506)	0.1481 (0.0384)	0.1998 (0.0373)	0.1794 (0.0365)
Bumiputra	-0.0677 (0.0459)	-0.0647 (0.0462)	-0.2269 (0.0446)	-0.1947 (0.0450)
Chinese	0.2031 (0.0547)	0.2025 (0.0525)	0.2217 (0.0554)	0.2358 (0.0516)
Secondary	0.4914 (0.0557)	0.4895 (0.0465)	0.4948 (0.0698)	0.3964 (0.0389)
Beyond secondary	1.3481 (0.0678)	1.3461 (0.0553)	1.1670 (0.0788)	1.0645 (0.0513)
Experience	0.0311 (0.0046)	0.0310 (0.0042)	0.0253 (0.0079)	0.0377 (0.0047)
Experience square	-0.0006 (0.0001)	-0.0006 (0.0001)	-0.0003 (0.0002)	-0.0007 (0.0001)
Married	0.0397 (0.0628)	0.0445 (0.0450)	0.0752 (0.1002)	0.2413 (0.0459)
Self-selection correction	-0.0387 (0.1608)		-0.8123 (0.4381)	
Self-selection correction square	0.0331 (0.1128)		0.4428 (0.2879)	
Constant	5.6209 (0.0985)	5.6166 (0.0759)	6.1220 (0.1401)	5.8938 (0.0739)
Pseudo R^2	0.2695	0.2695	0.2678	0.2670
N	2379		3204	

Figure 6.7: Kernel Density Earning Differential between Men and Women if Women had Men's Characteristics and Women's Returns, 1994

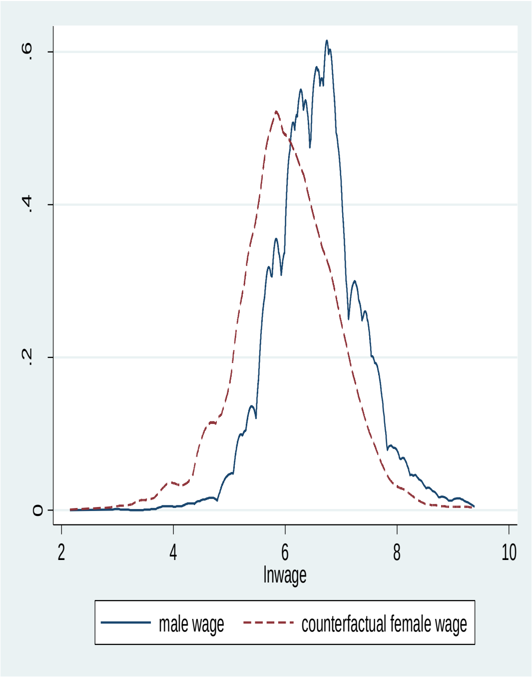


Figure 6.8: Kernel Density Earning Differential between Men and Women if Women had Men's Characteristics and Women's Returns, Restricted Sample 2004

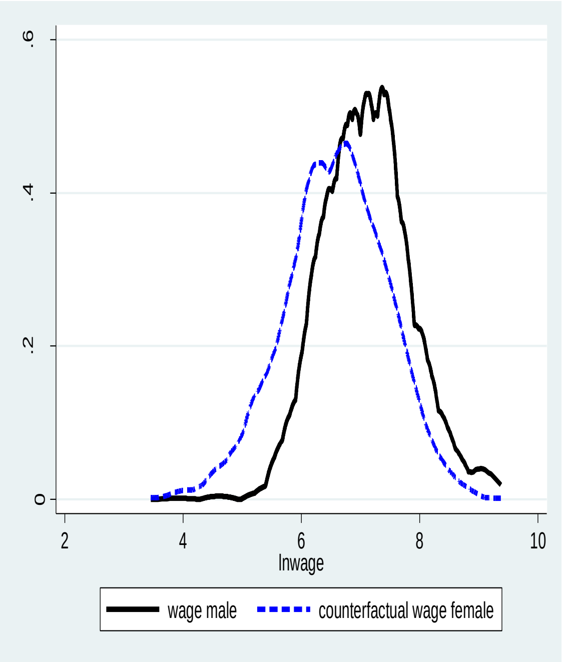
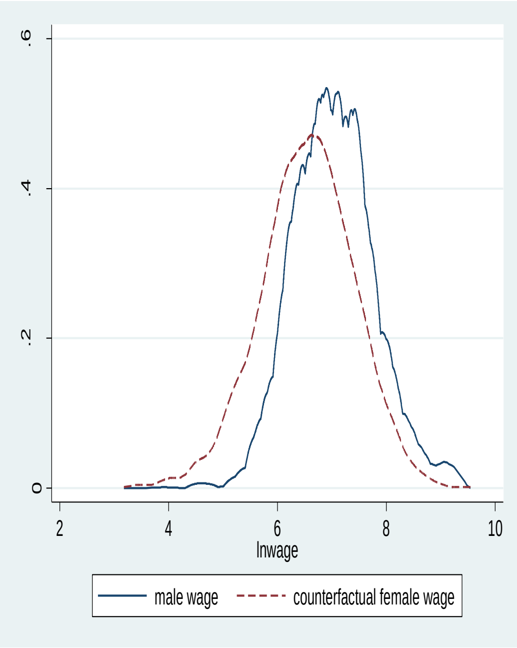
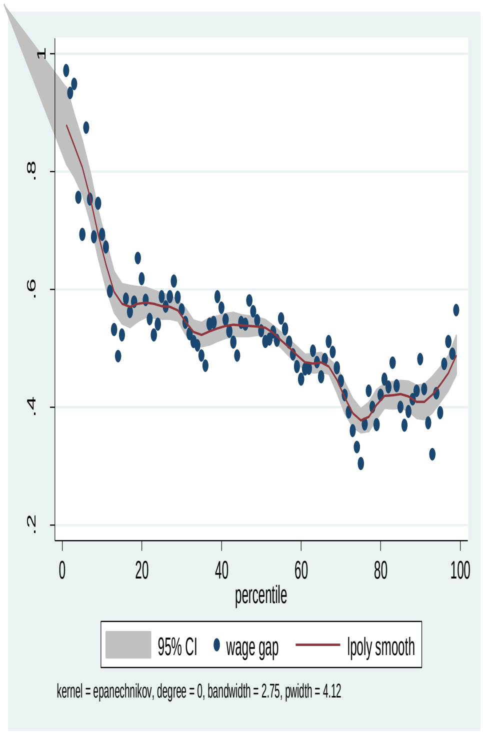


Figure 6.9: Kernel Density Earning Differential between Men and Women if Women had Men's Characteristics and Women's Returns, Unrestricted Sample 2004

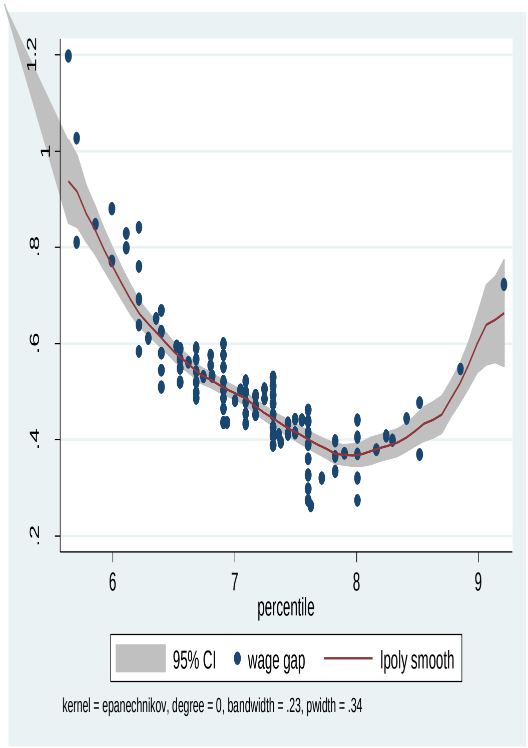


Wage Differential between Men and Women that had Men's Characteristics and Paid Like Women

1994



Restricted sample



Unrestricted sample

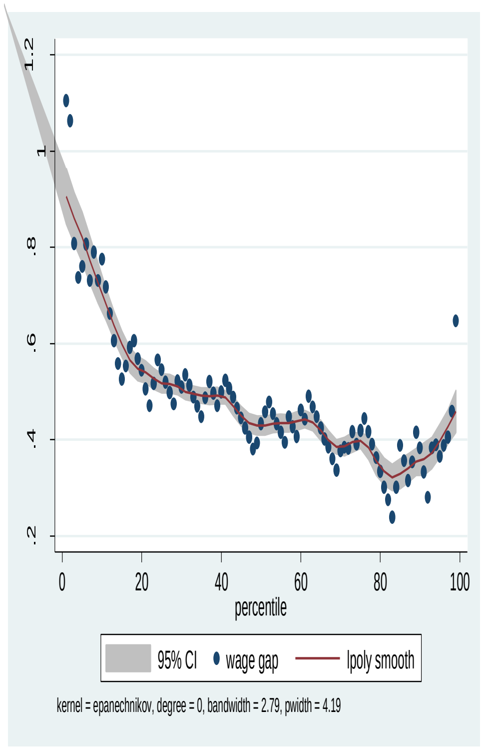


Table 4: Quantile Wage Decomposition Based on Counterfactual Density, 1994-2004

	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile	Mean (OLS)
Panel A: 1994						
Differences in characteristics	-0.0275	-0.0183	-0.0033	-0.1084	0.0914	-0.0218
Differences in rewards	0.7206 (0.6931)	0.6069 (0.5878)	0.5629 (0.5298)	0.3961 (0.3041)	0.4194 (0.4813)	0.5514
Total raw wage gap	0.6931	0.5878	0.5596	0.2877	0.5108	0.5296
Panel B: restricted sample 2004						
Differences in characteristics	-0.0385	-0.1081	-0.0340	0.0030	-0.0894	-0.0500
Differences in rewards	0.7316 (0.7601)	0.5499 (0.5203)	0.4394 (0.4777)	0.2847 (0.4149)	0.3771 (0.3200)	0.5020
Total raw wage gap	0.6931	0.4418	0.4054	0.2877	0.2877	0.4520
Panel C: unrestricted sample 2004						
Differences in characteristics	0.0309	-0.0430	-0.0661	-0.0075	-0.0016	-0.0483
Differences in rewards	0.6622 (0.7747)	0.6026 (0.5457)	0.4491 (0.4335)	0.2439 (0.4197)	0.3260 (0.3829)	0.4731
Total raw wage gap	0.6931	0.5596	0.3830	0.2364	0.3102	0.4248

Note: Figures in parentheses are computed based on counterfactual women density if they had men's characteristics but were paid like women.

Conclusions

- We have found that the gender earnings gap declines as we move up the wage distribution. The gap is bigger at the bottom of earnings distribution.

- Most of the earnings gap is explained by differences between the price that the market pays to male and female endowments. But the extent of the price effect is larger at the bottom end of the distribution than at the top.

- In terms of policy, these findings suggest that the focus should be on finding ways to improve the returns to characteristics earned by women at the bottom end of the distribution.
- One possibility would be to require large employers to undertake job evaluations and to remunerate their workers, regardless of gender. Another possibility is to introduce affirmative action.

THANK YOU