The Application of Quantile Regression in Analysis of Gender Earnings Gap in China

Fang Wang*
Master's Degree Candidate
Department of Economics
East Carolina University

June 27th, 2002

Abstract

The goal of this paper is to analyze the gender differentials in the returns of various characteristics and gender earnings differentials along the conditional earnings distribution in Chinese case. The data we use is the Chinese Household Income Project for the year 1988 and 1995. We use the modified standard Mincerian wage equation to estimate the marginal earnings distribution as a function of workers' characteristics applying quantile regression approach. We then break down the unadjusted gender earnings gap into the "explained part" (due to differential in characteristics) and the "unexplained part" (due to the different returns to the endowments) by using Oaxaca decomposition method. We found that the returns to experience, childbirth, marriage and education do not vary much at different quantiles for 1988, but show substantial variation in 1995. In contrast, little difference in discrimination exists across quantiles in 1995 while significant variation in 1988. The results also suggest that low earnings women suffer greater discrimination under the socialist conditions of 1988 than under the emerging market conditions of 1995; higher earnings women are faced up with greater discrimination in 1995 than in 1998.

* The author would thank Dr. John Bishop and Dr. Frank Luo for their input and guidance with all respects of this paper.

I. Introduction

Gustafsson and Li asked two important questions, "How are women faring in the transformation from planned economies to market economies that began in the 1980s?" and "Is the gender gap increasing or decreasing?" They found that in 1988 the average gender earnings gap was 15.6%, and even smaller among the youngest and higher educated workers. They also found that the gender differential in earnings increased to 17.5 percent in 1995. They argued that a substantial and increasing part of the average earnings gap could not be explained by differences in characteristics, but they also pointed out that it was not clear that the differentials in coefficients were due to discrimination. This paper extends their work by using the quantile regression to model the marginal earnings distribution as a function of workers' characteristics at different earnings levels, as well as combining the standard decomposition technique with quantile regressions to determine the gender earnings gap component along the earning distribution.

Quantile regression is a statistical technique to estimate and to conduct inference of conditional quantile functions – models in which quantiles of the conditional distribution of the response variable are expressed as functions of observed covariates. Using for the quantile regression methods, Garcia, Hernandez, and Lopez-Nicolas (2001) found that the absolute wage gap of the Spanish workers increases with the pay scale. Moreover, the decomposition of the wage gap in the spirit of the Oaxaca (1973) methodology shows that the "unexplained part" is greater at higher wage levels. They conclude that there is evidence that more frequent and greater discrimination exists for women at high wage levels. Claudio Montenegro (2001) used quantile regression approach to analyze the gender differential in the returns to education, the return to experience and gender wage differentials in Chile. The results show systematic differences in the returns to education and to experience by gender along the conditional wage distribution. The paper also found that the unexplained wage differential is higher in the upper quantiles.

The contribution of this paper is that we estimate the earnings distribution at different earnings scales for males and females using quantile regression approach and examine gender earnings gaps at various points in the earnings distribution. We suspect that whether the gender earnings gaps are the same across different quantiles of the earnings distributions in China. The results show that low earnings women suffer greater discrimination under the socialist conditions of 1988 (0.90 adjusted earnings gap ratio) than under the emerging market conditions of 1995 (0.67 adjusted earnings gap ratio). In contrast, the higher earnings women suffer greater discrimination in 1995 (0.74) than in 1988 (0.55).

The rest of the paper is organized as follows: section II discusses the data and gives a brief introduction to the methodology of quantile regression and wage decomposition. Section III discusses the empirical results. Section IV gives the conclusion.

II. Data and Methods

(i) The data

We obtain our estimates from the Chinese Household Income Project, 1988 and 1995. The surveys were carried out by Carl Riskin, Zhao Renwei and Li Shi and intended to measure and estimate the distribution of income in both rural and urban areas of the People's Republic of China. The data collection of both surveys consists of two distinct samples of the urban and rural population of the People's Republic of China, which were selected from significantly larger samples drawn by the State Statistical Bureau.

We restrict our sample into the urban area population. We only focus on those whose age is between 18 years and 60 years, and those with positive earnings as well as some labor market experience during the survey year. We define the earnings as the sum of regular wages, floating wages, all kinds of bonuses, subsidies, cash income and allowances. There are in total 17,558 and 11,927 observations in the 1988 data set and 1995 data set, respectively.

Table 1 shows the descriptive statistics of main variables by gender, as well as the male-female ratio. The summary of the 1988 dataset is presented in Table 1-a. It can be shown that on average males earn 19% more than females. Males have around two years experience and 0.83 year education more than females. Within the occupations, a much higher proportion of the males are factory managers, government officials, and office workers than that of females, while females are more prevalent in manual labor jobs. We can also see that males are more likely to be distributed in certain economic sectors such as agriculture, mining, construction, transport and communications, and science and technology; on the other hand, females are more likely found in trade and restaurant, personal services, health and social welfare, and education and cultural area. Further, the region that has the highest proportion of males is Beijing, followed by Shanxi and Gansu.

Table 1-b presents the statistics of 1995 dataset. Males' earnings in 1995 are 20% higher than females, 1 percentage point higher than in 1988. The average experience for both genders is roughly one year less than in 1988, while the male-female differential does not change. Schooling increases dramatically for both males and females between 1988 and 1995, and the females now have 0.65 year's less schooling than males. For the occupation distribution, the table shows that relative to 1988, more workers are now in the professional work, head and office worker categories while less are working as unskilled laborers. More females are now private owners and institution heads; however, a higher proportion of unskilled manual laborers are females in 1995 than in 1988.

(ii) The Quantile Regression Model

To analyze earning differentials, we use the standard model which is based on the human capital earnings function developed by Mincer (1974):

$$\ln(Y_i) = j(X_i) + m_i , \qquad (1)$$

where $\ln(Y_i)$ is the natural logarithm of the earnings for observation i, and X_i is a vector of characteristics including a measure of school years, experience, gender, occupation dummies, economic sector dummies, region dummies, etc. For 1988 dataset, we include the dummy variable to identify whether there is a newly born baby in the household, and for 1995 dataset, we use a marital status dummy instead of baby birth. We also proxy experience by its potential term: age minus years of education minus six. Further, there are some differences in the dummy variables for occupation, economic sector and province for the two years because of the adjustment in the survey.

The traditional method to estimate the above Mincerian equation is the OLS, which characterizes the wage distribution only at the mean of the distribution; therefore it cannot investigate the effects of independent variables on the "shape" of the distribution. For example, the rate of return to schooling might not be identical at different earning levels. As we have mentioned in the introduction, quantile regression, introduced by Koenker and Bassett (1978), provides a mechanism for estimating models for the conditional median function, and the full range of other conditional quantile functions.

We say that the earning of a worker at the τth quantile if his earning is higher than the proportion τ of the reference group of workers and lower than the proportion $(1-\tau)$. Hence, half of the workers have higher earnings than the median worker and half have lower earnings. Likewise, the quartiles divide the population into four groups with equal proportions of the population in each group; the quintiles divide the population into five segments and the deciles into ten parts. The above cases can be generalized to the quantiles, or the percentiles.

Classical linear regression is a method of estimating conditional mean functions by minimizing sums of squared residuals, which is an extension of the idea of estimating an unconditional mean parameter. Similarly, quantile regression could be viewed as a way of extending the univariate quantile estimation to estimation of conditional quantile functions via an optimization of a piecewise linear objective function of residuals. The median regression method, also known as least absolute deviation (LAD) estimator, fits the regression line that minimizes the sum of absolute residuals instead of the sum of squared residuals. Since the solution to the problem of minimizing a sum of absolute residuals could yield the median, minimizing a sum of asymmetrically weighted absolute residuals would yield the quantiles. By solving

$$\min_{b \in Rp} \sum \rho_\tau (y_i - \xi(x_i, \beta)), \tag{2}$$

where the function $\rho_\tau(\bullet)$ is the absolute value function that yields the τth sample quantile, and $\xi(x_i, \beta)$ is the linear function of parameters, we can obtain estimates of the general conditional quantile functions.

By estimating an entire range of conditional quantile functions, quantile regression is capable of offering a more complete statistical analysis of the stochastic relationships, especially estimating the upper and lower quantile curves of interest as a function of a set of covariates without imposing parametric assumptions on the relationships among those curves.

(iii) Decomposition

One of the most common methodologies for determining earnings differentials by gender was developed by Oaxaca (1973) and Blinder (1973). It helps distinguish the unequal treatment of females outside the labor market (differentials in variables) from the unequal treatment inside the labor market (differentials in coefficients). The average unadjusted logarithmic differential in gender earnings may be decomposed into an "explained" portion and an "unexplained" portion, which represents the same characteristics being rewarded differently.

$$\overline{\ln Y_m} - \overline{\ln Y_f} = (\overline{X_m} - \overline{X_f})\hat{b}_m + (\hat{b}_m - \hat{b}_f)\overline{X_f} \tag{3}$$

or

$$\overline{\ln Y_m} - \overline{\ln Y_f} = (\overline{X_m} - \overline{X_f})\hat{b}_f + (\hat{b}_m - \hat{b}_f)\overline{X_m}\ , \tag{4}$$

where m stands for the male worker and f stands for the female worker. X is a vector of the characteristics of the workers, and $\hat{b}$ is a vector of the estimated coefficients. The first term on the right-hand side stands for the earning differential explained by the characteristic differential, while the second term is usually interpreted as discrimination because it represents the different returns for the same characteristics. The decomposition is performed based on the assumption that the "discriminated" group should be paid the same as another group, which means that males and females with the same characteristics such as education, experience, occupation, etc., should earn the same. Equation (3) assumes that the returns to characteristics for males should be the "true" coefficient given certain characteristics, while equation (4) assumes returns to characteristics for females are the true betas.

The standard decomposition above, however, pays little attention to the underlying earnings distribution.[1] Such decomposition may show that the average males may get earnings premium, but such average earning gaps and decompositions might not be

[1] Gustafsson and Li used Jenkins's (1994) method for analyzing discrimination that differently weights different portions of the earnings distribution but does not directly measure discrimination at various points in the distribution.

representative of the gaps at different quantiles of the earnings distributions for the reference populations. In fact, the size of the earnings gap and the weights of factors that make it up might not be constant along the whole of earnings scale.

The basic quantile regression model specifies the conditional quantile as a linear function of covariates. For the conditional earnings distribution, the formal econometric representation is given by

$$\ln Y_i = X_i{}'\boldsymbol{b}_t + u_{ti} \tag{5}$$

$$Q_t(\ln Y_i | X_i) = \ln Y_i^t = X_i{}'\boldsymbol{b}_t \, , \tag{6}$$

assuming that the τth quantile of the error term, conditional on X_i, is zero. Under this representation, the measure of discrimination for different quantiles can be given as

$$\hat{Q}_t(\ln Y_m | X_i) - \hat{Q}_t(\ln Y_f | X_i) = X_i{}'(\hat{\boldsymbol{b}}_{tm} - \hat{\boldsymbol{b}}_{tf}) \, , \tag{7}$$

where we compare the quantiles of the two earnings distributions of males and females conditional on the same set of characteristics as an approximation to the unobservable measure, as adopted by Garcia, Hernandez and Lopez-Nicolas (2001). They argue that the choice of X_i is arbitrary. The basic assumption for the standard Oaxaca decomposition is that the male beta is true, which is very restrictive. The male coefficient might increase and female coefficient might decrease in absence of the discrimination. The true beta might be lying somewhere between the male beta and female beta. Therefore we follow them to choose the sample average vector of characteristics to carry out the decomposition. We also experimented decompositions by using average characteristics around a symmetric neighborhood of every quantile, which are presented in Appendix Table 3 (a-e) and Table 4 (a-e) for 1988 and 1995 respectively. However, the patterns do not show much difference.

III. Empirical Results

(i) Quantile Regression

We estimate the modified Mincerian equation for different values of τ, which are 10%, 25%, 50%, 75% and 90% for both years, and for both males and females respectively. We also estimate the traditional OLS regression to make comparison. We are particularly interested in the influence of experience, the square of experience, childbirth or marital status and year of schooling, therefore we include only the above coefficient estimates are in Table 2-a and Table 2-b for 1988 and 1995 dataset separately. The complete estimation is presented in Appendix Table 1-a and b. Figure 1-3 illustrates the returns to those characteristics for males and females.

Table 2-a presents the earnings returns to experience of males and females at the different quantiles. We also find that both males and females in the 1988 sample have a diminishing rate of return to experience as the earnings increase. Figure 1 shows that the pattern of return to experience by quantiles evaluated at a level of 10 years of experience is very similar for both groups in 1988, which demonstrates that there is not much of a differential between genders. It is clear that the schedule is quite flat across various earnings levels, indicating that little variation in returns to experience along the earnings distribution. While according the OLS estimates, on average, males have only a slightly

higher return to 10 years' experience (about 3%) than females regardless of their wage levels. We also present the estimates of quantile regression for return to experience for 20 years' experience in Appendix Figure 1, which tells a similar story. It seems that OLS method is appropriate enough to depict the picture of earnings returns to experience in 1988 for Chinese men and women workers.

The quantile regression estimates for return to experience as well as OLS results for 1995 dataset are given in Table 2-b. Figure 1 illustrates that the return to 10 years of experience is consistent higher for females than males along the earnings distribution, while the difference between the returns keeps getting narrower when we move from the bottom quantile to the top quantile. The gain from experience for females declines sharply from about 84% at the first quantile to only 22% at the ninth quantile. The males show a similar decreasing pattern but only at a very slow speed, with 32.8% at the top and 13.8% at the bottom. Although the OLS shows a 23.6% differential in the rate of return to experience for males and females, it cannot clearly shows the greatest gap exists in the bottom earnings group and the smallest in the top earnings group. In 1995, Chinese women workers who earn least could gain more from more experience relative to those have higher income. Further, compared to higher income workers, the females at the bottom of the earnings distribution receive more benefit from having more experience than males. The picture of 20 years experience is presented in the Appendix Figure 1, which displays similar pattern.

In 1988, males have a positive earnings effect from the baby birth variable at all points along the distribution, while females generally experience negative earning effects, which are only significant at the median and 0.75 quantile. There is no clear pattern for the effect of a newly born baby at different quantiles, however, as shown in the Figure 2. Males with the lowest earnings level have the greatest premium about 6% from having a baby, followed by males at 75% quantile and median. The males located at the top of the earnings distribution gain least from a baby birth, only 2.41%. Although our quantile regression demonstrates that effect of a newly born baby is not identical at different quantile levels of earnings for males in the 1988's China, the variation is apparently very small. Hence the OLS estimate, around 5.31%, might be suitable enough to explain the pattern for earnings gain for men from childbirth. For females, on the contrary, generally they have a negative effect from giving birth to children, but only significant at the median and 75% quantile. The picture shows that the line pattern is almost horizontal, barely with any variation. Further, the OLS gives a significant average effect of childbirth for females approximately negative 3%. Although in many countries an important reason for why women earn less than their male counterparts is that they have to pause their working for baby birth, those interruptions are relatively short in urban China. Hence it is not surprising that the effect of childbirth for women would not be great.

For 1995 we substitute the analysis of baby birth with an analysis of the marriage effect. It can be shown that marriage provides males with a substantial increase in their earnings, as illustrated in Figure 2. The bottom quantile male workers enjoy the greatest marriage premium around 46% while the effect shrinks when moving up to the higher quantiles to around 14% at the top quantile. For females, on the contrary, except for those who have higher earnings at the 0.75 and 0.90 quantiles, they usually do not earn more after they marry. Marriage has no significant effect on female workers with low earnings but provides some mild positive effect on those with high income. We also see that the gender differential in the marriage premium is the greatest for those who earn least, and then declines to about 7 percent for those earn most. In contrast to quantile regression, OLS gives significant positive marriage effect for both males and females around 26% and 6% respectively, which does not capture the wide variations at different quantiles of the earnings.

We also examine the return to education. Table 2-a gives the coefficients estimated for 1988. The quantile regression shows that at the quantiles below the median earnings level, the females have a slightly higher return to schooling than males, while beyond the median, males' turn to education is slightly higher than females. At the 50% quantile, the returns to education are also equal for both genders. However, we can see that there is very little differential earnings gain from education between females and males at every quantile by looking at the Figure 3. It is interesting to see that the OLS estimates for males and females are also the same at 1.98%. This is quite different from the empirical results of most of the other countries, where females have higher return to education than males. Furthermore, Figure 3 also indicates that in 1988 either females or males have a quite flat pattern of returns to education across earnings distribution, which makes OLS a more efficient method to estimate the education effects.

As to 1995, we find that returns to schooling for females are higher than for males at all the quantiles, while they decrease as earnings increase. Those at top quantile have the greatest gender gap in returns to education around1.67%, followed by those at 75% quantile. Workers located at the 0.25 quantile make smallest differential regarding to the earnings return to schooling. It is clear from Figure 3 that the difference in returns to schooling between genders is greater at the two tales and smaller in the middle, and hence for those females with lowest earnings and highest earnings, they would earn more from their schooling compared to the males whose earnings is in the same location of the distribution.

For the 1988 dataset we find that there are not much great variations for our coefficient estimates along the earning distributions for males and females, though the effects of independent variables on the earnings are not identical at different quantiles. This means that the fitted lines at different quantiles might be almost parallel with each other as well as OLS line. Therefore we might consider that the OLS is appropriate enough for estimating the earnings equation for the 1988 data. Further, the effects of occupation, economic sector and region in our regression are fairly small and lots of are insignificant, as presented in the Appendix Table 1. This may be because in 1988, China was still in centrally planned economy, in which the labor was bureaucratically allocated and wages were administratively regulated. Such an egalitarian wage system eliminated or minimized earnings differences across occupations, economic sectors, regions and genders. However, the quantile regression method is useful for the analysis of 1995.

In 1995, China has been on the track of transition from the redistributive economy to a more market-oriented economy. We can see that there are much richer variations in the coefficients for the independent variables of interest along the earnings distribution than 1988. It makes more sense to estimate the Mincerian equation at different quantiles instead of OLS just at the conditional mean of earnings, since the former technique is more capable of capturing a plenty of the variations in the effects of regressors.

(ii) Earnings Gap Decompositions

In this section we use an Oaxaca-type decomposition to analyze the unexplained part of the total male-female earnings gap at the mean and the selected quantiles. Tables 3-a and 3-b present the log earnings for males and females at the mean and the exact 10th, 25th, 50th, 75th, and 90th quantiles of each sample. The tables also include the unadjusted gender earnings gap. Using the parameter estimates derived from the quantile regression and OLS (see Appendix Table 1) we decompose the earnings gap at the mean of the vector of independent variables. The detailed decomposition at each quantile is presented in Appendix Tables 2 and 3.

Table 3-a shows the decomposition result for the 1988 dataset. Column 4 presents the unadjusted earnings gap between males and females, which is the highest at the lowest quantile (20.82 percent) and decreases until the third quartile (15.97 percent). At the ninth quantile the unadjusted earnings gap rises slightly to 17.57 percent. Column 5 provides the "unexplained part" of the earnings gap, $X'(\beta_m-\beta_f)$, which shows a pattern similar to the unadjusted earnings gap. This implies that the earnings differential

unexplained by the observable characteristics (the X's) between genders is the largest at the lowest earnings level, then gradually decreases as earnings increase, and finally increases slightly at the 90th percentile. Column 6 provides the ratio of unexplained part of earnings differential to the total earnings gap. This ratio reaches it peak of nearly 90 percent at the bottom quantile. The lowest ratio occurs at the top quantile (55 percent), indicating that there is only slightly more that half of the total earnings differential can be explained by the different treatment to the characteristics (X's) for males and females.

The 1995 earnings decompositions are presented in Table 3-b. Column 4 presents the 1995 unadjusted earnings gap. Comparing the earnings gaps over time we find that the average earnings gap has increased slightly over time from 18.5 percent in 1988 to 22 percent in 1995. However, it is also apparent that variation in the unadjusted earnings gap across quantiles also increased. For 1988 the bottom to top deciles unadjusted earnings gap varied from 21 percent to 18 percent; for 1995, the bottom to top deciles unadjusted earnings gap varied from 33 percent to 15 percent. Column 6 provides the ratio of unexplained part of earnings differential to the total earnings gap. Here we find the opposite results to those for the unadjusted earnings gap – in 1995 there is very little variation in the ratio of unexplained part of earnings differential to the total earnings gap across quantiles (0.67 at the bottom decile and 0.74 at the top decile).

As noted above, since the ratio $X'(\beta_m-\beta_f)$/earnings gap captures the earnings gap resulting from differences in the returns to characteristics based on the same vector the characteristics for both males and females, it is often viewed as evidence of discrimination.[1] Our results show substantial differences in the degree of discrimination at the tails of the distribution in 1988; 90 percent of the lowest decile's earnings gap is "unexplained", while only 55 percent of the top decile earnings gap is due to discrimination.[2] In contrast, we find much less variation in discrimination across quantiles in 1995.

This finding of little difference in discrimination across quantiles in 1995 deserves some comment. Recall the regression results as illustrated in Figures 1-3 showed relatively little quantile variation in the returns to schooling and experience in 1988 and substantial variation by quantile in 1995. From this one might conclude that quantile regression is unnecessary in 1988 and extremely useful in 1995. However, we find the opposite results in Table 3 when we examine the adjusted earnings gap ratios-there is little variation in the adjusted earnings gap ratio by quantile in 1995 and substantial quantile variation in l988. Hence it is difficult to predict from partial regression results the efficacy of quantile regression. Finally, if we examine column 6 in both Table 3-a and Table 3-b together we find that low earnings women suffer greater discrimination under the socialist conditions of 1988 (0.90 adjusted earnings gap ratio) than under the emerging market conditions of 1995 (0.67 adjusted earnings gap ratio). In contrast the higher earnings women suffer greater discrimination in 1995 (0.74) than in 1988 (0.55). All of this is hidden in the OLS results, which show a slight increase in the adjusted earnings gap ratio, from 0.66 in 1988 to 0.71 in 1995.

---

[1] An alternative explanation is that on average men are more productive than women.

[2] These findings are not consistent with those of Garcia, Hernandez and Lopez-Nicolas (2001) and Montenegro (2001), who report that the unexplained part is the highest at the upper quantile of the wage earnings distribution.

IV. Conclusion

The principal goal of this paper is to analyze gender differentials in the earnings along the conditional earnings distribution in China for the year 1988 and 1995. We use quantile regression as well as OLS to estimate the adjusted Mincer equation. We also decompose the unadjusted earnings gap on every quantile based on Oaxaca method.

Our quantile regression results show that the returns to experience, childbirth, marriage and education do not vary much at different quantiles for 1988, but show substantial variation in 1995.

In contrast, the decomposition results show that in 1995 there is little difference in discrimination across quantiles; however, in 1988, we find significant variation in discrimination by quantile. These results suggest that it is difficult to predict from partial regression results the efficacy of quantile regression.

Across time we find that low earnings females suffer greater discrimination under the socialist conditions of 1988 than under the emerging market conditions of 1995. On the contrary, the higher earnings females suffer greater discrimination in 1995 than in 1988. All of this is incorporated in the OLS results, which exhibit a slight increase in the adjusted earnings gap ratio.

References:

Eide, R. Eric, Mark H. Shhowalter and David P. Sims. 2002. "The Effects of Secondary School Quality on the Distribution of Earnings." *Contemporary Economic Policy*. Volume 20, Number 2. pp. 160-170.

Garcia, Jaume, Pedro J. Hernandez and Angel Lopez-Nicolas. 2001. "How Wide is the Gap? An Investigation of Gender Wage Differences Using Quantile Regression." *Empirical Economics.* 26:149-167.

Gustafsson, Bjorn and Li Shi. 2001. "Economic Transformation and the Gender Earnings Gap in Urban China." *China's Retreat from Equality*. pp. 184-209.

Koenker, Roger. 2000. "Quantile Regression." *International Encyclopedia of the Social Sciences.*

Koenker, Roger and Kevin F. Hallock. 2001. "Quantile Regression." *Journal of Economic Perspectives.* Volume 15, Number 4. pp. 143-156.

Mueller, E. Richard. 1998. *"Public-private Sector Wage Differentials in Canada: Evidence from Quantile Regressions."* Economic Letters. 60. pp. 229-235.

Montenegro, Claudio. 2001. *"Wage Distribution in Chile: Does Gender Matter? A Quantile Regression Approach."* Policy research report on Gender and Development.
The World Bank Working Paper Series No.20.

Table 1-a **Descriptive Statistics of Main Variables by Gender - 1988**

Variable	Male (n=9114)	Female (n=8444)	M/F Ratio
earnings	1971.5310	1656.3790	1.1903
exp	22.8126	20.8019	1.0967
exp2	663.2137	553.6380	1.1979
babybirth	0.1071	0.1099	0.9744
schoolyr	9.5493	8.7226	1.0948
Occupation			
prvt_owner	0.0066	0.0039	1.6845
prvt_owner_mngr	0.0011	0.0009	1.1581
prof_tech	0.1565	0.1592	0.9830
gvmt_offi	0.0745	0.0152	4.9147
fctry_mngr	0.0297	0.0045	6.6074
office_wkr	0.2580	0.2110	1.2223
manual_wkr	0.4700	0.5956	0.7892
Economic Sector			
agri	0.0111	0.0077	1.4396
mining	0.0406	0.0219	1.8530
manufacture	0.4115	0.4467	0.9211
geological	0.0095	0.0082	1.1682
construction	0.0409	0.0276	1.4832
trans_post	0.0831	0.0499	1.6659
trade_cater	0.1109	0.1733	0.6402
estate_utility	0.0148	0.0137	1.0782
personal_consult	0.0066	0.0126	0.5244
health_sport	0.0335	0.0587	0.5697
edu_cult_art	0.0669	0.0783	0.8550
sci_tech	0.0251	0.0155	1.6196
finan_insur	0.0155	0.0152	1.0206
party_gvmt_soci	0.1163	0.0505	2.3053
other	0.0066	0.0071	0.9265
dontknow	0.0011	0.0024	0.4632
Provinces			
Bejing	0.0508	0.0456	1.1142
Shanxi	0.1124	0.1037	1.0830
Liaoning	0.1019	0.1072	0.9511
Jiangsu	0.1284	0.1260	1.0188
Anhui	0.0978	0.0972	1.0055
Henan	0.1155	0.1194	0.9678
Hubei	0.1083	0.1105	0.9801
Guangdong	0.1172	0.1193	0.9826
Yunnan	0.0984	0.1054	0.9338
Gansu	0.0678	0.0642	1.0564

Source: Chinese Household Income Project, 1988

Table 1-b **Descriptive Statistics of Main Variables by Gender - 1995**

Variable	Male (n=6222)	Female (n=5705)	M/F Ratio
earnings	6734.6940	5592.5520	1.2042
exp	21.5005	19.7534	1.0884
exp2	572.9296	482.1800	1.1882
married	0.8709	0.8792	0.9906
schoolyr	12.1395	11.4855	1.0569
Occupation			
owner	0.0063	0.0053	1.1920
manager	0.0019	0.0014	1.3753
profwork	0.2136	0.2282	0.9359
head	0.1659	0.0564	2.9387
office	0.1893	0.2193	0.8634
skilled	0.2493	0.1774	1.4053
unskil	0.1189	0.2168	0.5485
other	0.0283	0.0649	0.4362
Economic Sector			
manf	0.4176	0.3975	1.0503
mining	0.0116	0.0093	1.2456
const	0.0317	0.0254	1.2457
tspcomm	0.0561	0.0400	1.4035
trade	0.1149	0.1593	0.7212
real	0.0307	0.0412	0.7452
health	0.0347	0.0566	0.6132
educ	0.0632	0.0834	0.7570
tech	0.0257	0.0203	1.2647
finance	0.0180	0.0214	0.8417
govt	0.1337	0.0910	1.4699
other	0.0264	0.0179	1.4742
Provinces			
Beijing	0.0736	0.0699	1.0525
Liaoning	0.1078	0.1043	1.0340
Henan	0.0829	0.0780	1.0632
Jiangsu	0.1128	0.1120	1.0073
Anhui	0.0678	0.0715	0.9484
Hubei	0.1061	0.1059	1.0019
Guangdong	0.0826	0.0820	1.0070
Shanxi	0.0972	0.0918	1.0586
Gansu	0.0561	0.0554	1.0127
Sichuan	0.1194	0.1283	0.9307
Yunnan	0.0935	0.1008	0.9281

Source: Chinese Household Income Project, 1995

Table 2-a **Quantile Regression Estimates for 1988 Dataset**

Variables	τ=0.10		τ=0.25		τ=0.50		τ=0.75		τ=0.90		OLS	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Constant	6.1263	6.0016	6.4446	6.4137	6.7191	6.7045	6.9199	6.9361	7.1344	7.1353	6.5907	6.5815
	(0.0273)	(0.0454)	(0.0219)	(0.0296)	(0.0171)	(0.0227)	(0.0176)	(0.0186)	(0.0263)	(0.0288)	(0.0198)	(0.0253)
exp	0.0536	0.0598	0.0437	0.0448	0.0369	0.0344	0.0336	0.0290	0.0308	0.0240	0.0444	0.0424
	(0.0015)	(0.0026)	(0.0012)	(0.0017)	(0.0010)	(0.0013)	(0.0011)	(0.0011)	(0.0017)	(0.0017)	(0.0012)	(0.0015)
exp2	-0.0007	-0.0010	-0.0005	-0.0007	-0.0004	-0.0004	-0.0004	-0.0003	-0.0003	-0.0003	-0.0005	-0.0006
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
babybirth	0.0604	-0.0201	0.0275	-0.0187	0.0477	-0.0262	0.0590	-0.0244	0.0241	-0.0107	0.0531	-0.0293
	(0.0156)	(0.0253)#	(0.0130)	(0.0169)#	(0.0105)	(0.0134)	(0.0111)	(0.0112)	(0.0169)#	(0.0169)#	(0.0121)	(0.0149)
schoolyr	0.0218	0.0261	0.0192	0.0194	0.0172	0.0172	0.0168	0.0162	0.0147	0.0146	0.0198	0.0198
	(0.0014)	(0.0023)	(0.0011)	(0.0015)	(0.0008)	(0.0012)	(0.0009)	(0.0010)	(0.0013)	(0.0015)	(0.0010)	(0.0013)

Notes: a. Standard errors in parenthesis.
b. # stands for not significant at 95% level.
c. Omitted variables: Manual_worker, Manufacture, Jiangsu
d. The complete regression results are presented in Appendix Table 1-a.
Source: Chinese Household Income Project, 1988

Table 2-b **Quantile Regression Estimates for 1995 Dataset**

Variables	τ=0.10		τ=0.25		τ=0.50		τ=0.75		τ=0.90		OLS	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Constant	6.5766	6.0557	7.2117	6.9749	7.7215	7.4387	8.1006	7.8509	8.4735	8.1531	7.5497	7.2543
	(0.0703)	(0.1167)	(0.0553)	(0.0699)	(0.0395)	(0.0511)	(0.0498)	(0.0508)	(0.0652)	(0.0725)	(0.0450)	(0.0589)
exp	0.0388	0.1071	0.0292	0.0652	0.0254	0.0457	0.0180	0.0294	0.0148	0.0249	0.0270	0.0596
	(0.0043)	(0.0070)	(0.0034)	(0.0042)	(0.0025)	(0.0031)	(0.0031)	(0.0031)	(0.0041)	(0.0045)	(0.0028)	(0.0036)
exp2	-0.0006	-0.0023	-0.0003	-0.0013	-0.0003	-0.0008	-0.0001	-0.0004	-0.0001	-0.0003	-0.0003	-0.0012
	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)#	(0.0001)	(0.0001)	(0.0001)
married	0.4595	0.0347	0.3220	0.0479	0.1517	0.0498	0.1909	0.0808	0.1364	0.0657	0.2569	0.0603
	(0.0392)	(0.0594)#	(0.0305)	(0.0352)#	(0.0223)	(0.0265)#	(0.0285)	(0.0267)	(0.0381)	(0.0400)	(0.0254)	(0.0306)

schoolyr	0.0443	0.0591	0.0412	0.0491	0.0372	0.0458	0.0300	0.0453	0.0278	0.0445	0.0376	0.0470
	(0.0042)	(0.0078)	(0.0034)	(0.0045)	(0.0024)	(0.0034)	(0.0031)	(0.0034)	(0.0041)	(0.0049)	(0.0027)	(0.0039)

Notes: a. Standard errors in parenthesis.
b. # stands for not significant at 95% level.
c. Omitted variables: unskilled, Manufacture, Jiangsu
d. The complete regression results are presented in Appendix Table 1-b.
Source: Chinese Household Income Project, 1995

Table 3-a **Unadjusted Earnings Gaps and Decomposition for 1988 Data Set**

Quantiles	log_earnings_Male	log_earnings_Female	Earnings Gap	$X'(\beta_m - \beta_f)$	$X'(\beta_m - \beta_f)$/Earnings Gap
(1)	(2)	(3)	(4)	(5)	(6)
τ=0.10	6.9836	6.7754	0.2082	0.1866	0.8961
τ=0.25	7.2881	7.0961	0.1920	0.1202	0.6257
τ=0.50	7.5247	7.3563	0.1684	0.0974	0.5780
τ=0.75	7.7383	7.5787	0.1597	0.0966	0.6050
τ=0.90	7.9593	7.7836	0.1757	0.0970	0.5520
Mean	7.4923	7.3068	0.1854	0.1221	0.6582

Note: log_earnings_Male and log_earnings_Female are calculated from the observation at the exact quantile.

Source: Chinese Household Income Project, 1988

Table 3-b **Unadjusted Earnings Gaps and Decomposition for 1995 Data Set**

Quantiles	log_earnings_Male	log_earnings_Female	Earnings Gap	$X'(\beta_m - \beta_f)$	$X'(\beta_m - \beta_f)$/Earnings Gap
(1)	(2)	(3)	(4)	(5)	(6)
τ=0.10	8.0226	7.6898	0.3327	0.2241	0.6735
τ=0.25	8.3802	8.1513	0.2289	0.1508	0.6587
τ=0.50	8.6995	8.5206	0.1789	0.1190	0.6650
τ=0.75	9.0042	8.8537	0.1505	0.1076	0.7147
τ=0.90	9.2951	9.1484	0.1468	0.1083	0.7381
Mean	8.6664	8.4466	0.2199	0.1564	0.7113

Note: log_earnings_Male and log_earnings_Female are calculated from the observation at the exact quantile.
Source: Chinese Household Income Project, 1995

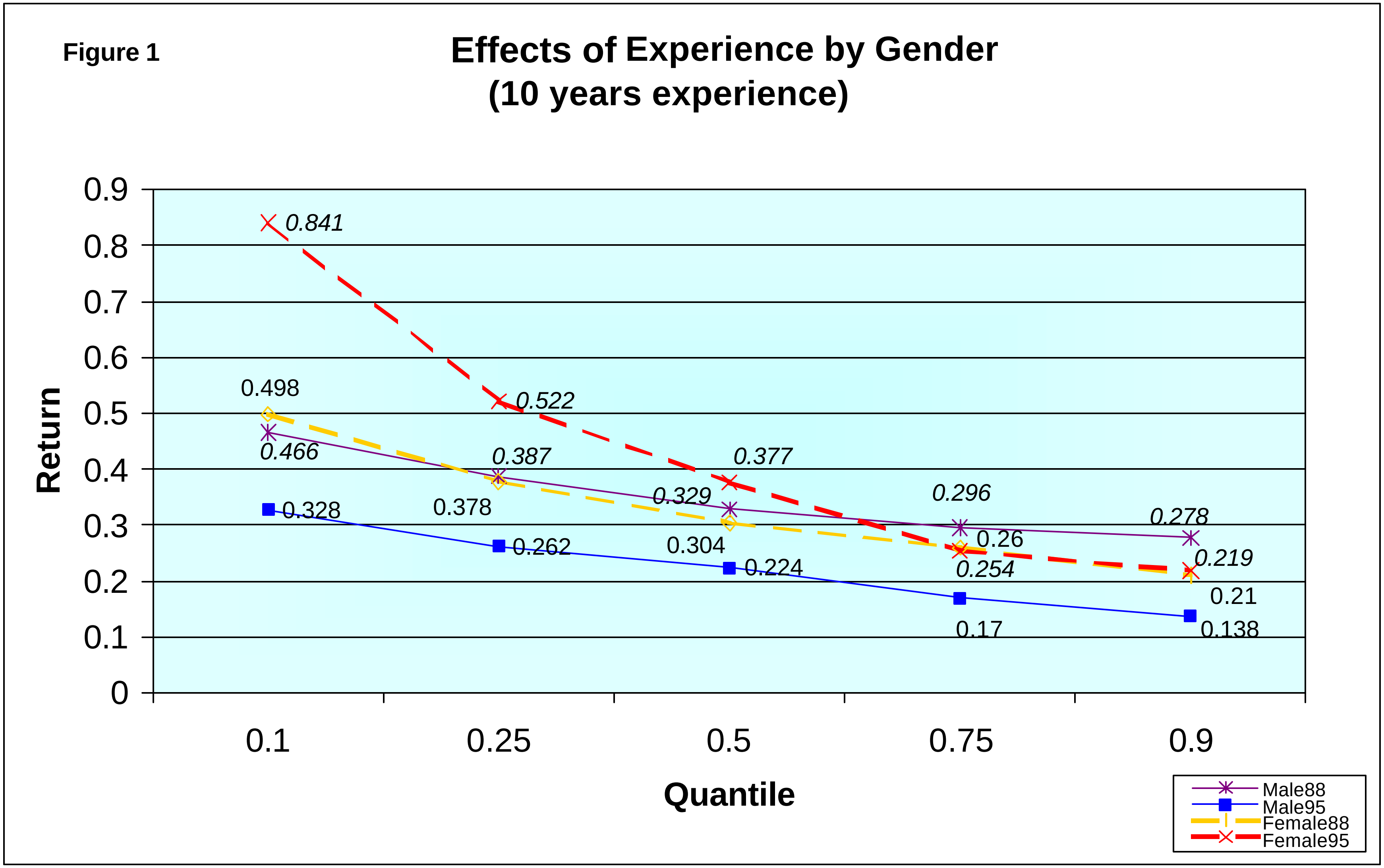


Figure 1
Effects of Experience by Gender
(10 years experience)
Return
Quantile
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.1
0.25
0.5
0.75
0.9
0.841
0.522
0.377
0.254
0.219
0.498
0.378
0.304
0.26
0.21
0.466
0.387
0.329
0.296
0.278
0.328
0.262
0.224
0.17
0.138
Male88
Male95
Female88
Female95

Figure 2

Effects of Family Structure by Gender
Baby Birth (1988), Marriage (1995)

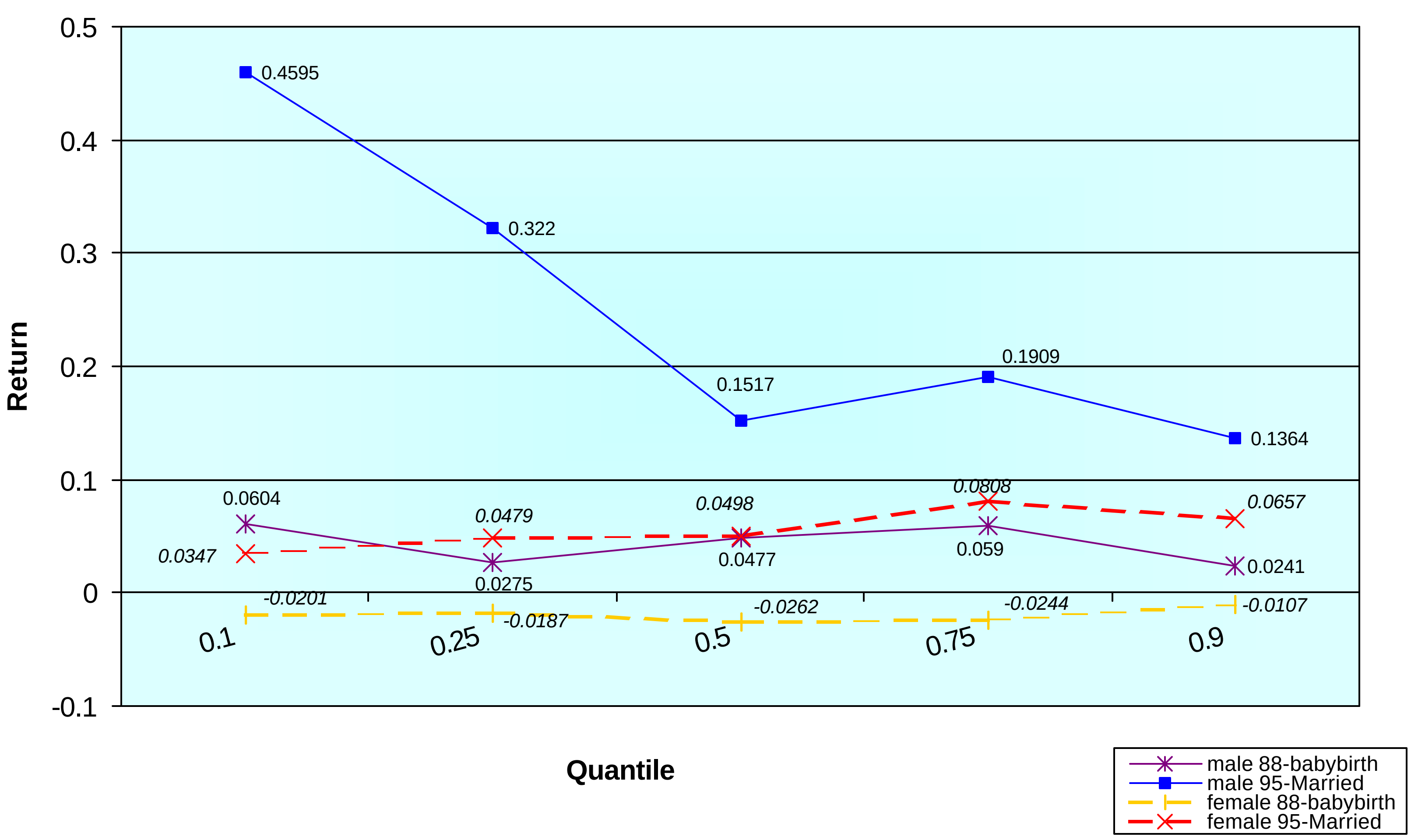
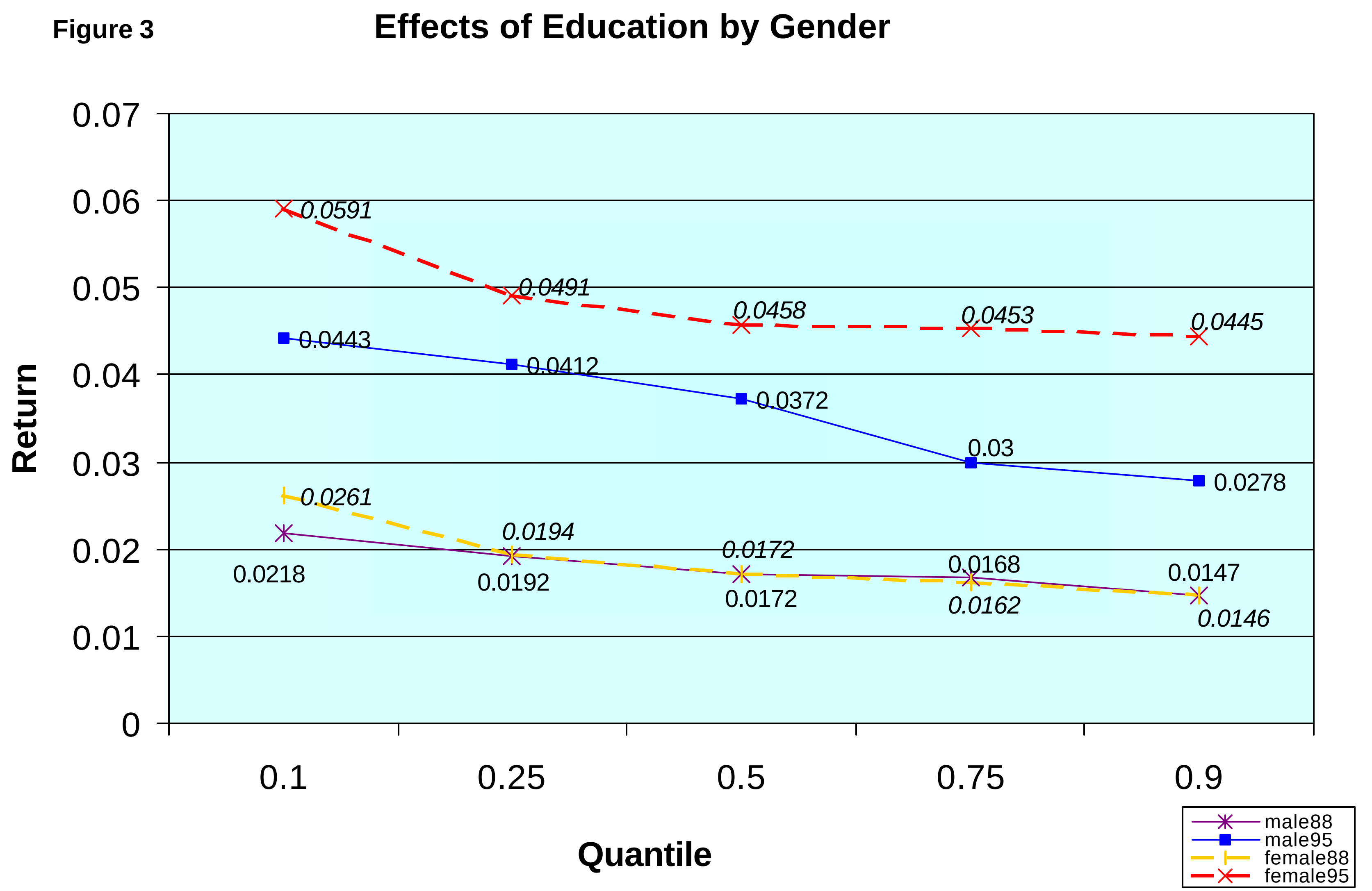


Figure 3



Effects of Education by Gender
Return
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0.0591
0.0491
0.0458
0.0453
0.0445
0.0443
0.0412
0.0372
0.03
0.0278
0.0261
0.0194
0.0172
0.0168
0.0147
0.0218
0.0192
0.0172
0.0162
0.0146
0.1
0.25
0.5
0.75
0.9
Quantile
male88
male95
female88
female95

Appendix Table 1-a

Quantile Regression Estimates for 1988 Dataset

Variables	τ=0.10		τ=0.25		τ=0.50		τ=0.75		τ=0.90		OLS	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
exp	0.0536	0.0598	0.0437	0.0448	0.0369	0.0344	0.0336	0.0290	0.0308	0.0240	0.0444	0.0424
	(0.0015)	(0.0026)	(0.0012)	(0.0017)	(0.0010)	(0.0013)	(0.0011)	(0.0011)	(0.0017)	(0.0017)	(0.0012)	(0.0015)
exp2	-0.0007	-0.0010	-0.0005	-0.0007	-0.0004	-0.0004	-0.0004	-0.0003	-0.0003	-0.0003	-0.0005	-0.0006
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
babybirth	0.0604	-0.0201	0.0275	-0.0187	0.0477	-0.0262	0.0590	-0.0244	0.0241	-0.0107	0.0531	-0.0293
	(0.0156)	(0.0253)#	(0.0130)	(0.0169)#	(0.0105)	(0.0134)	(0.0111)	(0.0112)	(0.0169)#	(0.0169)#	(0.0121)	(0.0149)
schoolyr	0.0218	0.0261	0.0192	0.0194	0.0172	0.0172	0.0168	0.0162	0.0147	0.0146	0.0198	0.0198
	(0.0014)	(0.0023)	(0.0011)	(0.0015)	(0.0008)	(0.0012)	(0.0009)	(0.0010)	(0.0013)	(0.0015)	(0.0010)	(0.0013)
Occupation												
prvt_owner	0.0097	0.2738	0.0238	0.1857	0.0418	0.0807	0.0811	0.0633	0.0948	-0.0604	-0.0046	0.0349
	(0.0586)#	(0.1182)	(0.0487)#	(0.0810)#	(0.0392)#	(0.0643)#	(0.0420)#	(0.0541)#	(0.0649)#	(0.0803)#	(0.0457)#	(0.0724)#
prvt_owner	-0.0570	0.3622	-0.0482	0.4188	-0.0238	0.1734	0.2076	0.1944	0.0731	0.1254	0.0162	0.1936
_mngr	(0.1418)#	(0.0820)	(0.1078)#	(0.1609)	(0.0908)#	(0.1239)#	(0.0930)	(0.1094)#	(0.1557)#	(0.0559)	(0.1107)#	(0.1470)#
prof_tech	0.1570	0.2579	0.1277	0.1914	0.1137	0.1585	0.0986	0.1297	0.0958	0.1322	0.1263	0.2016
	(0.0156)	(0.0270)	(0.0133)	(0.0182)	(0.0108)	(0.0143)	(0.0113)	(0.0117)	(0.0167)	(0.0179)	(0.0125)	(0.0159)
gvmt_offi	0.1641	0.3374	0.1305	0.2458	0.1072	0.2337	0.0817	0.1526	0.0702	0.0901	0.1197	0.2319
	(0.0208)	(0.0632)	(0.0176)	(0.0428)	(0.0144)	(0.0346)	(0.0153)	(0.0293)	(0.0234)	(0.0444)	(0.0167)	(0.0387)
fctry_mngr	0.1594	0.1705	0.1574	0.1578	0.1623	0.1826	0.1944	0.1081	0.2071	0.2312	0.2021	0.1961
	(0.0277)	(0.1030)#	(0.0237)	(0.0748)#	(0.0194)	(0.0602)	(0.0205)	(0.0501)	(0.0313)	(0.0706)	(0.0225)	(0.0676)
office_wkr	0.1256	0.1862	0.0918	0.1297	0.0549	0.1039	0.0370	0.0841	0.0209	0.0628	0.0777	0.1336
	(0.0130)	(0.0218)	(0.0107)	(0.0146)	(0.0087)	(0.0115)	(0.0093)	(0.0097)	(0.0144)#	(0.0146)	(0.0101)	(0.0128)
Economic Sector												
agri	-0.0383	0.1325	-0.0137	0.0340	0.0017	-0.0214	-0.0108	-0.0432	-0.0397	-0.0449	-0.0035	0.0084
	(0.0453)#	(0.0848)#	(0.0376)#	(0.0586)#	(0.0305)#	(0.0465)#	(0.0323)#	(0.0390)#	(0.0480)#	(0.0580)#	(0.0354)#	(0.0519)#
mining	0.0805	-0.0221	0.1159	0.0232	0.1509	-0.0310	0.1617	-0.0422	0.2120	-0.0130	0.1406	0.0186
	(0.0256)	(0.0526)#	(0.0212)	(0.0356)#	(0.0170)	(0.0284)#	(0.0177)	(0.0240)#	(0.0272)	(0.0363)#	(0.0197)	(0.0316)#
geological	0.0883	0.0350	0.0548	-0.0923	0.0249	-0.0685	-0.0141	-0.0573	-0.0184	-0.0718	0.0320	-0.0195

	(0.0474)#	(0.0805)#	(0.0403)#	(0.0574)#	(0.0330)#	(0.0455)#	(0.0348)#	(0.0389)#	(0.0525)#	(0.0572)#	(0.0383)#	(0.0512)#
construction	-0.0503	0.0621	0.0036	0.0318	0.0260	0.0180	0.0249	0.0222	0.0342	0.0292	-0.0068	0.0494
	(0.0243)	(0.0457)#	(0.0201)#	(0.0318)#	(0.0164)#	(0.0252)#	(0.0174)#	(0.0211)#	(0.0268)#	(0.0323)#	(0.0190)#	(0.0280)#
trans_post	0.0308	0.0645	0.0272	0.0479	0.0194	0.0263	0.0314	0.0211	0.0483	0.0857	0.0374	0.0584
	(0.0180)#	(0.0361)#	(0.0149)#	(0.0243)	(0.0121)#	(0.0193)#	(0.0129)	(0.0163)#	(0.0198)	(0.0247)	(0.0140)	(0.0214)
trade_cater	-0.0649	0.0009	-0.0390	-0.0269	-0.0376	-0.0275	-0.0458	-0.0256	-0.0412	0.0076	-0.0429	0.0021
	(0.0161)	(0.0216)#	(0.0135)	(0.0146)#	(0.0109)	(0.0116)	(0.0115)	(0.0097)	(0.0176)	(0.0148)#	(0.0127)	(0.0129)#
estate_utility	0.0049	-0.0040	-0.0160	-0.0411	-0.0540	-0.0779	-0.0671	-0.0865	-0.1370	-0.0942	-0.0439	-0.0734
	(0.0387)#	(0.0645)#	(0.0325)#	(0.0439)#	(0.0265)	(0.0354)	(0.0279)	(0.0300)	(0.0433)	(0.0452)	(0.0308)#	(0.0394)#
personal_consult	-0.0342	-0.4945	-0.0423	-0.3572	-0.0696	-0.1914	-0.0715	-0.1696	0.0248	-0.1630	-0.0668	-0.2308
	(0.0585)#	(0.0675)	(0.0484)#	(0.0459)	(0.0390)#	(0.0367)	(0.0418)#	(0.0308)	(0.0644)#	(0.0465)	(0.0455)#	(0.0409)
health_sport	-0.0043	0.0426	-0.0346	0.0140	-0.0341	0.0102	-0.0664	-0.0162	-0.0808	-0.0285	-0.0400	0.0175
	(0.0273)#	(0.0366)#	(0.0230)#	(0.0248)#	(0.0186)#	(0.0198)#	(0.0196)	(0.0163)#	(0.0295)	(0.0255)#	(0.0216)#	(0.0219)#
edu_cult_art	-0.0063	0.0027	-0.0151	0.0202	-0.0237	0.0147	-0.0415	-0.0302	-0.0532	-0.0675	-0.0285	0.0174
	(0.0205)#	(0.0337)#	(0.0176)#	(0.0226)#	(0.0144)#	(0.0180)#	(0.0152)	(0.0150)	(0.0232)	(0.0229)	(0.0167)#	(0.0200)#
sci_tech	0.0501	0.0324	0.0426	0.0513	0.0213	0.0441	0.0275	0.0237	0.0418	0.0778	0.0247	0.0672
	(0.0312)#	(0.0653)#	(0.0262)#	(0.0430)#	(0.0213)#	(0.0340)#	(0.0225)#	(0.0284)#	(0.0338)#	(0.0421)#	(0.0246)#	(0.0379)#
finan_insur	-0.0290	0.0143	-0.0088	0.0095	-0.0129	0.0148	-0.0364	-0.0373	-0.0818	-0.0422	-0.0302	-0.0130
	(0.0393)#	(0.0639)#	(0.0326)#	(0.0431)#	(0.0264)#	(0.0343)#	(0.0280)#	(0.0287)#	(0.0431)#	(0.0433)#	(0.0306)#	(0.0383)#
party_gvmt_soci	-0.0028	0.0031	-0.0217	0.0002	-0.0518	-0.0260	-0.0949	-0.0742	-0.1147	-0.1190	-0.0503	-0.0064
	(0.0174)#	(0.0384)#	(0.0146)#	(0.0261)#	(0.0119)	(0.0208)#	(0.0127)	(0.0176)	(0.0195)	(0.0265)	(0.0137)	(0.0232)#
other	-0.4203	-1.1047	-0.1754	-0.5354	-0.0606	-0.4108	-0.0603	-0.1176	0.0225	-0.0828	-0.3308	-0.5500
	(0.0584)	(0.0916)	(0.0473)	(0.0598)	(0.0392)#	(0.0484)	(0.0409)#	(0.0401)	(0.0601)#	(0.0583)#	(0.0457)	(0.0541)
dontknow	-1.4474	-1.6196	-0.8520	-1.5605	-0.5228	-1.0162	-0.3726	-0.7410	-0.4991	-0.5137	-0.7430	-1.0625
	(0.1414)	(0.1584)	(0.1076)	(0.1056)	(0.0908)	(0.0818)	(0.0929)	(0.0703)	(0.1557)	(0.1083)	(0.1107)	(0.0931)
Provinces												
Bejing	0.0262	-0.0177	0.0423	0.0236	0.0578	0.0635	0.1248	0.0974	0.2082	0.1655	0.0896	0.0370
	(0.0243)#	(0.0419)#	(0.0204)	(0.0280)#	(0.0167)	(0.0223)	(0.0176)	(0.0188)	(0.0272)	(0.0288)	(0.0193)	(0.0248)#
Shanxi	-0.1787	-0.3432	-0.1365	-0.2393	-0.1292	-0.2025	-0.1029	-0.1529	-0.0818	-0.0685	-0.1295	-0.2146
	(0.0194)	(0.0321)	(0.0162)	(0.0215)	(0.0131)	(0.0171)	(0.0138)	(0.0143)	(0.0209)	(0.0215)	(0.0151)	(0.0190)
Liaoning	0.0063	0.0214	-0.0070	-0.0119	-0.0081	-0.0066	-0.0012	0.0065	0.0012	0.0386	0.0106	0.0127
	(0.0195)#	(0.0314)#	(0.0164)#	(0.0213)#	(0.0133)#	(0.0169)#	(0.0141)#	(0.0142)#	(0.0216)#	(0.0216)#	(0.0154)#	(0.0188)#

Anhui	-0.0645	-0.1898	-0.0585	-0.1372	-0.0590	-0.0916	-0.0225	-0.0627	0.0196	-0.0224	-0.0330	-0.1147
	(0.0199)	(0.0319)	(0.0165)	(0.0217)	(0.0135)	(0.0174)	(0.0144)#	(0.0147)	(0.0220)#	(0.0226)#	(0.0156)	(0.0193)
Henan	-0.1940	-0.2797	-0.1646	-0.2193	-0.1527	-0.1969	-0.1220	-0.1524	-0.0784	-0.1145	-0.1603	-0.2119
	(0.0190)	(0.0306)	(0.0158)	(0.0205)	(0.0128)	(0.0164)	(0.0136)	(0.0138)	(0.0209)	(0.0209)	(0.0148)	(0.0182)
Hubei	-0.0914	-0.0488	-0.0962	-0.0668	-0.0958	-0.0601	-0.0721	-0.0529	-0.0385	-0.0071	-0.0795	-0.0545
	(0.0194)	(0.0313)#	(0.0161)	(0.0210)	(0.0131)	(0.0167)	(0.0140)	(0.0141)	(0.0213)	(0.0214)#	(0.0152)	(0.0186)
Guangdong	0.1058	0.0980	0.2039	0.1945	0.3034	0.2859	0.4311	0.3852	0.5925	0.5401	0.3143	0.2808
	(0.0189)	(0.0308)	(0.0158)	(0.0207)	(0.0129)	(0.0164)	(0.0136)	(0.0138)	(0.0207)	(0.0207)	(0.0149)	(0.0182)
Yunnan	0.0014	-0.0284	0.0276	0.0025	0.0317	0.0168	0.0498	0.0393	0.0763	0.0889	0.0404	0.0194
	(0.0197)#	(0.0315)#	(0.0166)#	(0.0213)#	(0.0135)	(0.0169)#	(0.0144)	(0.0143)	(0.0221)	(0.0217)	(0.0156)	(0.0188)#
Gansu	-0.0476	-0.3267	0.0006	-0.1239	0.0321	-0.0554	0.0774	0.0052	0.0997	0.0558	0.0144	-0.1236
	(0.0226)	(0.0366)	(0.0188)#	(0.0251)	(0.0152)	(0.0202)	(0.0162)	(0.0170)#	(0.0248)	(0.0259)	(0.0177)#	(0.0225)
Constant	6.1263	6.0016	6.4446	6.4137	6.7191	6.7045	6.9199	6.9361	7.1344	7.1353	6.5907	6.5815
	(0.0273)	(0.0454)	(0.0219)	(0.0296)	(0.0171)	(0.0227)	(0.0176)	(0.0186)	(0.0263)	(0.0288)	(0.0198)	(0.0253)
(Pseudo) R2	0.3352	0.2245	0.3033	0.2131	0.2604	0.1995	0.2343	0.1875	0.2315	0.1958	0.4002	0.2857

Notes: a. Standard errors in parenthesis.
b. # stands for not significant at 95% level.
c. Omitted variables: Manual_worker, Manufacture, Jiangsu
Source: Chinese Household Income Project, 1988

Appendix Table 1-b

Quantile Regression Estimates for 1995 Dataset

Variables	τ=0.10		τ=0.25		τ=0.50		τ=0.75		τ=0.90		OLS	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
exp	0.0388	0.1071	0.0292	0.0652	0.0254	0.0457	0.0180	0.0294	0.0148	0.0249	0.0270	0.0596
	(0.0043)	(0.0070)	(0.0034)	(0.0042)	(0.0025)	(0.0031)	(0.0031)	(0.0031)	(0.0041)	(0.0045)	(0.0028)	(0.0036)
exp2	-0.0006	-0.0023	-0.0003	-0.0013	-0.0003	-0.0008	-0.0001	-0.0004	-0.0001	-0.0003	-0.0003	-0.0012
	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)#	(0.0001)	(0.0001)	(0.0001)
married	0.4595	0.0347	0.3220	0.0479	0.1517	0.0498	0.1909	0.0808	0.1364	0.0657	0.2569	0.0603
	(0.0392)	(0.0594)#	(0.0305)	(0.0352)#	(0.0223)	(0.0265)#	(0.0285)	(0.0267)	(0.0381)	(0.0400)	(0.0254)	(0.0306)
schoolyr	0.0443	0.0591	0.0412	0.0491	0.0372	0.0458	0.0300	0.0453	0.0278	0.0445	0.0376	0.0470
	(0.0042)	(0.0078)	(0.0034)	(0.0045)	(0.0024)	(0.0034)	(0.0031)	(0.0034)	(0.0041)	(0.0049)	(0.0027)	(0.0039)
Occupation												
owner	0.0448	0.3485	0.3065	0.0364	0.2943	0.0798	0.3240	0.0179	0.2216	0.0845	0.2934	0.1317
	(0.0991)#	(0.1967)#	(0.0882)	(0.1170)#	(0.0654)	(0.0878)#	(0.0827)	(0.0870)#	(0.1027)	(0.1122)#	(0.0753)	(0.1028)#
manager	0.2319	-0.0645	0.0445	0.1152	0.1268	0.0003	0.0447	0.2718	0.0961	0.2108	0.0776	0.2520
	(0.1823)#	(0.1252)#	(0.1585)#	(0.1836)#	(0.1127)#	(0.1605)#	(0.1489)#	(0.1676)#	(0.1867)#	(0.0847)	(0.1335)#	(0.1973)#
profwork	0.2524	0.4486	0.1261	0.2804	0.1031	0.2010	0.0876	0.1384	0.0846	0.1362	0.1574	0.2800
	(0.0330)	(0.0499)	(0.0276)	(0.0305)	(0.0200)	(0.0228)	(0.0259)	(0.0231)	(0.0354)	(0.0354)	(0.0228)	(0.0263)
head	0.3075	0.3854	0.1653	0.2387	0.1429	0.2301	0.1415	0.1847	0.1216	0.1459	0.2002	0.2837
	(0.0344)	(0.0699)	(0.0285)	(0.0436)	(0.0209)	(0.0327)	(0.0272)	(0.0331)	(0.0374)	(0.0502)	(0.0238)	(0.0377)
office	0.1476	0.3024	0.0758	0.1987	0.0642	0.1413	0.0428	0.0966	0.0466	0.0837	0.0987	0.1819
	(0.0322)	(0.0453)	(0.0263)	(0.0276)	(0.0191)	(0.0205)	(0.0244)#	(0.0204)	(0.0328)#	(0.0306)	(0.0217)	(0.0237)
skilled	0.2001	0.2049	0.1153	0.1483	0.0870	0.1058	0.0929	0.1134	0.0785	0.0868	0.1356	0.1457
	(0.0280)	(0.0437)	(0.0233)	(0.0269)	(0.0170)	(0.0201)	(0.0218)	(0.0203)	(0.0293)	(0.0299)	(0.0194)	(0.0233)
other	-0.0015	0.1436	0.0314	0.0075	0.0226	0.0355	-0.0111	0.0362	-0.0768	-0.0041	0.0071	0.0582
	(0.0562)#	(0.0641)#	(0.0460)#	(0.0389)#	(0.0334)#	(0.0288)#	(0.0425)#	(0.0286)#	(0.0576)#	(0.0426)#	(0.0382)#	(0.0332)#
Economic Sector												
mining	0.0949	0.1715	0.1663	0.1810	0.1097	0.1698	0.0060	0.0684	0.0396	-0.0346	0.1056	0.1131

	(0.0802)#	(0.1457)#	(0.0649)	(0.0904)	(0.0481)	(0.0668)	(0.0622)#	(0.0666)#	(0.0826)#	(0.0985)#	(0.0550)#	(0.0776)#
construction	0.0610	0.0263	0.0388	-0.0246	0.0325	0.0292	-0.0067	0.0193	0.0781	-0.0362	0.0467	-0.0246
	(0.0496)#	(0.0898)#	(0.0410)#	(0.0556)#	(0.0298)#	(0.0413)#	(0.0382)#	(0.0408)#	(0.0508)#	(0.0591)#	(0.0339)#	(0.0477)#
tspcomm	0.0592	0.0372	0.0717	-0.0204	0.1105	0.0540	0.1603	0.1763	0.1383	0.1693	0.0933	0.0850
	(0.0385)#	(0.0722)#	(0.0316)	(0.0448)#	(0.0230)	(0.0336)#	(0.0295)	(0.0334)	(0.0395)	(0.0493)	(0.0261)	(0.0388)
trade	-0.1265	0.0081	-0.0239	-0.0105	-0.0134	-0.0043	-0.0073	-0.0192	0.0114	-0.0127	-0.0311	-0.0032
	(0.0293)	(0.0435)#	(0.0241)#	(0.0266)#	(0.0174)#	(0.0196)#	(0.0223)#	(0.0194)#	(0.0298)#	(0.0290)#	(0.0199)#	(0.0226)#
real	0.0608	-0.0786	0.0472	-0.0650	0.0980	-0.0771	0.0585	-0.0273	0.0924	-0.0013	0.0469	-0.0543
	(0.0510)#	(0.0727)#	(0.0416)#	(0.0443)#	(0.0303)	(0.0332)	(0.0390)#	(0.0330)#	(0.0511)#	(0.0499)#	(0.0346)#	(0.0383)#
health	0.1155	0.2071	0.1274	0.1423	0.0609	0.1211	0.0586	0.1300	-0.0097	0.0460	0.0571	0.0920
	(0.0473)	(0.0664)	(0.0397)	(0.0404)	(0.0291)	(0.0298)	(0.0373)#	(0.0298)	(0.0501)#	(0.0446)#	(0.0332)#	(0.0344)
education	0.2185	0.2139	0.1173	0.1295	0.0493	0.0972	0.0210	0.0309	0.0171	-0.0221	0.0683	0.0956
	(0.0388)	(0.0580)	(0.0314)	(0.0352)	(0.0228)	(0.0261)	(0.0289)#	(0.0261)#	(0.0386)#	(0.0389)#	(0.0260)	(0.0301)
technique	0.2656	0.3128	0.2055	0.1694	0.1383	0.1531	0.0772	0.0831	0.0271	0.0038	0.1624	0.1605
	(0.0556)	(0.1030)	(0.0459)	(0.0628)	(0.0337)	(0.0470)	(0.0428)#	(0.0472)#	(0.0563)#	(0.0699)#	(0.0385)	(0.0543)
finance	0.3193	0.2487	0.3062	0.2527	0.2692	0.3456	0.2767	0.3561	0.1473	0.2635	0.2779	0.3352
	(0.0659)	(0.1011)	(0.0544)#	(0.0613)	(0.0394)	(0.0456)	(0.0503)	(0.0458)	(0.0658)	(0.0686)	(0.0452)	(0.0529)
govt	0.1596	0.2088	0.0963	0.1337	0.0452	0.1117	-0.0139	0.0683	-0.0443	0.0560	0.0493	0.1284
	(0.0299)	(0.0563)	(0.0241)	(0.0344)	(0.0177)	(0.0254)	(0.0227)#	(0.0254)	(0.0307)#	(0.0374)#	(0.0201)	(0.0294)
esother	0.0861	0.0207	0.0404	0.0414	0.0681	0.0124	0.0781	0.0081	0.1329	0.0906	0.0532	-0.0019
	(0.0547)#	(0.1028)#	(0.0447)#	(0.0655)#	(0.0328)	(0.0488)#	(0.0420)#	(0.0486)#	(0.0566)	(0.0720)#	(0.0374)#	(0.0564)#
Provinces												
Beijing	0.1230	0.0269	0.1472	0.0722	0.1685	0.1653	0.2065	0.1384	0.1981	0.1616	0.1896	0.1384
	(0.0404)	(0.0674)#	(0.0333)	(0.0416)#	(0.0243)	(0.0311)	(0.0312)	(0.0311)	(0.0420)	(0.0467)	(0.0277)	(0.0359)
Liaoning	-0.1935	-0.3054	-0.2175	-0.2854	-0.1820	-0.2589	-0.1418	-0.2289	-0.1512	-0.2007	-0.1676	-0.2732
	(0.0365)	(0.0609)	(0.0299)	(0.0373)	(0.0219)	(0.0276)	(0.0281)	(0.0276)	(0.0381)	(0.0412)	(0.0249)	(0.0318)
Henan	-0.3214	-0.5313	-0.3414	-0.4835	-0.2938	-0.3890	-0.2569	-0.3941	-0.3112	-0.3482	-0.3172	-0.4332
	(0.0391)	(0.0654)	(0.0321)	(0.0401)	(0.0235)	(0.0298)	(0.0301)	(0.0298)	(0.0404)	(0.0440)	(0.0267)	(0.0344)
Anhui	-0.1925	-0.3250	-0.2689	-0.3500	-0.2721	-0.3229	-0.2651	-0.3343	-0.3196	-0.3317	-0.2546	-0.3427
	(0.0414)	(0.0664)	(0.0340)	(0.0410)	(0.0249)	(0.0305)	(0.0319)	(0.0306)	(0.0428)	(0.0456)	(0.0283)	(0.0352)
Hubei	-0.1750	-0.2245	-0.2043	-0.1689	-0.1626	-0.1338	-0.1437	-0.1744	-0.1936	-0.1502	-0.1771	-0.1698
	(0.0365)	(0.0599)	(0.0301)	(0.0368)	(0.0219)	(0.0275)	(0.0282)	(0.0277)	(0.0379)	(0.0413)	(0.0250)	(0.0317)

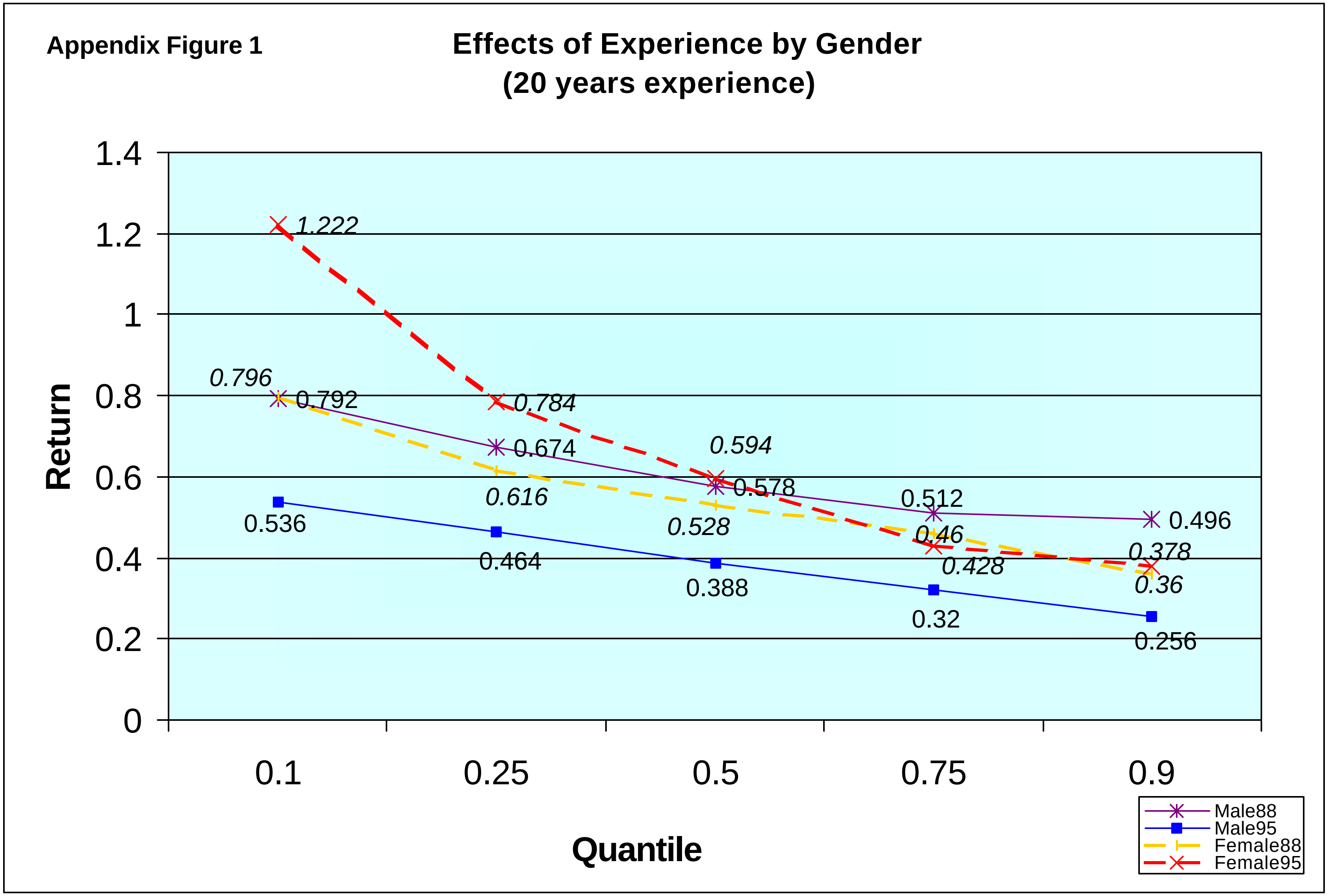
Guangdong	0.3498	0.2926	0.3480	0.3195	0.4325	0.4304	0.5628	0.4636	0.6277	0.5837	0.4716	0.4172
	(0.0389)	(0.0650)	(0.0322)	(0.0396)	(0.0235)	(0.0295)	(0.0301)	(0.0296)	(0.0408)	(0.0442)	(0.0267)	(0.0340)
Shanxi	-0.3214	-0.4813	-0.3443	-0.4596	-0.2910	-0.4133	-0.2510	-0.4367	-0.2293	-0.4175	-0.2798	-0.4913
	(0.0374)	(0.0634)	(0.0307)	(0.0386)	(0.0225)	(0.0287)	(0.0289)	(0.0286)	(0.0394)	(0.0430)	(0.0257)	(0.0330)
Gansu	-0.2299	-0.4397	-0.3187	-0.4227	-0.3777	-0.4103	-0.3969	-0.4696	-0.4384	-0.4543	-0.3613	-0.4464
	(0.0445)	(0.0734)	(0.0364)	(0.0446)	(0.0266)	(0.0334)	(0.0342)	(0.0335)	(0.0465)	(0.0505)	(0.0304)	(0.0385)
Sichuan	-0.2057	-0.3054	-0.2204	-0.2687	-0.2023	-0.2128	-0.1147	-0.1811	-0.1724	-0.1611	-0.1809	-0.2160
	(0.0354)	(0.0579)	(0.0292)	(0.0354)	(0.0213)	(0.0263)	(0.0273)	(0.0262)	(0.0368)	(0.0389)	(0.0242)	(0.0303)
Yunnan	-0.0735	-0.1139	-0.1526	-0.1573	-0.1947	-0.1921	-0.1983	-0.2458	-0.2765	-0.2306	-0.1664	-0.1957
	(0.0377)#	(0.0606)#	(0.0310)	(0.0375)	(0.0227)	(0.0279)	(0.0293)	(0.0280)	(0.0397)	(0.0421)	(0.0259)	(0.0322)
Constant	6.5766	6.0557	7.2117	6.9749	7.7215	7.4387	8.1006	7.8509	8.4735	8.1531	7.5497	7.2543
	(0.0703)	(0.1167)	(0.0553)	(0.0699)	(0.0395)	(0.0511)	(0.0498)	(0.0508)	(0.0652)	(0.0725)	(0.0450)	(0.0589)
(Pseudo R2)	0.2436	0.2018	0.2224	0.2120	0.2118	0.2125	0.2129	0.2208	0.2417	0.2403	0.3572	0.3198

Notes: a. Standard errors in parenthesis.

b. # stands for not significant at 95% level.

c. Omitted variables: unskilled, Manufacture, Jiangsu

Source: Chinese Household Income Project, 1995


Appendix Figure 1
Effects of Experience by Gender
(20 years experience)
Return
Quantile
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.1
0.25
0.5
0.75
0.9
1.222
0.796
0.792
0.784
0.674
0.594
0.578
0.616
0.536
0.528
0.512
0.46
0.428
0.496
0.378
0.36
0.464
0.388
0.32
0.256
Male88
Male95
Female88
Female95