An Exploration of Gender Gap In Labor Market

Instructor: Professor Frank Stafford

Xuanzhong Wang

March 2009

1. Introduction

Gender gap is described as the systematic difference between men and women in the labor market. It is often reflected by the difference in percentage of men and women participating in the workforce—the labor force participation rate; the different occupations that men and women engage in—occupational exclusion; and the difference in wage rate between men and women doing the same job. It is of great interest for labor economists, sociologists and women right activist. It is commonly believed that "the gender gap in employment, earnings, and occupations has narrowed in the twentieth century, but with increasing significance, it seems, in the eighties. Whether or not the gap will continue to narrow and eventually disappear is uncertain, and probably depends on the gender gap in time spent in child care and in the home[1]."

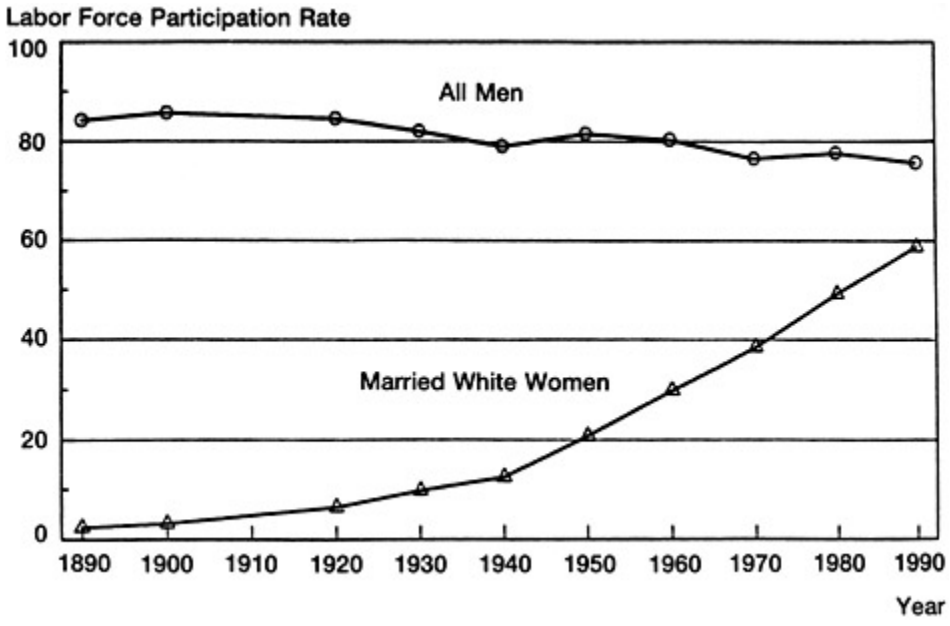
In this project, empirical evidence of the evolving gender gap in the United States from the 1980s to 2000 was examined. The data were extracted from the Panel Study of Income Dynamics (PSID). The PSID gathers information about families and all individuals in those families through its annual interviews (bi - annual since 1997). The study's original households constitute a national probability sample of U.S. households as of 1967.

[1] *Gender Gap*, Claudia Goldin, http://www.econlib.org/Library/Enc/GenderGap.html

2. Results and discussion

In the past one hundred years, a drastic increasing trend in the labor force participation rate in women can be observed. As shown from Chart 1 below, the gap in labor force participation rate between genders narrowed—especially after the forties last century, the labor force participation rate for married white women increased steadily at a rate of approximately ten percent per decade.

Chart 1. Labor Force Participation Rates of Men and Women, 1890-1990



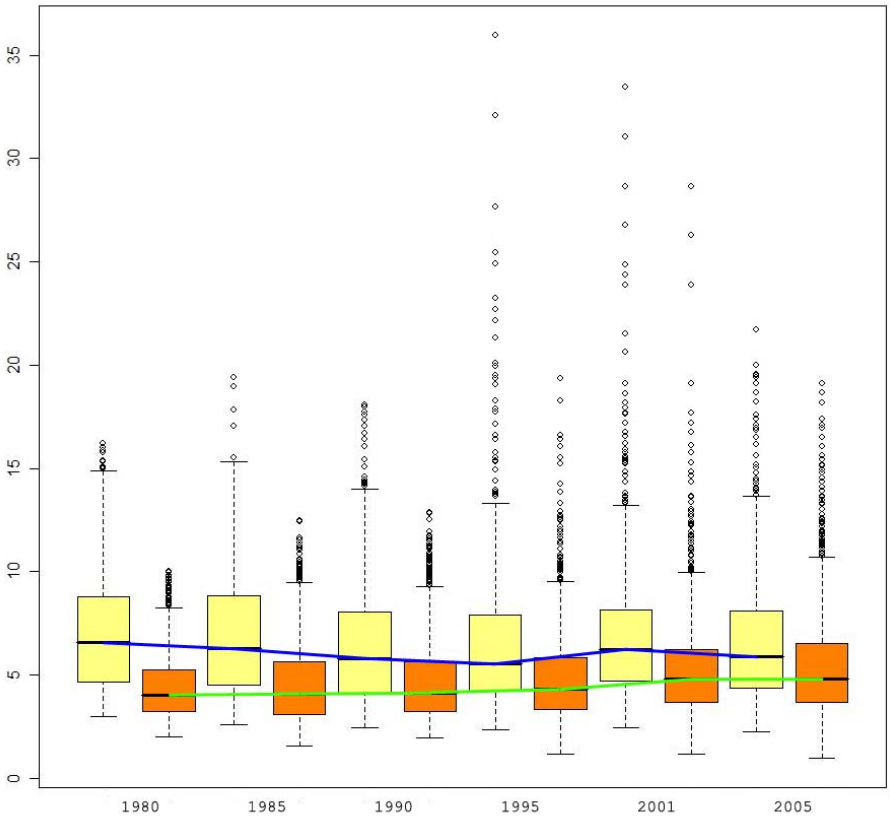
SOURCES:

Men: *1890 to 1970, U.S. Bureau of the Census,* Historical Statistics of the United States, Colonial Times to 1970. *Government Printing Office, 1975; and 1980 to 1990, U.S. Department of Labor, Bureau of Labor Statistics*, Employment and Earnings.

Women: *C. Goldin,* Understanding the Gender Gap: An Economic History of American Women, *table 2.1., 1990.*

Better educational coverage for women, increase in demand of clerical and administrative workers, drop in amount of manual work in the modern labor market and etc; these are all possible factors that contributed to the rising labor force participation rate for women. However, a converging trend in the labor force participation rate between genders does not imply the narrowing of gender gap. Rather, the gender gap persisted in a less obvious manner—differential and occupational exclusion existing in the United States economy.

Chart 2. Hourly Wage Rate (corrected for inflation) Distribution of Men and Women, 1980-2005



SOURCES:

PSID Main Family Data: PAY/HR HRLY DH (E) and PAY/HR HRLY WF (E)

Inflation Data: www.bls.gov (CPI for urban wage earners)

Chart 2 examines the distribution of hourly wage rate paid to those who are working for money. It presented a series of box-and-whisker-plots of the hourly wage rate (corrected for inflation) of men and women in the past twenty-five years. The yellow color boxes represent the distribution of hourly wage rate for men while the orange ones represent that for women. The thick black line at the center of the box represents the median of the sample; the lower and upper boundary of the boxes represents the lower and upper quartile of the sample respectively—the height of each box is proportional to the inter quartile range of the sample that the box represents; the two horizontal lines outside the boxes indicates a region that ninety five percent of the data points from the sample fall into. Points outside the two lines are considered as outliers.

Several interesting features can be inferred from this plot. The blue (green) line joins the median

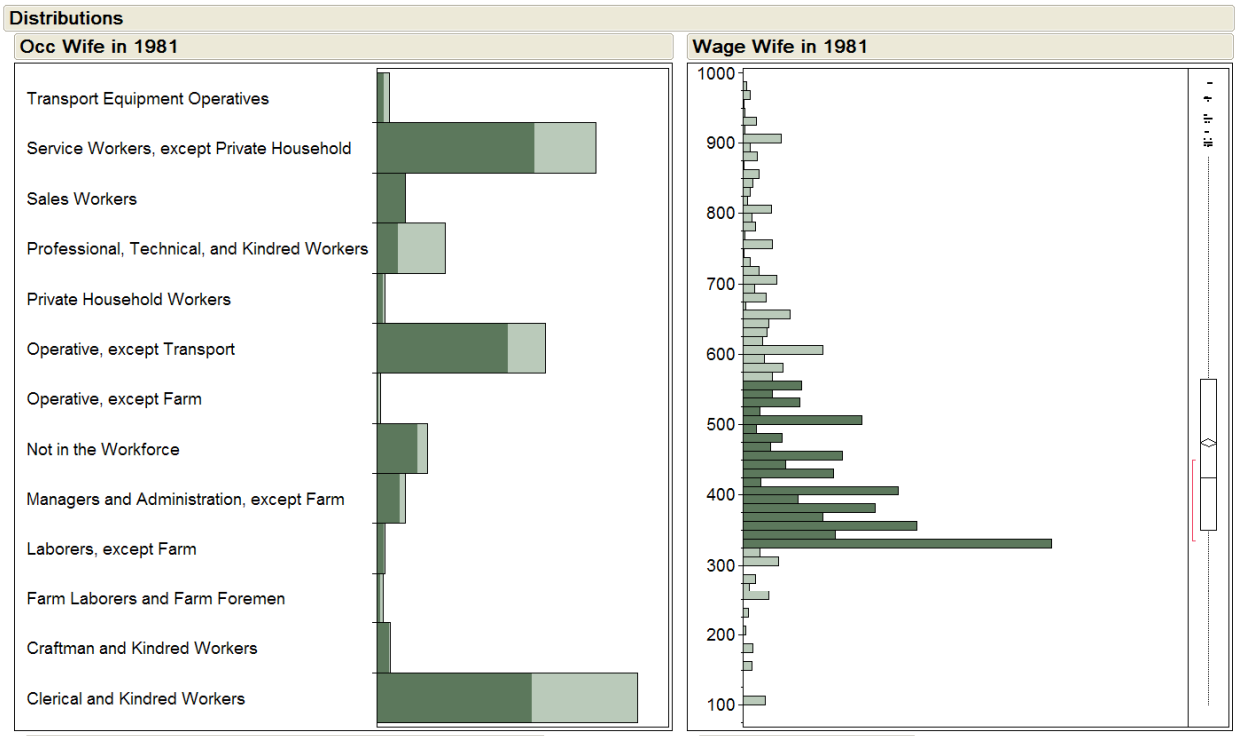
hourly wage rate for men (women) over the twenty-five years. The purchasing power of the hourly earning for both genders did not change much over the period. However, except the period from 1995 to 2000, the purchasing power of the hourly earnings for men slowly dropped while on another hand, the purchasing power of the hourly earning for women constantly increased. Despite the fact that men still earn more than women in 2005, there was some vague hope of the convergence in wage rate.

Comparing the distribution of hourly wage rate for women, a more optimistic prospect of the narrowing gender gap can be observed. In 1980, the distribution of hourly wage rate for women was compact; it has small variance and few outliers. The more homogeneous distribution indicates that most women who participated in the labor force were bounded to some stereotype—women are more likely to be employed in positions like clerical workers, service workers and etc.

Chart 3 shows the distribution of hourly wage rate and occupations for married women aged in between 20 and 42 who worked for money in 1981. It is obvious that the three industries: Clerical and Kindred Workers, Operatives except Transport and Service Workers except Private Household constituted the majority (73.60%) of married women who worked for money. Also, those women who belong to the center 50% of the hourly wage rate distribution (highlight in dark green) are mainly from the three industries listed above. This to a large extent proved that less variation in the hourly wage rate is an indication of lacking variation in occupation choices available to women.

Chart 2 shows that there are not only increasing variation in hourly wage rate for women; there is increasing number of outliers which span a larger range as time passes by as well. The existence of large amount of outliers has similar implications as the existence of fatter tails of the distribution—that is more occupation options are available to women.

Chart 3 Distribution of Wage and Occupations of Married Women, 1981



Frequencies

Level	Count	Prob
Clerical and Kindred Workers	1003	0.29613
Craftman and Kindred Workers	54	0.01594
Farm Laborers and Farm Foremen	26	0.00768
Laborers, except Farm	33	0.00974
Managers and Administration, except Farm	111	0.03277
Not in the Workforce	197	0.05816
Operative, except Farm	15	0.00443
Operative, except Transport	647	0.19102
Private Household Workers	34	0.01004
Professional, Technical, and Kindred Workers	265	0.07824
Sales Workers	110	0.03248
Service Workers, except Private Household	843	0.24889
Transport Equipment Operatives	49	0.01447
Total	3387	1.00000

N Missing 22157
13 Levels

Quantiles

100.0%	maximum	985.00
99.5%		964.00
97.5%		900.00
90.0%		715.00
75.0%	quartile	565.00
50.0%	median	425.00
25.0%	quartile	350.00
10.0%		335.00
2.5%		260.00
0.5%		100.00
0.0%	minimum	100.00

Moments

Mean	473.99793
Std Dev	165.93153
Std Err Mean	2.8511583
upper 95% Mean	479.5881
lower 95% Mean	468.40777
N	3387

Source: PSID Main Family Data: MAIN OCC: 3DIG (WF-E) and PAY/HR HRLY WF (E)

Income disparity within each occupation is of great interest. To examine this, a 95% confidence interval of the difference in mean hourly wage rate of men and women were constructed. The degree of freedom is adjusted so that the 95% confidence interval takes into consideration the possibility of unequal variance. The 95% confidence interval is given by the following formulae:

$$\bar{W}_M - \bar{W}_F \pm t_{\frac{\alpha}{2},v}\sqrt{\frac{S_M^2}{n_M} + \frac{S_F^2}{n_F}} \quad \text{Where } v = \frac{\left(\frac{S_M^2}{n_M} + \frac{S_F^2}{n_F}\right)^2}{\frac{\left(\frac{S_M^2}{n_M}\right)^2}{n_M - 1} + \frac{\left(\frac{S_F^2}{n_F}\right)^2}{n_F - 1}}$$

$\bar{W}_M$: *the* mean hourly wage rate for males working in a specific occupation during a specific year
$\bar{W}_F$: *the* mean hourly wage rate for females working in a specific occupation during a specific year
n_M : the number of males working in a specific occupation during a specific year
n_F : the number of females working in a specific occupation during a specific year
S_M^2 : *the* sampel variance of hourly wage rate for males working in a specific occupation during a specific year
S_F^2 : *the* sampel variance of hourly wage rate for females working in a specific occupation during a specific year

Chart 4 presented a series plots of the upper bound and the lower bound of the 95% confidence interval of occupations arranged in a way such that the mean hourly wage rate is in ascending order. If zero is contained in the interval, it can be concluded that there is no significant difference in mean hourly wage rate between men and women in that occupation; if the entire interval is above zero, then men are significantly better paid than women in the same occupation. Another matter that one needs to take note is that for the case only one gender is employed in one industry, the confidence interval does not make sense, so such occupations will be eliminated from the plots.

Chart 4 A 95% Confidence Interval of mean wage difference between men and women in various industries, 2001

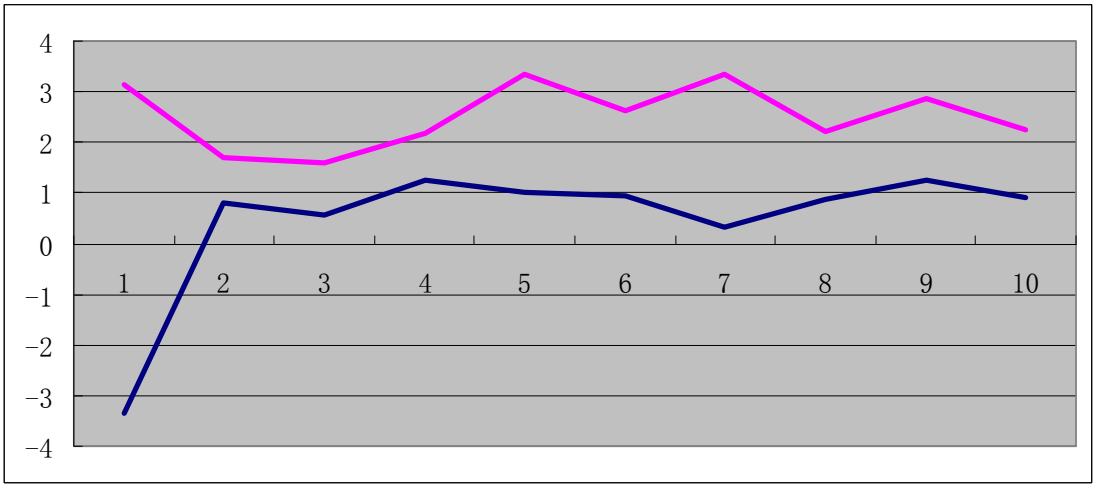
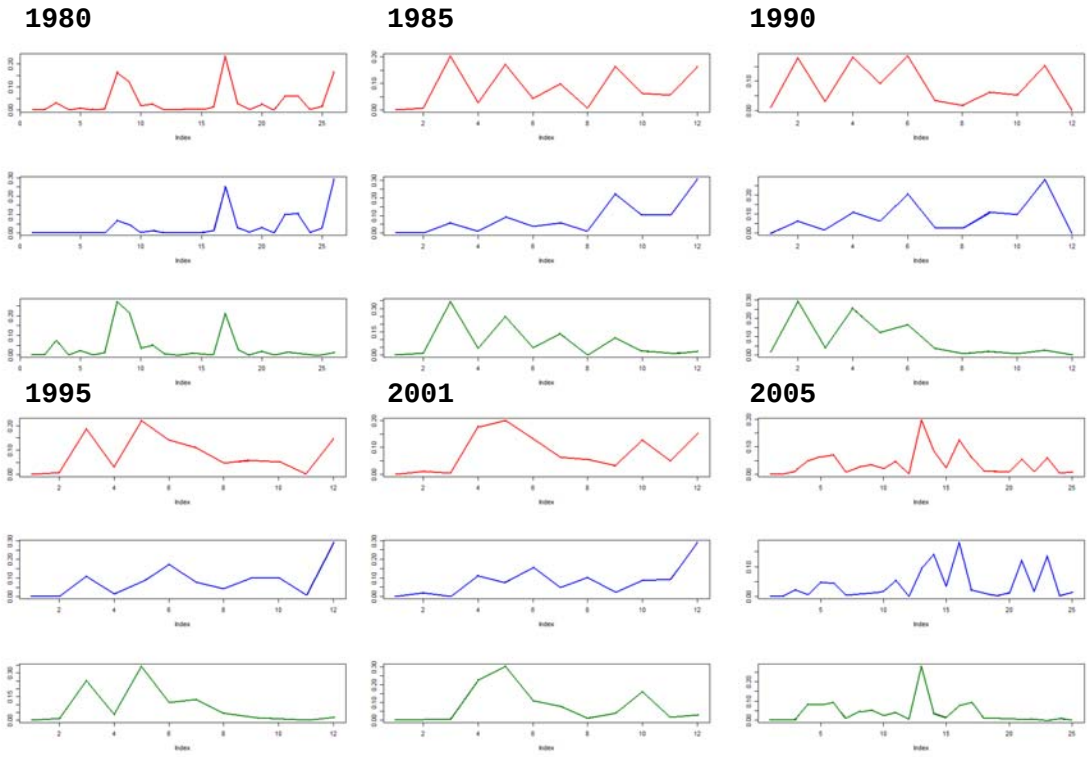


Chart 4 shows the plot of 95% confidence interval in 2001. The wage rate disparity disappears in the low paid occupations but persisted through the higher paid occupations. Together with the increasing labor force participation rate, the distribution of hourly wage rate for women provoked a doubt about narrowing gender gap in the labor market—if the increase in the labor force participation

rate is merely driven by the increase in women employment in female dominated low paid occupations like clerical work or service works and etc, probably it is not a positive sign of narrowing gender gap.

Chart 5 Distribution of Population Proportion in Various Industries (Mean Hourly Wage Rate Ranked Lowest to Highest), 1980-2005



Source: PSID Main Family Data: MAIN OCC: 3DIG (WF-E) and PAY/HR HRLY WF (E)

In Chart 5, for every given year, the first plot (red line) is the population proportion of people working in different industries. The positive direction of the horizontal axis indicates increasing mean hourly wage rates. The second plot (blue line) is the proportion of men employed in corresponding occupations while the third plot (green line) represents the employment pattern of women.

From the series of plots, some features are rather obvious. First, from 1980 to 2005, the overall employment pattern are getting more affected by the employment pattern of women—as a result of increasing labor force participation rate. Second, the employment pattern of men is generally more stable than that of women—it has two peaks—also, the location of the peaks is more or less centered on the medium wage rate to high wage rate industries. Third, there is drastic shift in the employment pattern of women. In 1980s and the 1990s, most women working for money are employed in the lower

end of the spectrum; after 2001, the female employment pattern started to have one very significant peak located near the medium and another less obvious peak at high wage rate industries.

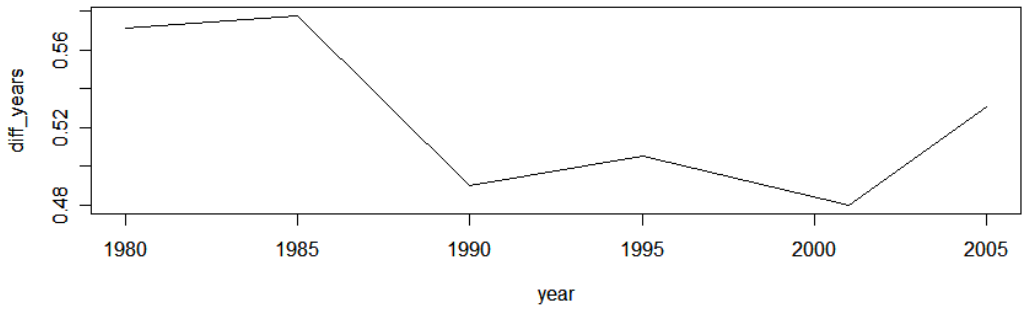
One quantified measure of the occupation exclusion is the index of occupational difference which is represented using the formula below:

$$index = \sum_{i=1}^{n} \frac{| w_{mi} - w_{fi} |}{2}$$ [2]

Where

n = *total number of occupations*
w_{mi} = *percentage of males employed in occupation i*
w_{fi} = *percentage of females employed in occupation i*

Chart 6 Changes in Index of Occupational Difference



Theoretically, the value of the index of occupational difference varies between zero (while there is no occupational exclusion at all) and one (when there is complete occupational exclusion); the lower the value of the index, the less occupational exclusion exists in the economy. In Chart 6, the index of occupational difference is plotted on the vertical axis with the corresponding year on the horizontal axis. We can see that from 1980 to 2001, the index of occupational difference decreases gradually; however, in 2005, there is a sharp rise in the index of occupational difference.

One problem about the index of occupational difference that needs to be taken into consideration is that the value of the index depends heavily on the classification of occupations. The more categories of occupations there are, the higher the value of the index of occupational difference will be[3]. Therefore, it is naive to conclude that there is increasing occupational exclusion in 2005 since there are 25 categories of occupations comparing with year 1985 to 2001 which had only 12 occupation categories.

In general, it can be concludes that in the past twenty-five years, occupational exclusion decreases.

[2] *Alternative Approaches to Occupational Exclusion*, George E. Johnson and Frank P. Stafford
[3] Please refer to Appendix A for a mathematical proof of the statement.

Chart 5 and Chart 6 can be seen as an evidence of narrowing gender gap in the labor market. However, at the same time, the question about mobility of women in the labor market rises. In another word, is the changing pattern in the employment of women a result of young women who are better educated entering the workforce or is it more of a reason that existing women moving around in the labor market?

To look at the movement of women in the labor market, as well as their relative position in the earning spectrum, the following transition matrices are constructed. Group of women of age between 25 and 35 (including 25 and 35 years old) in 1981 are selected. Their occupation and hourly wage rate in 1981 and 1991 is recorded; a balanced panel is taken in this case—women who left or entered the PSID sample in 1991 were discarded from the data. The occupations are categorized to 12 categories based on the modified 3-digit occupation code from 1970 Census of Population provided in the codebook in alphabetical order. Wage deciles of the entire population in 1981 and 1991 are constructed. Two transition matrices are tabulated with a computer algorithm. The first one is a 12 by 12 matrix indicating how women moving around occupations from 1981 to 1991; while the second one is an 11 by 11 matrix showing how women changed their relative position in the hourly wage rate distribution.

Similarly, transition matrices for women of age 25 to 35 in 1991 to 2001 are constructed. At the same time, equivalent transition matrices of men who are in the same age group are constructed as a mean of comparison[4].

For each cell in the transition matrix, the corresponding row label indicates the occupation category or wage rate decile at the beginning of the period and the corresponding column label indicates the occupation category or wage decile at the end of the transition period. The number in each cell is the count of individuals who falls into the specific transition group represented by the cell. The diagonal entries in each transition matrices are the count of individuals who stayed in the same occupation or wage decile at the beginning and the end of a transition period that spans over ten years. Note that it is only the beginning and end point of occupation and wage decile that are taken into consideration; we have ignored all the movements within the time span.

First the two transition matrices of changing occupations for women were shown below. By examining the 6th row (not in the work force) of each matrix, the following change could be observed:

	1	2	3	4	5	Not in the Workforce in 1991	7	8	9	10	11	12	Total
Not in the Workforce in 1981	605	74	0	7	123	992	232	50	309	93	427	32	2944

1981-1991

[4] Please refer to Appendix B for the constructed transition matrices.

	1	2	3	4	5	Not in the Workforce in 2001	7	8	9	10	11	12	Total
Not in the Workforce in 1991	255	64	6	21	135	361	103	5	176	42	272	40	1480

1991-2001

In 1991, around 33.7% of women who were not in the workforce in 1981 remained so; the same ratio reduced to approximately 24.4% in 2001. Comparing with the equivalent rows extracted from matrices of changing employment for men, we find that in 1991, approximately 28.2% of men remained out of workforce after a ten years' period, and the ratio reduced to 11.8%. Men did not only have a smaller proportion of people remained out of the workforce in both transition periods; they also had a much faster damping ratio. Thus, the decrease in proportion of women remained out of workforce is more of a reason that there is lower barrier to enter the workforce in general; despite so, when comparing with men of the same age group, women enjoys less benefit from the lowered barrier.

	1	2	3	4	5	Not in the Workforce in 1991	7	8	9	10	11	12	Total
Not in the Workforce in 1981	27	67	0	59	23	153	45	7	41	20	59	41	542

1981-1991

	1	2	3	4	5	Not in the Workforce in 2001	7	8	9	10	11	12	Total
Not in the Workforce in 1991	26	57	1	14	31	44	70	0	28	4	53	45	373

1991-2001

Chart 7 Transition of Occupation for Women Between 25 to 35 (1981-1991)

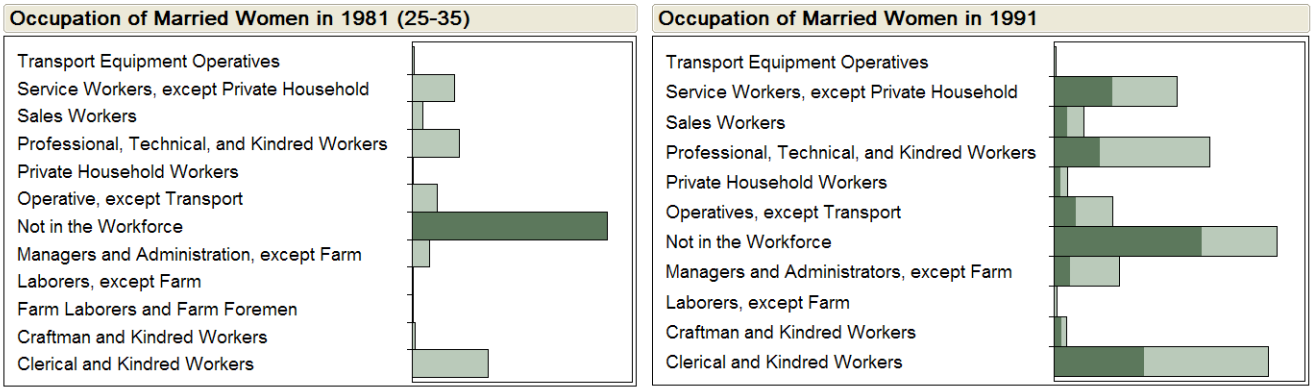
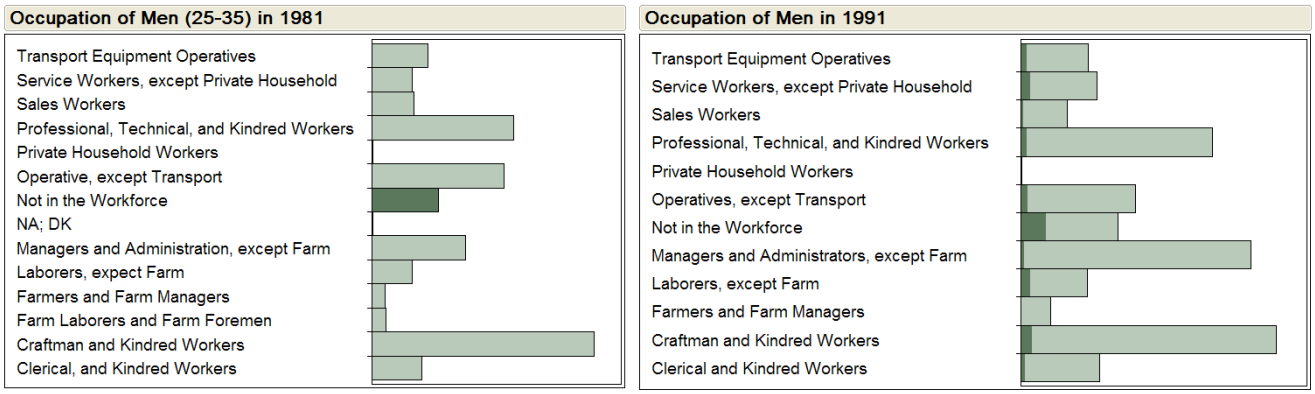


Chart 8 Transition of Occupation for Men Between 25 to 35 (1981-1991)



Comparing Chart 7 and Chart 8, most women entered the workforce did so by entering Clerical and Service related occupations (the relatively large proportion of women who entered the workforce as Professionals are usually students at the beginning of the transition period); however, this pattern does not exist in the transition of occupation for men—when men entered the occupation, there is a more evenly distributed frequency in each occupation category. This is another illustration of the existence of occupational exclusion in the labor market—women can easily get into occupations that are stereotyped as feminine industries. Many people believed that office is the place for women; this might explains the reason that most women enter the workforce as clerical workers or service providers.

	1	2	3	4	5	6	7	8	9	10	11	12	Total
Managers and Administration in 1981	117	54	1	12	99	351	30	12	79	11	149	14	929
Professional, Technical in 1981	38	12	0	0	70	82	4	4	413	21	44	0	688

1981-1991

	1	2	3	4	5	6	7	8	9	10	11	12	Total
Managers and Administration in 1991	41	6	0	4	85	28	18	0	54	21	15	5	277
Professional, Technical in 1991	77	3	0	0	100	91	7	0	415	33	68	4	798

1991-2001

By looking at the 5th and 9th row of the occupation transition matrices, we will explore the more prestigious occupations including managers and administrative staff as well as professionals and technical workers. The cells highlighted in yellow indicate the number of women who consistently remain in the prestigious occupations. We can see that a large amount of women who worked as professionals remained in the field—being a professional usually requires a long period of training and

the cost of leaving the profession is usually higher than those do not require training. Besides the question of who stayed in the highly regarded occupations, we would also like to take a look at those who were in some other occupations at the beginning and moved into better occupations at the end of the transition period.

From Chart 9 and Chart 10, we can see that those who were already in professional fields at the beginning of the transition period and those who were going through training program or graduate school (usually reflected by not in the workforce at first) are the main groups that make up the professional and manager population at the end of the transition period; there are two specific occupations that serve as the pathway between entry level occupations and more prestigious occupations. Service workers and clerical workers are the two occupations that yield particularly high proportions of women entering prestigious occupations. Therefore, they can be described as the pathway occupations.

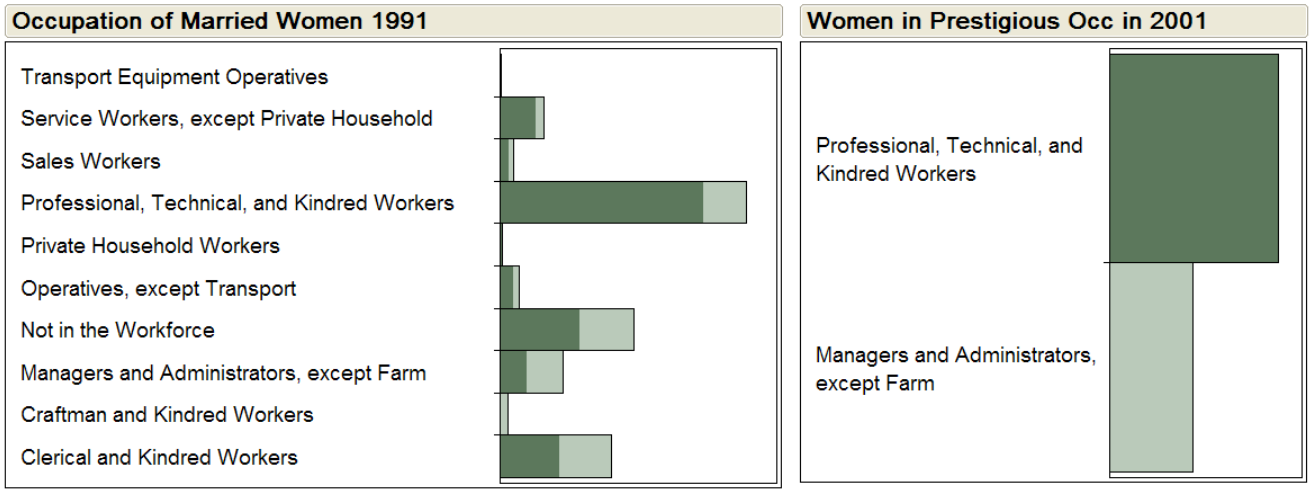
The existence of pathway occupations may attribute to the following factors: as mentioned before, entering the more prestigious occupation requires training and experience. As a woman who entered the workforce at a relatively young age—25 to 35 years' old, they are not as well endowed in terms of experience or training comparing to those who already worked in the labor force for some time.

Chart 9 Occupation Distributions of Married Women

Who End up in Prestigious Occupations 1981 – 1991

Chart 10 Occupation Distributions of Married Women

Who End up in Prestigious Occupations 1991 - 2001



Therefore, it is probably more rational to take up a job that requires little or no training when an individual seeks employment in the labor market. That explained why many women started their occupation on service or clerical positions. These occupations have relatively low workload comparing to those occupations demand high labor input—this allows women who were employed as clerical workers or service workers to take up some training during their free time. This also can be reflected from the high mobility associated with the two pathway occupations. Less than half of women who started as service workers or clerical workers stayed in the same position throughout the transition period. Women from various occupations at the beginning filled the empty positions at the end of the transition period.

The existence of pathway occupations is another evidence of occupational exclusion—there is no obvious pathway occupation for men through both transition periods. However, with the evolution of various MBA and graduate programs, there are more formal pathways established for those who desired to get into more prestigious occupations. Instead of entering a pathway occupation at the beginning of the transition period, nowadays it is easier and more common to enter as professionals or managers. Therefore, it is reasonable to predict that the pathway occupations will get less significant over the years.

Earning is another feature that may reflect the gender gap in the labor market. By looking at the transition matrices relative positions in the population wage decile, we can compare the gender difference in movement between wage deciles. The Transition matrix of relative position change in wage deciles for women from 1981 to 1991 (when only looking at the population wage deciles) behaves more like an upper triangular matrix—the cells above the diagonal entries in general have

more counts than those below the diagonal entries. This shows that during the period from 1981 to 1991, women are moving up in their relative earning positions compare to the entire population.

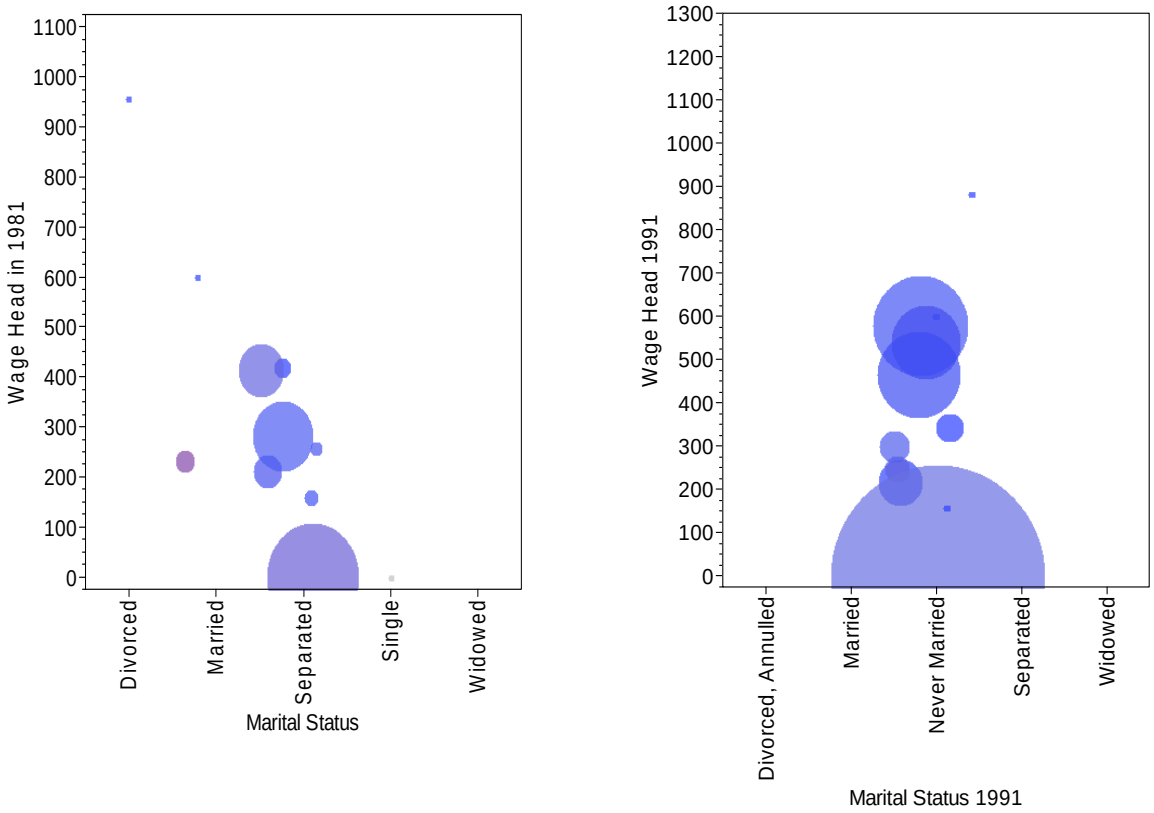
	1	2	3	4	5	6	7	8	9	10
1	42	45	71	17	10	21	19	0	13	15
2	32	30	29	26	32	24	21	8	0	4
3	10	16	44	30	23	18	8	10	0	0
4	21	6	13	25	31	38	26	5	8	0
5	9	15	20	3	28	6	33	25	19	2
6	17	10	0	2	22	24	29	39	18	13
7	5	0	0	2	15	9	6	22	19	5
8	0	9	0	0	0	6	9	23	32	20
9	0	0	0	0	0	6	0	0	5	2
10	0	0	63	0	0	6	0	0	0	1

However, the similar upper triangular property cannot be observed from other transition matrices of relative position in wage deciles. Instead, the dense cells are located close to the diagonal entries. This indicates that the wage rate distribution is approximately stable that people only move about in wage deciles close to where he/she belonged to at the beginning of the transition period. The wage decile transition matrices for men in both periods demonstrate this property; and the matrix of wage decile transition for women only started to show this property in the transition period from 1991 to 2001. This goes in line with Goldin's statement at the beginning of this project that the gender gap narrows, with increasing significance, in the eighties; after which it appears to be a more stable and slower evolution.

Earning a living is not the only role that a women play in society. More importantly, they bear the responsibility of family and housework—especially child care. Although there are few men who stayed home and took over the wives responsibility in more recent years, the social norm remained unchanged.

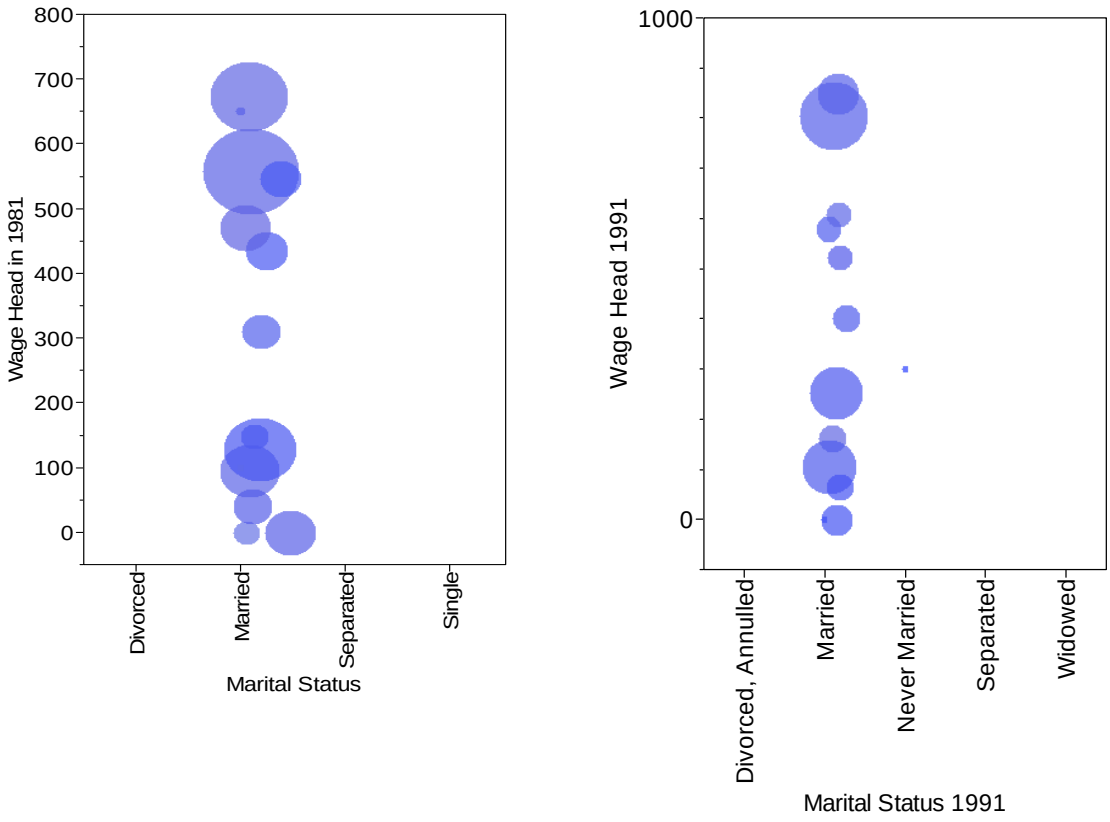
Chart 11 Bubble plots of wage rate of women between 25-35

Against marital status and number of children below 2 and 6



In Chart 11, three bubble plots of wage against marital status and number of children below 2 and 6 are plotted. The marital status is plotted on the X-axis and the hourly wage rate on the Y-axis. Each bubble represents a specific occupation; the diameter of the bubble indicates the mean number of children below two years old for all women in that industry and the color of the bubble indicates the mean number of children below six years old. It is obvious that in both years, the number of children below two years old has significant impact on one's occupation and wage rate. Large bubbles close to the bottom implied that the more children of age below two years old, probably it requires the mother to spend more time with the baby, thus the mother committed less time to her job, and earn less. This is especially true in case of single mothers. However, in Chart 12, the same pattern disappears in the bubble plot of the male counterpart of Chart 11. The wage rate of men does not depend heavily on the number of children below two or six years old.

Chart 12 Bubble plots of wage rate of men between 25-35 against marital status and number of children below 2 and 6



3. Conclusion

From the above analysis, the following conclusion was reached. To a certain extent, the gender gap is narrowing as the gap between labor force participation rates slowly closes; wage differential started to disappear in lower tail occupations. However, these are rather superficial indicators. When the phenomenon was examined in greater detail, the seemingly sign of converging gender gap reflects the hard core of the gender gap. The existence of pathway occupations, the effect of number of children on wage rate and employment shows that the employment pattern and wage rate of women are more vulnerable to family life comparing to men. The need of women to take care of the family and the children is one of the factors that constantly drive women in and out the workforce.

The significant narrowing in the gender gap in eighties probably is the outcome of a developed economy. The increasing demand of administrative workers, computer operators; the readily available media that assists job searching as well as a set of well defined education programs enabled women to find a job and earn some decent wage. However, it appears to be the case that men are more blessed by the economic development. It will take much longer for women to really cross the gender gap—as long as it takes for the entire society to change their perspective on the socio-economic roles played by women.

Appendix A

Statement: The more categories of occupations there are, the higher the value of the index of occupational difference will be.

Mathematical Proof:

if $n<l$

$index_n <= index_l$ *for the same economy*

let $d_n = \sum_{i=1}^{n} | w_{mi} - w_{fi} | = | w_{m1} - w_{f1} | + | w_{m2} - w_{f2} | + ... + | w_{mn} - w_{fn} |$

Suppose, the new categories are developed from the original categories.

which means each original category i is broken to k_i *sub-categories*

w'_{mi_j} *is the proportion of males employed in occupation* i_j

w'_{fi_j} *is the proportion of females employed in occupation* i_j

$$d_l = \sum_{i=1}^{n} \sum_{j=1}^{k_i} | w'_{mi_j} - w'_{fi_j} |$$

$$w_{mi} = \sum_{j=1}^{k_i} w'_{mi_j}$$

$$w_{fi} = \sum_{j=1}^{k_i} w'_{fi_j}$$

$$d_n = | (\sum_{j=1}^{k_1} w'_{m1_j}) - (\sum_{j=1}^{k_1} w'_{f1_j}) | + ... + | (\sum_{j=1}^{k_n} w'_{mn_j}) - (\sum_{j=1}^{k_n} w'_{fn_j}) |$$

$$= | (w'_{m1_1} + ... + w'_{m1_{k_1}}) - (w'_{f1_1} + ... + w'_{f1_{k_1}}) | + ... + | (w'_{mn_1} + ... + w'_{mn_{k_n}}) - (w'_{fn_1} + ... + w'_{fn_{k_n}}) |$$

$$= | (w'_{m1_1} - w'_{f1_1}) + ... + (w'_{m1_{k_1}} - w'_{f1_{k_1}}) | + ... + | (w'_{mn_1} - w'_{fn_1}) + ... + (w'_{mn_{k_n}} - w'_{fn_{k_n}}) |$$

$$<= | w'_{m1_1} - w'_{f1_1} | + ... + | w'_{m1_{k_1}} - w'_{f1_{k_1}} | + ... + | w'_{mn_1} - w'_{fn_1} | + ... + | w'_{mn_{k_n}} - w'_{fn_{k_n}} |$$

$$= d_l$$

$$index_n = \frac{d_n}{2} <= index_l = \frac{d_l}{2}$$

Appendix B

Occupation Code

Clerical and Kindred Workers	**1**
Craftman and Kindred Workers	**2**
Farm Laborers and Farm Foremen	**3**
Laborers, except Farm	**4**
Managers and Administration, except Farm	**5**
Not in the Workforce	**6**
Operative, except Transport	**7**
Private Household Workers	**8**
Professional, Technical, and Kindred Workers	**9**
Sales Workers	**10**
Service Workers, except Private Household	**11**
Transport Equipment Operatives	**12**

Wage Deciles Code

In the 1st Wage Decile	**1**
In the 2nd Wage Decile	**2**
In the 3rd Wage Decile	**3**
In the 4th Wage Decile	**4**
In the 5th Wage Decile	**5**
In the 6th Wage Decile	**6**
In the 7th Wage Decile	**7**
In the 8th Wage Decile	**8**
In the 9th Wage Decile	**9**
In the 10th Wage Decile	**10**
The Individual Has No Income	**No Income**
Missing Data	**NA**

Transition matrix of changing occupations for women from 1981 to 1991:

	1	2	3	4	5	6	7	8	9	10	11	12	Total
1	604	34	0	12	159	207	53	6	168	41	62	11	1357
2	6	12	0	4	7	13	8	0	1	0	21	0	72
3	6	5	0	3	2	10	11	0	0	5	7	6	55
4	6	13	0	2	14	22	19	0	12	14	5	7	114
5	117	54	1	12	99	351	30	12	79	11	149	14	929
6	605	74	0	7	123	992	232	50	309	93	427	32	2944
7	55	0	0	8	13	63	178	3	16	6	29	0	371
8	21	9	0	6	26	3	6	0	64	8	21	6	170
9	38	12	0	0	70	82	4	4	413	21	44	0	688
10	58	31	7	10	27	110	36	10	25	30	115	39	498
11	53	14	0	3	23	97	9	28	80	31	259	3	600
12	8	0	0	0	12	2	4	0	0	2	3	3	34
Total	1577	258	8	67	575	1952	590	113	1167	262	1142	121	7832

Transition matrix of changing occupations for men from 1981 to 1991:

	1	2	3	4	5	6	7	8	9	10	11	12	Total
1	161	26	0	4	78	35	7	0	51	25	12	7	406
2	96	876	7	62	260	112	132	0	97	52	62	40	1796
3	0	18	134	7	12	25	8	0	5	10	0	5	224
4	14	68	0	71	17	31	37	0	13	2	48	27	328
5	19	83	17	20	398	31	23	0	75	52	18	21	757
6	27	67	0	59	23	153	45	7	41	20	59	41	542
7	50	219	11	83	115	81	370	0	29	12	58	38	1066
8	0	0	0	0	0	0	0	0	0	0	0	7	7
9	20	44	4	12	246	21	25	0	718	25	20	10	1145
10	6	23	3	15	149	16	17	0	47	55	5	5	341
11	11	25	0	22	21	32	0	0	40	9	157	12	329
12	61	52	1	36	36	33	11	1	14	11	13	186	455
Total	465	1501	177	391	1355	570	675	8	1130	273	452	399	7396

Transition matrix of relative position change in wage deciles for women from 1981 to 1991:

	1	2	3	4	5	6	7	8	9	10	No Income	NA	Total
1	42	45	71	17	10	21	19	0	13	15	166	9	428
2	32	30	29	26	32	24	21	8	0	4	159	0	365
3	10	16	44	30	23	18	8	10	0	0	145	5	309
4	21	6	13	25	31	38	26	5	8	0	69	0	242
5	9	15	20	3	28	6	33	25	19	2	107	0	267
6	17	10	0	2	22	24	29	39	18	13	91	0	265
7	5	0	0	2	15	9	6	22	19	5	138	4	225
8	0	9	0	0	0	6	9	23	32	20	87	3	189
9	0	0	0	0	0	6	0	0	5	2	60	0	73
10	0	0	63	0	0	6	0	0	0	1	0	0	70
No Income	270	228	157	159	188	160	141	105	68	144	3669	56	5345
NA	0	0	0	0	0	14	4	0	0	0	33	3	54
Total	406	359	397	264	349	332	296	237	182	206	4724	80	7832

Transition matrix of relative position change in wage deciles for men from 1981 to 1991:

	1	2	3	4	5	6	7	8	9	10	No Income	NA	Total
1	12	8	11	7	2	1	1	0	0	5	48	13	108
2	9	12	3	7	19	9	14	0	0	2	44	12	131
3	2	8	6	5	5	11	13	12	5	0	56	0	123
4	7	3	11	24	5	43	12	3	0	10	109	0	227
5	5	12	13	32	13	24	43	10	17	5	148	4	326
6	1	16	12	18	25	45	21	59	4	16	142	0	359
7	14	8	4	8	8	22	62	55	98	25	194	0	498
8	6	12	9	4	17	20	14	105	150	49	188	0	574
9	0	16	6	9	25	7	19	33	93	96	161	4	469
10	4	8	15	5	0	15	30	15	47	236	233	6	614
No Income	33	64	55	88	66	81	99	86	143	162	3003	18	3898
NA	0	0	12	0	7	0	7	12	0	6	19	6	69
Total	93	167	157	207	192	278	335	390	557	612	4345	63	7396

Transition matrix of changing occupations for women from 1991 to 2001:

	1	2	3	4	5	6	7	8	9	10	11	12	Total
1	480	31	0	11	117	96	22	2	148	31	80	16	1034
2	28	24	0	4	20	24	14	0	12	11	8	0	145
3	0	0	0	0	0	0	0	0	0	0	0	0	0
4	15	2	4	8	0	0	0	0	0	0	3	0	32
5	41	6	0	4	85	28	18	0	54	21	15	5	277
6	255	64	6	21	135	361	103	5	176	42	272	40	1480
7	49	19	0	13	16	43	112	0	27	3	29	10	321
8	3	0	0	1	0	6	3	0	7	0	5	0	25
9	77	3	0	0	100	91	7	0	415	33	68	4	798
10	37	6	0	6	18	14	14	0	21	32	0	0	148
11	110	30	5	13	32	85	36	9	82	34	240	28	704
12	5	0	0	0	5	0	3	0	0	0	4	13	30
Total	1100	185	15	81	528	748	332	16	942	207	724	116	4994

Transition matrix of changing occupations for men from 1991 to 2001:

	1	2	3	4	5	6	7	8	9	10	11	12	Total
1	58	10	0	0	33	8	12	0	30	7	10	5	173
2	42	473	4	52	102	33	51	11	82	27	64	34	975
3	0	15	69	6	8	1	0	0	4	4	6	2	115
4	22	76	0	50	11	13	21	0	22	6	37	42	300
5	39	56	0	4	361	4	18	0	87	41	20	9	639
6	26	57	1	14	31	44	70	0	28	4	53	45	373
7	43	77	0	21	27	26	136	0	23	7	23	37	420
8	1	0	0	0	0	0	0	0	0	0	0	0	1
9	31	29	0	6	121	2	1	0	378	34	13	0	615
10	4	5	2	0	53	9	4	0	12	85	3	1	178
11	10	16	0	13	30	34	25	0	18	9	151	3	309
12	24	29	5	16	30	4	27	3	16	1	30	97	282
Total	300	843	81	182	807	178	365	14	700	225	410	275	4380

Transition matrix of relative position change in wage deciles for women from 1991 to 2001:

	1	2	3	4	5	6	7	8	9	10	No Income	NA	Total
1	54	64	37	46	22	16	0	0	5	0	47	0	291
2	39	47	45	26	37	35	4	3	8	0	53	0	297
3	33	69	26	34	18	18	29	9	14	3	23	1	277
4	16	14	22	25	37	10	8	5	11	4	25	0	177
5	11	21	28	52	37	31	9	14	4	4	36	5	252
6	5	12	2	45	21	18	36	8	13	4	22	0	186
7	8	5	17	6	23	30	41	20	8	9	14	0	181
8	5	6	18	10	9	17	8	38	30	0	21	0	162
9	5	14	6	5	4	19	30	25	16	6	19	0	149
10	5	0	0	0	0	0	10	33	22	16	20	0	106
No Income	265	285	244	237	194	235	200	273	216	168	580	6	2903
NA	0	0	0	0	8	0	0	0	0	0	5	0	13
Total	446	537	445	486	410	429	375	428	347	214	865	12	4994

Transition matrix of relative position change in wage deciles for men from 1991 to 2001:

	1	2	3	4	5	6	7	8	9	10	No Income	NA	Total
1	10	2	10	6	7	7	4	0	3	0	7	0	56
2	15	28	14	8	1	17	5	0	10	1	6	11	116
3	4	12	12	14	8	20	1	7	0	0	17	5	100
4	6	19	2	14	21	23	1	15	1	2	5	0	109
5	14	3	23	34	46	33	9	13	0	11	7	6	199
6	3	8	26	45	17	27	35	10	7	4	7	6	195
7	8	9	10	16	32	74	45	37	31	6	14	6	288
8	10	4	25	16	31	26	69	50	12	9	12	15	279
9	12	3	9	7	12	10	38	57	79	42	15	10	294
10	15	4	2	5	3	8	11	46	56	57	7	12	226
No Income	69	95	119	128	159	184	280	237	412	559	113	121	2476
NA	10	1	3	0	0	6	4	0	4	0	3	11	42
Total	176	188	255	293	337	435	502	472	615	691	213	203	4380

Reference:

1) *C. Goldin,* "**Understanding the Gender Gap: An Economic History of American Women**", *1990.*

2) Martha Hill "***The Panel Study of Income Dynamics: A User's Guide****", volume 2. 1991*

3) R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

4) JMP, Version 7. SAS Institute Inc., Cary, NC, 1989-2007

Money Resource Allocation To Children in PSID

Instructor: Professor Frank Stafford

Xuanzhong Wang

March 2009

1. Introduction

Money resource allocation to children as a combination of money spent by family and school can be considered a form of investment in human capital. The distribution of money resource allocation varies across nation. The disparity in money resource allocation is a possible factor that can explain the difference children's future achievement. In this project, the distribution of money resource allocation to children is studied; and in the second part of this project, possible factors that may affect the distribution of money resource allocation within family/school will be discussed.

Various models and theories have been developed to explain the inequality in terms of money resource allocated to children. In this project, the quality of child will be the primary factor of interest in predicting money resource allocating to children. A hypothetical model in the school context has been proposed in the 1970s, that if a school is elitist, it will be more likely to allocate more resources on students who are cognitively better endowed; but if a school prefers less variation in student's performance, more resources will be allocated to less prepared students in order to smooth out the performance disparity (Brown and Saks 1975). This model is equally relevant in a family context—parents usually wish to smooth out the achievement between children in the family rather than having one child stands out from all other siblings. Evidence had been found that within the family with a child having Down Syndrome (who was poorly endowed both physically and intellectually), the time devoted to the normal siblings is less than otherwise, as illustrated by time diary studies (Barnett 1993). In the second part of this project, the mechanisms behind the allocation of money resource will be tested.

2. Data & Method

2.1) Data Source

The data used for this project consist of three parts—the family level data used for the analysis were extracted from the Panel Study of Income Dynamics (PSID) core data and child level data were from Child Development Supplement wave II (CDS-II) interview in 2002; school level data were obtained from National Center for Educational Statistics (NCES) Common Core of Data (CCD). The PSID gathers information about families and all individuals in those families

through its annual interviews (bi-annual since 1997). The study's original households constitute a national probability sample of U.S. households as of 1967. The CCD School District Finance Survey consists of data submitted annually to NCES by state education agencies in the 50 states and the District of Columbia. In 2008, a sensitive data set that links the PSID CDS sample to the CCD was made available. In this project, the expenditure related variables in CDS II and per pupil expenditure in CCD were linked using the link file provided by the PSID.

2.2) PSID Sample

The initial sample for the PSID in 1968 consisted of two independent samples: a cross-sectional, national sample (based on stratified multistage selection of the civilian non-institutional population of the U.S.) and a national sample of low-income families. The cross-section sample was drawn by the Survey Research Center (SRC) of the University of Michigan and it is commonly called the SRC sample; it was an equal probability sample of households in the 48 coterminous states. The second sample of responding PSID families, known as the SEO sample, originated from the Survey of Economic Opportunity (SEO), conducted by the Bureau of the Census for the Office of Economic Opportunity. The PSID selected from the SEO's sample, the goal was to obtain about 2,000 low-income families with heads under 60 years old. The SEO sample was confined to Standard Metropolitan Statistical Areas (SMSAs) and to non-SMSAs in the Southern region, and it involves unequal selection probabilities. The PSID sample is a combination of the SRC and SEO samples, results a nationally representative sample of families in the United States with an oversample of low-income families.

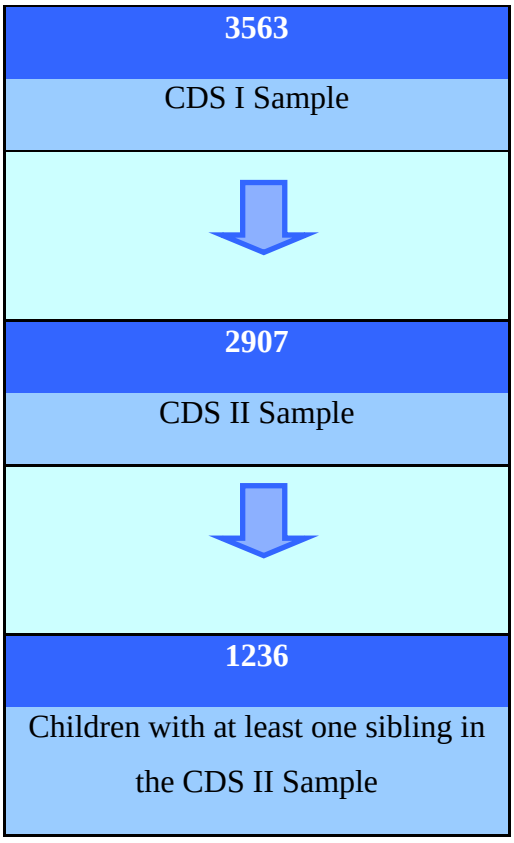
2.3) CDS Sample

In 1997, all PSID families with children aged 0-12 years were sampled for a Child Development Supplement (CDS). In the CDS study, up to 2 children per PSID family are chosen for the interview. 2394 families participated in the interview of CDS-I yielding a total sample of 3563 children. In 2002-2003, 2021 families were successfully re-interviewed, resulting in total of 2907 child interviews. The child level weight is re-calculated to take care of the re-classification of the eligibility and non-response problem in the wave II study of CDS. (PSID Website)

2.4) Analytic Sample

The analysis in this project will only make use of the CDS-II sample. In the first part of the analysis about the overall distribution of money resource allocation, all children in the CDS-II interview are included resulting a sample size of 2907. In the second part of the analysis where the focus is the pattern of in family allocation of money resource between multiple children, only observations with at least one sibling who was also included in the CDS-II were included. This results a sample size of 1236. By the design of the CDS, if there are more than two children in a CDS family, only two will be interviewed, the 1236 observations are in fact 618 pairs of siblings.

2.5) Analytic Sample Selection



2.6) Variables

Two variables were constructed from CDS-II interview, namely, the school related expenditure on child in 2002 and the total expenditure on child in 2002. In CDS-II, a series

questions were asked about the expenditure on various areas for the child being interviewed. The response to those questions serves as an estimate of amount of money resource allocated to child by family. This is an under-estimate of the actual amount of money resources allocated to children—for the majority of children in the United States attend free public schools. This downward bias can be corrected by including the amount of per pupil expenditure reported in the CCD School District Finance Survey for all public schools. Per pupil expenditure is the average amount of school expenditure on the entire student body.

Other variables such as sex and age of child are also included in the analysis as a possible factor that may have an effect in the money resource allocation. The Family Identification Mapping System (FIMS) is used to match the family level data to child. A summary of all variables employed in the analysis in this project is presented on page 5.

2.7) Statistical Analysis

When properly weighted, the PSID provides a national representative sample. The standard error estimation in this project does not take the complex survey design feature into consideration. First, descriptive statistics were calculated to explore the characteristics of the data—the mean of continuous variables; weighted frequencies for categorical variables were tabulated. Descriptive statistics of the data assuming a simple random sample—i.e. estimation without weight were also presented in order to capture the bias induced by ignoring the complex sample design.

Following the descriptive statistics, this project went further to contrast the mean expenditures between different groups of children—different geographical locations and different ability brackets. Based the contrast of the means, several regression models were tested. Indicator variables were heavily employed in the construction of regression model and will be further discussed in the result and discussion section. Another point to note is case-wise deletion was used in cases of missing data.

3. Results & Discussions

3.1) Descriptive Statistics

A table summarizing all variables is shown on page 6. It can be observed that assuming simple random sample brings bias in the estimation and deflates standard error estimation.

<table>
<tr><th colspan="2">Sampling Variables</th><th colspan="2">Resource Allocation Variables</th><th colspan="2">Explanatory Variables</th></tr>
<tr><td rowspan="2">Unique Child ID</td><td rowspan="6">PSID</td><td>School Related Expenditure (by Family) on Child in 2002</td><td rowspan="2">Calculated from PSID</td><td>Sex of Child</td><td>Nominal 0 for Female, 1 for Male</td></tr>
<tr><td>Total Expenditure (by Family) on Child in 2002</td><td>Age of Child</td><td>Ordinal</td></tr>
<tr><td rowspan="4">Child Level Weight</td><td>Per Pupil Expenditure (by School) in 2002</td><td rowspan="2">Common Core of Data</td><td>Total Number of Children Reside in the Family Unit</td><td rowspan="4">From PSID</td></tr>
<tr><td rowspan="2">Total Money Resources Allocated to Children in 2002</td><td>WJR Score 2002</td></tr>
<tr><td rowspan="2">Combining Expenditure by Family and Per Pupil Expenditure by School</td><td>Family Income in the Previous Year</td></tr>
<tr><td>Variation in Family Income in the Past 15 Years</td></tr>
<tr><td colspan="2"></td><td></td><td></td><td>Parents Expectation on Child</td><td>Ordinal 1 Lowest, 8 Highest</td></tr>
<tr><td colspan="2"></td><td></td><td></td><td>Type of School Attended</td><td>Nominal
0 Not in School
1 Public School
2 Private</td></tr>
</table>

					School 3 Home School
				State in which the Child Lives	FIPS State Code
				Beale Code of where the Child Lives	Ordinal 1 Urban with Population >= 1 Million 10 Completely Rural

Summary of all Variables Used in the Analysis

Descriptive Statistics of Variables Used in the Analysis

	Sample Size Used in Estimation	Assume Simple Random Sample	Weight Adjusted Estimates
		Estimated Mean	
		(s.e.)	
Family Income in the Previous Year	2907	758.27	773.0457
		(8.416527)	*(11.36436)*
Variation in Family Income	2907	1004.263	1015.574
		(10.77199)	*(13.64655)*
Total Expenditure (by Family)	2907	3499.008	3805.428
		(56.63853)	*(76.07048)*
School Related Expenditure (by Family)	2907	430.6966	501.8507
		(22.61445)	*(28.59978)*
Per Pupil Expenditure (by School)	2101[i]	8935.227	8995.448
		(51.44929)	*(72.10581)*
WJR Score 2002	2907	101.2883	91.70461
		(1.779688)	*(2.206258)*

Sex		**Male**				**Female**			
	w/o Weight	50.64%				49.36%			
	Weight Adjusted	50.35%				49.65%			
Age Group		**Below 6**		**Grade School**		**Junior High School**		**High School or Above**	
	w/o Weight	6.19%		55.66%		14.55%		23.60%	
	Weight Adjusted	6.71%		53.88%		15.97%		23.44%	
Parents Expectation		**1 (Lowest)**	**2**	**3**	**4**	**5**	**6**	**7**	**8 (Highest)**
	w/o Weight	1.24%	22.91%	2.10%	4.71%	9.77%	46.51%	8.02%	4.13%
	Weight Adjusted	1.00%	19.22%	2.17%	4.42%	9.87%	48.39%	8.74%	5.30%
Total Number of Children in Family Unit		**Missing**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	w/o Weight	1.20%	17.03%	44.31%	26.42%	7.33%	2.13%	1.24%	0.14%	0.07%	0.14%
	Weight Adjusted	1.32%	15.65%	42.21%	27.56%	8.11%	3.13%	1.41%	0.44%	0.04%	0.12%
Type of School Attended		**Not in School**		**Public School**		**Private School**		**Home School**	
	w/o Weight	4.61%		87.34%		6.57%		1.41%	
	Weight Adjusted	5.14%		85.07%		7.97%		1.78%	
Beale Code		**1 (Urban)**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 (Rural)**
	w/o Weight	28.17%	16.61%	25.50%	5.51%	3.62%	3.37%	5.87%	8.64%	1.23%	1.48%
	Weight Adjusted	22.28%	17.28%	24.17%	6.56%	6.82%	4.08%	6.53%	8.75%	1.72%	1.81%

From the summary of money resource allocation variables: i.e. total expenditure by family, school related expenditure by family and per pupil expenditure by school, it can be seen that expenditure by family only constitutes about one third of total money resource allocated to children. It can also be seen that school related expenditure is a minor component of total money resource allocated to children by family. Most of families' expenditure on child goes to toys, clothes, medical care, food and etc.

Figure 1 shows the distribution of mean money resource allocated to children in different areas—from urban areas with huge population to completely rural areas. Still, it can be observed that per pupil expenditure is the most significant component in terms of money resource allocation to children. It also can be seen that children reside in the more urban areas (with Beale code equal 2), as well as children in the most rural areas (Beale code 9 or 10) enjoyed more money resources. Besides children reside in most rural and urban areas, money resource allocated to children decreases as children moves from urban to rural areas.

Figure 1 Components of Money Resource Allocated to Children (Mean)

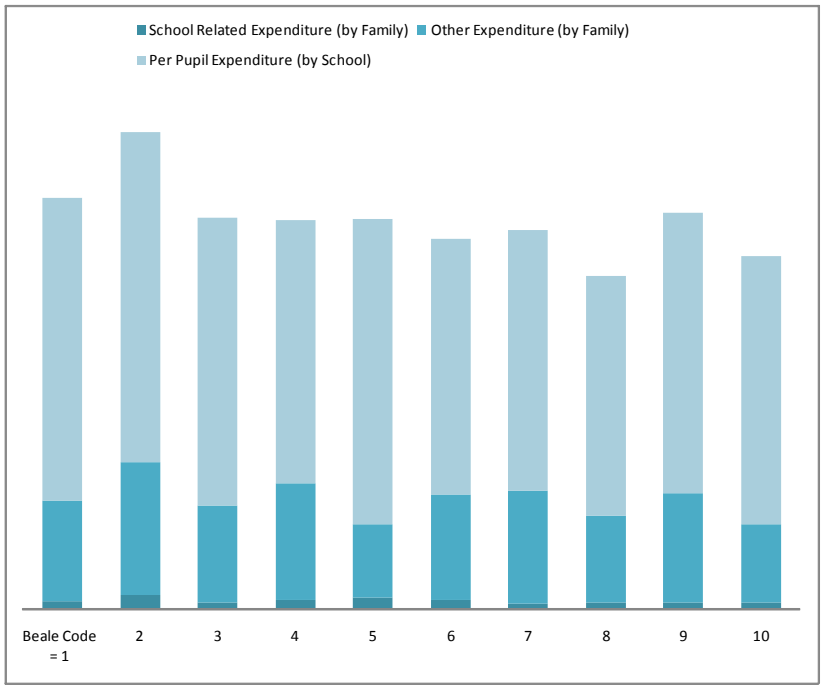
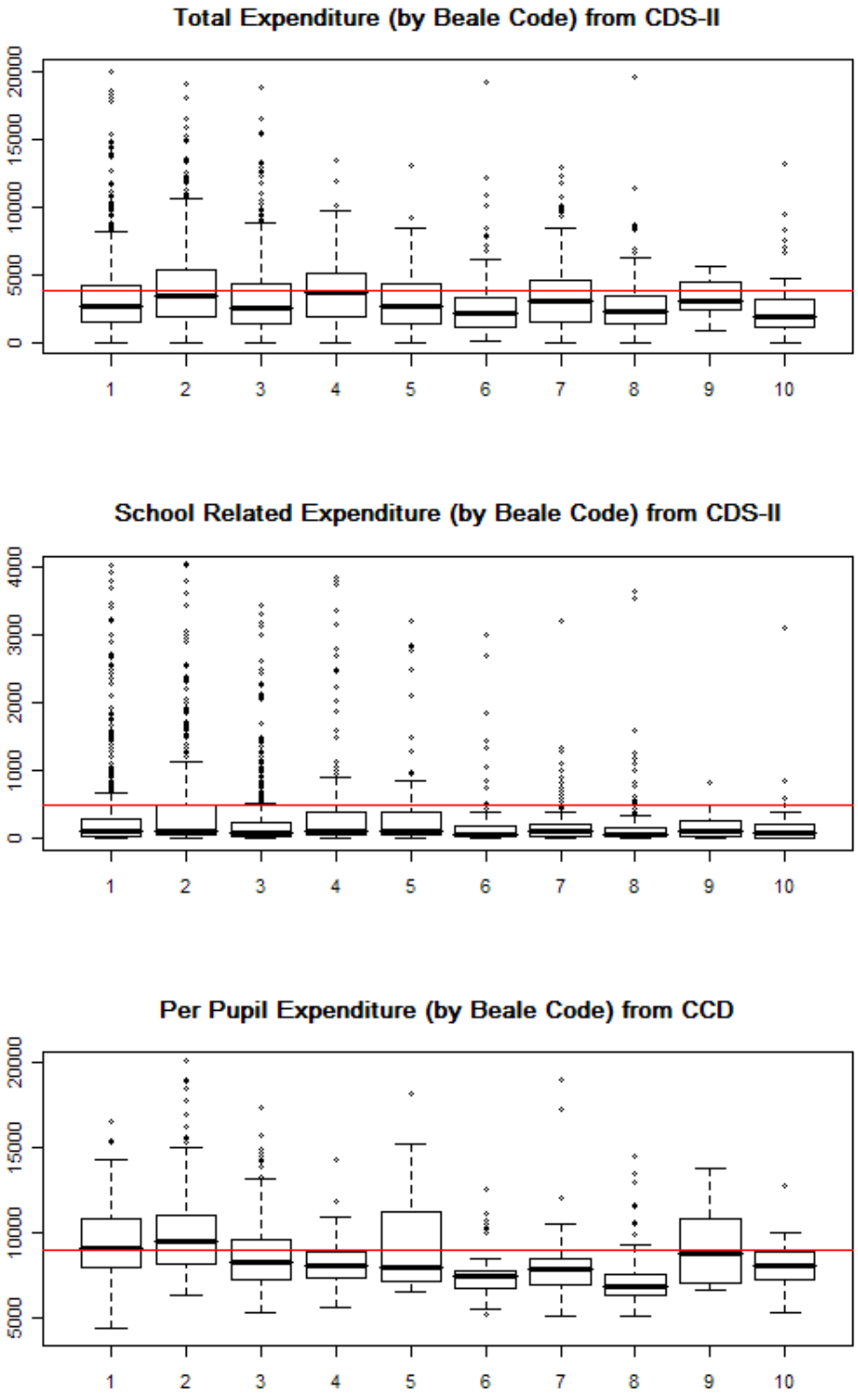


Figure 2 shows three box-and-whisker plots of the money resource allocation variables in order to compare the distribution of money resource allocation cross urban and rural areas. The national average of each variable is marked with the horizontal red line.

Figure 2 Box-and-Whisker Plot of Money Resource Allocation vs. Beale Code



Several features can be observed from the Figure 2. First of all, there is huge amount of variation in money resource allocated to children by family; this can be observed from the large number of outliers in the first two panels in Figure 2. Also, it can be observed that there is more variation in money resource allocation between children living in urban areas than those living in rural areas. This indicates that there are many more money resource allocation options available in more urban areas that lead to a diversified portfolio in terms of expenditure on child—for example, more choices of schools, bookshops and etc.

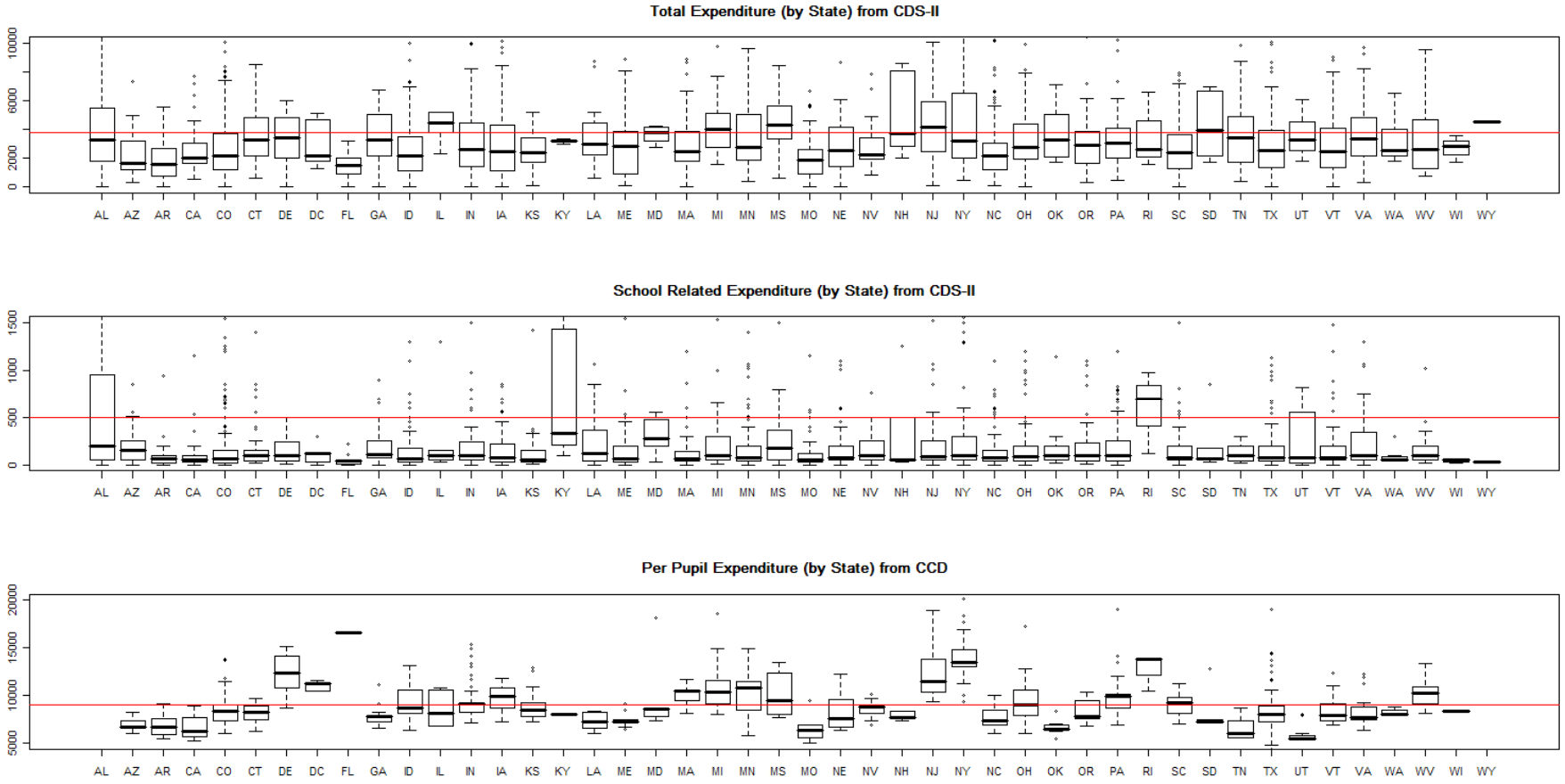
However, there is no obvious pattern in amount of money resource allocation with respect to geographical locations—the median of total expenditure by family, school related expenditure by family remained at similar level throughout the entire range of Beale Code. It can be seen that the national averages of total expenditure on child by family, and school related expenditure by family are well above the median of each distribution. Combing the two features, it can be concluded that the national averages of total expenditure, as well as school related expenditure on children by family are driven up by the outliers presented in population—they are either parents who spend a huge amount of money on their children or parents who over-stated the amount of money spent on children during the interview. Apart from a small proportion of families who invest heavily in children's daily life and education, there are not many disparities in term of money resource allocation to children by family.

In the last panel of Figure 2, the box-and-whisker plot of per pupil expenditure by school against Beale Code is presented. It can be observed that there is less variation in per pupil expenditure compare to the distribution of expenditure by family. From Beale Code 1 through 8, a decreasing trend in per-pupil expenditure is observed—this makes sense as more urban regions and larger cities usually have higher price levels than that of rural regions. However, for the two categories labeled as most rural areas—Beale Code 9 and 10, the mean per pupil expenditure goes up. This can possibly be attributed to the fact that most rural areas have smaller population, i.e. in this case, smaller number of children in public schools. When per pupil expenditure is calculated, despite the relatively lower price level, when dividing by a small population size, the average value is likely to be higher than regions with larger population and similar price level.

Figure 3 is similar to Figure 2 but instead of comparing the distribution across different Beale Codes, Figure 3 compares the money resource allocation distribution across states. From the first panel of Figure 3, it can be concluded that there are more variations in total expenditure

by family across states than the variation between urban-rural areas. However, in terms of school related expenditure, the distribution remained relatively consistent across states. This can be explained as total expenditure is closely related to the different price levels, but for most (around 85%) families who sent their children to public schools, school related expenditure is a more constant component of expenditure on children.

Figure 3 Box-and-Whisker Plot of Money Resource Allocation vs. State



In the third panel of Figure 3, a substantial amount of variability of per pupil expenditure can be observed between states. States like Delaware, D.C., New Jersey, New York and etc. have per pupil expenditure distribution completely above the national average; at the same time, and states like Arizona, Arkansas, California and etc. have per pupil expenditure distribution completely below the national average. The disparity can be explained as a combination of difference in educational input by state government between different states, as well as the difference between price levels across states. In states with high price level, the wage level of public school teachers is likely to be higher than states with low price level.

One possible way of accommodating the different price levels across states is the introduction of a variable of student teacher ratio in monetary units. Teacher's wage constitutes the most significant proportion of cost in most schools. The average annual wage of teachers in different states is provided by the Bureau of Labor Statistics, student teacher ratio is calculated using the following formula:

$$\frac{\text{Average Annual Wage of Teacher in 2002 (BLS)}}{\text{Per Pupil Expenditure in 2002 (CCD)}}$$

The student teacher ratio calculated as shown above can be interpreted as number of students assigned to one teacher in terms of monetary cost. The larger the number, the less the actual amount of teaching resource allocated to children. The distributions of student teacher ratio in monetary units with respect to urban-rural areas, and states are presented in Figure 4 and Figure 5.

In Figure 4, the variation in student teacher ratio decreases as the region goes from urban to rural. A more consistent distribution of student teacher ratio can be observed in Figure 4, indicating some of the variation in per pupil expenditure by school can be attributed to the difference in difference price levels. On another hand, it also shows that there is disparity in terms of real teaching resource allocation between and within urban and rural areas.

From Figure 5, it can be observed that some of the variation in per pupil expenditure are removed in student teacher ratio. This indicates that price level—i.e. teacher's wage level is a significance factor that may affect the amount of money resource allocated to children. The variation existing in Figure 5 indicates that there is still some difference in real teaching resource allocation between states.

Figure 4 Box-and-Whisker Plots of Student Teacher Ration vs. Beale Code

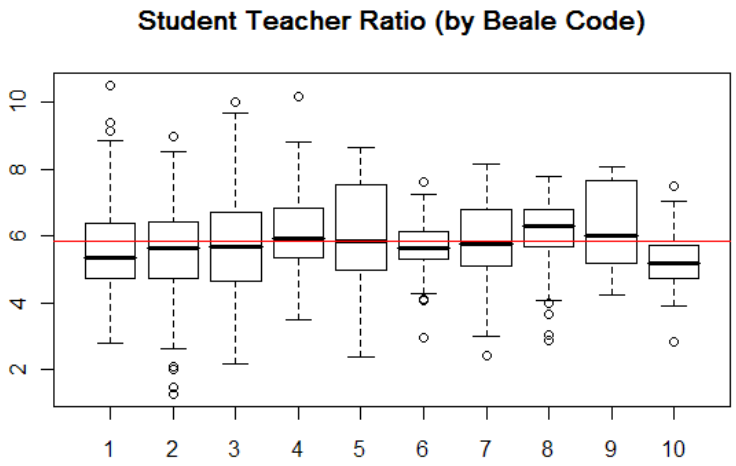
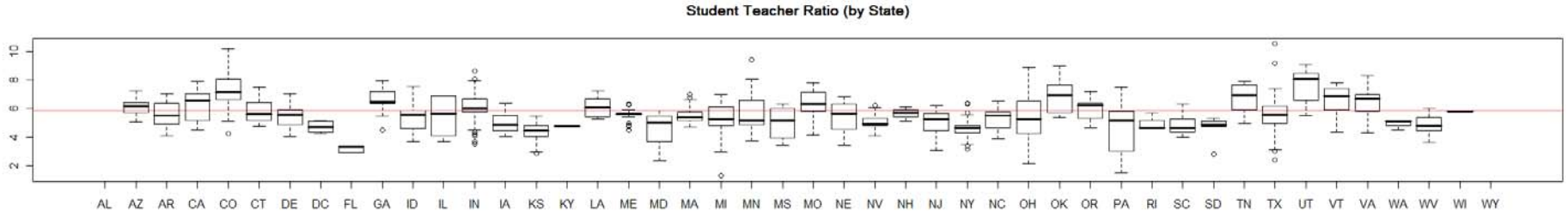


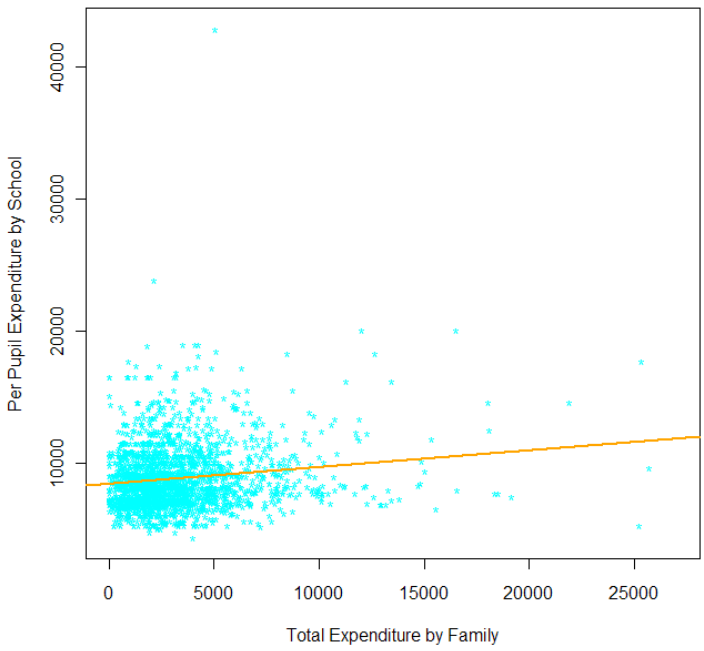
Figure 5 Box-and-Whisker Plots of Student Teacher Ration vs. State



One problem with the calculated student teacher ratio is the variable is calculated based on only the state level average annual wage for teachers. In reality, wage level differs within the same state. Bias will set in if the state average is used in place of the actual wage level.

Besides the descriptive statistics, it is also of great interest to find out the relationship between money resources allocated to children by family and by school. Figure 6 shows a scatter plot of per pupil expenditure against total expenditure by family.

Figure 6 Scatter Plot of Per Pupil Expenditure vs. Total Expenditure



The fitted line of the linear model with per pupil expenditure as response and total expenditure as the only predictor is plotted in orange. The positive slope of this line indicates a positive correlation between total expenditure and per pupil expenditure.

cor(Total Expenditure, Per Pupil Expenditure)=0.139

The total expenditure by family and the per pupil expenditure by school is then broken into 10 deciles and are tabulated as shown in the next page.

		Deciles in Total Expenditures by Family										
		1st	**2nd**	**3rd**	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**10th**	**Total**
Deciles in Per Pupil Expenditure	**1st**	0.44	1.01	1.03	1.39	1.01	0.83	1.32	1.36	0.93	0.57	**9.89**
	2nd	0.84	1.00	0.97	0.81	1.15	0.50	0.63	0.70	1.20	0.51	**8.30**
	3rd	2.21	1.63	1.20	1.10	0.50	1.29	0.92	1.41	0.60	0.73	**11.58**
	4th	1.11	0.37	0.67	1.17	1.44	1.31	2.11	1.40	1.53	1.45	**12.55**
	5th	0.99	0.52	1.33	0.94	1.11	0.92	0.79	0.73	0.95	1.12	**9.40**
	6th	1.54	1.75	1.30	0.64	0.84	1.30	0.77	0.79	0.66	0.67	**10.26**
	7th	0.70	0.77	0.82	0.50	1.32	1.20	0.90	0.71	0.92	0.74	**8.56**
	8th	0.81	0.59	0.46	0.83	0.59	1.25	0.91	1.12	1.09	0.94	**8.59**
	9th	0.85	0.72	0.86	0.74	0.61	0.51	1.01	1.26	0.98	0.60	**8.15**
	10th	0.50	0.95	0.70	1.39	1.09	1.30	1.16	2.07	1.48	2.10	**12.73**
	Total	**9.99**	**9.30**	**9.34**	**9.49**	**9.65**	**10.40**	**10.52**	**11.54**	**10.33**	**9.43**	**100.00**

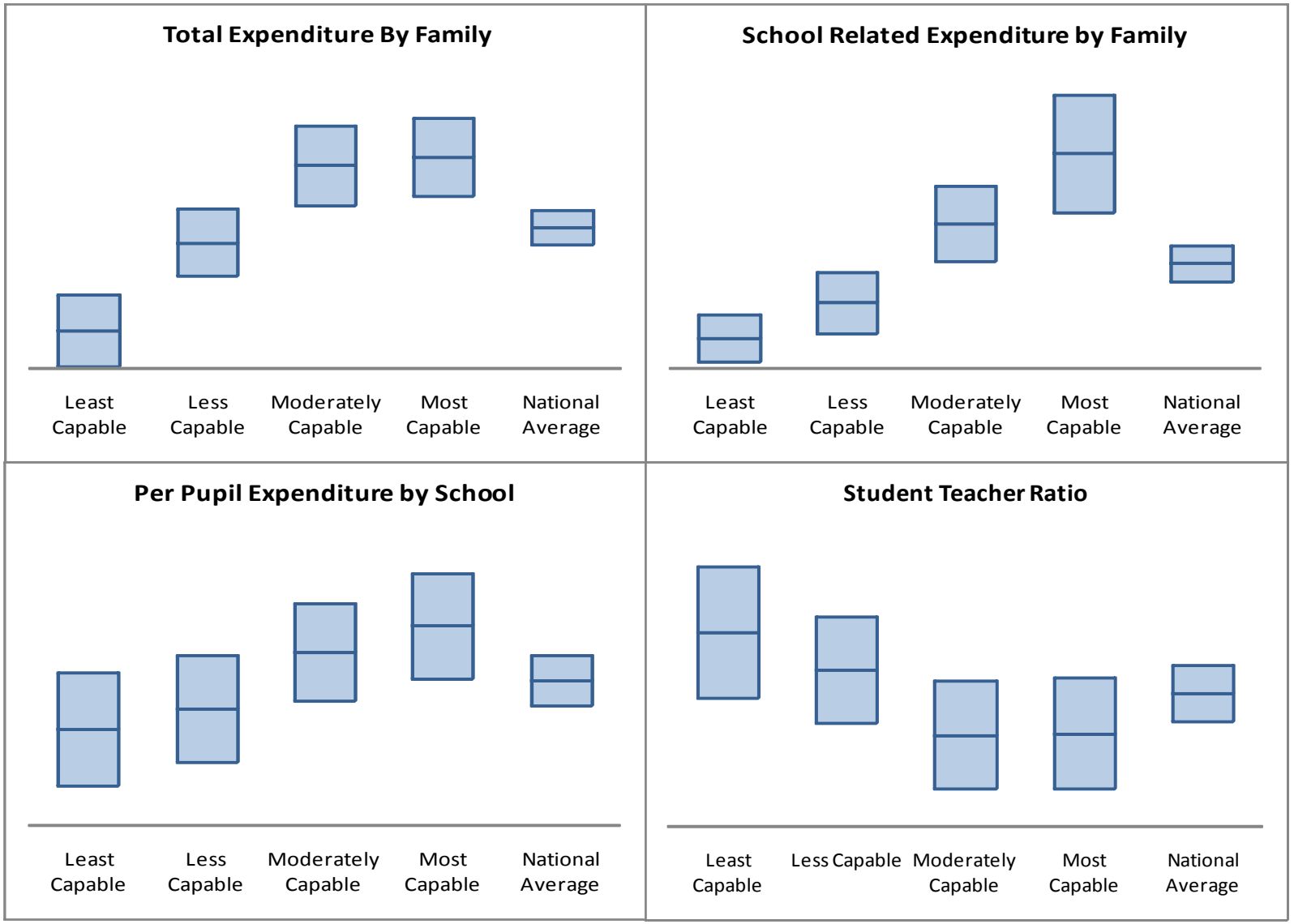
In the above table, each cell contains the weighted cell proportions of corresponding total expenditure decile and per pupil expenditure decile (with 1 being the lowest and 10 being the highest). From this table, only weak correlation can be observed—which goes in line with the flat slope in Figure 6.

3.2) Bivariate Analysis

There are many factors that may possibly affect the money resource allocation decision. As mentioned in the introduction, various models have been developed in order to capture the effect of such factors. Children's ability has been mentioned as a possible factor that might affect parents' money resource allocation decision. In the PSID CCD, the Woodcock Johnson Standardized Test (WJR) was administrated for every able aged respondent. The score from the WJR test was recorded, and in this case is used as an estimate of children's ability.

In the next part of the analysis, children are categorized into four ability brackets—least capable bracket corresponding to the children whose WJR score falls into the lowest 25% in the overall weighted WJR score distribution and etc. The mean total expenditure, mean school related expenditure, mean per pupil expenditure by school and teacher student ratio of children in each category were estimated. The 95% Confidence intervals of estimated mean for each category was calculated as well. Figure 7 shows the plotted 95% confidence intervals.

Figure 7 95% CI of Means



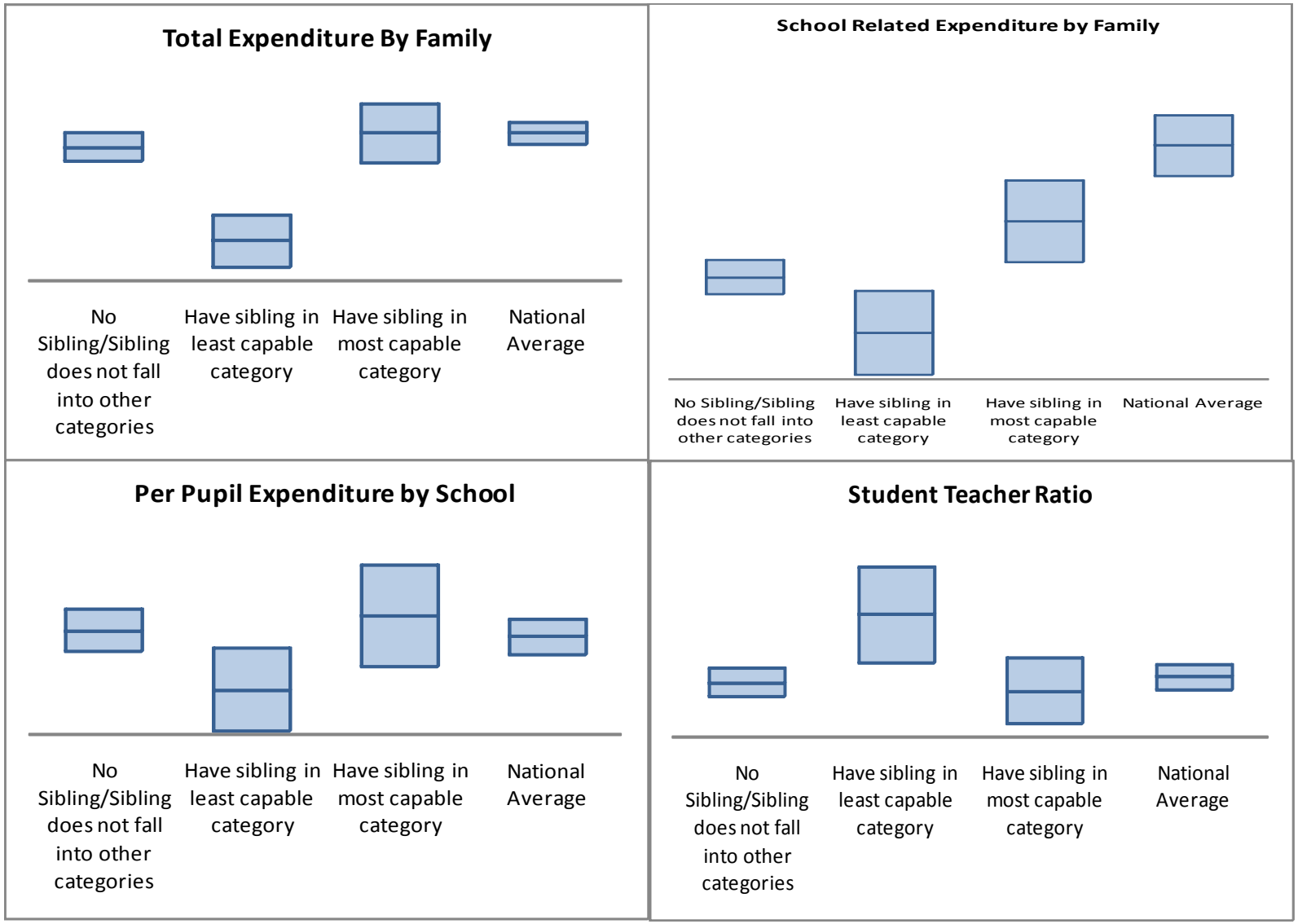
It can be observed from Figure 7 that there is an obvious trend in money resource allocation to children when the child moves up the ability brackets. In general, the money resource allocated to children is proportional to the ability that a child demonstrates in the WJR score. This is especially true in terms of money resource allocation by family—more capable children has significantly higher mean total expenditure, as well as school related expenditure comparing to their less capable counterparts.

Similar trends exist in per pupil expenditure by school and student teacher ratio as well. However, the trend is not as significant as that for in family money resource allocation. One possible explanation for that is policy and regulations ensure that each child receives a basic set of training; therefore the disparity in resource allocation taken care by school is not as significant as that between families.

Another factor that might have an effect on money resource allocation to children is the quality of siblings in the same family unit. As mentioned in the introduction, if there is multiple numbers of children in the family units, the decision to allocate money resource to children depends on whether parents have a preference of having more balanced development between children. Index variables indicating individual child have “no sibling or sibling not in any specified categories”, “one sibling in the least capable category” and “one sibling in the most capable category”. The money resource allocation variables are contrasted between the three groups. 95% CI’s of the means are constructed and plotted in Figure 8.

From Figure 8, it can be seen that children with a sibling in the least capable category are obviously at a disadvantage comparing to children in other categories—they have lower mean total expenditure, school related expenditure from family; in school, they have lower per pupil expenditure and higher student teacher ratio as well. This might be the outcome when parents try to “smooth out” children’s future performance by investing more heavily on less capable children. This is equivalent to the case where children with a Down ’s syndrome sibling enjoyed much less time devoted to them by parents. In the later section of regression analysis, it will examine the factor quality of sibling further.

Figure 8 95% CI of Means



3.3) Regression Analysis

First of all, factors that may affect the in family money resource allocation are examined. Four regression models are fitted for each response variable—school related expenditures and total expenditures. Ages of children are broken into four age categories and are coded as dummy variables in Model 1 and Model 5; similarly, ability brackets are also coded as dummy variables and used in the formation of Model 1, 2, 5 and 6. Types of school attended were also coded as dummy variable and Children who went to Public Schools were chosen as the reference category because about 85% of the children attend public schools. Sex of child is 0 if female and 1 if male, the females were chosen as the reference category. In Model 4 and Model 8, sibling's WJR score was also taken into the regression model results a much smaller sample size of 1026—out of the 1236 children in CDS-II whose sibling was also interviewed for CDS-II.

A summary of the regression models are shown in page 17 and 18. It is worth noting that all models are considered fairly fitted in the social science context. Approximately 43% of all variations in school related expenditures can be explained by the predictors selects, and more than 25% of the variation in total expenditure can be explained by the same set of predictors.

Regress School Related Expenditures on the Following Variables

Variables		Model 1	Model 2	Model 3	Model 4
		Coefficient *(Standard Error)*			
			(n=2225)		(n=1026)
Age of Child			17.26212 ~ *(9.554702)*	15.19986 *(9.358616)*	8.725293 *(10.72629)*
	Below 6	--	--	--	--
	Grade School	-47.6159 *(174.367)*			
	Junior High School	48.11075 *(191.538)*			
	High School and/or above	103.0268 *(212.1028)*			
WJR Score in 2002				5.959451 ** *(1.76801)*	1.988222 *(2.204387)*
	WJR Score 1st Quartile	--	--	--	--
	WJR Score 2nd Quartile	64.61093 *(53.96441)*	63.64477 *(55.44822)*		
	WJR Score 3rd Quartile	117.0763 (83.1511)	118.6787 (82.46895)		
	WJR Score 4th Quartile	259.918 *** *(79.79539)*	262.6258 *** *(83.53147)*		
Siblings' WJR Score					4.844903 ~ *(2.522347)*
School Type	*Not in School*	32.39679 *(63.24131)*	30.49039 *(62.34393)*	38.51924 *(60.37596)*	156.7968 ~ *(92.3291)*
	Public School	--	--	--	--
	Private	2675.889 *** *(256.565)*	2675.361 *** *(255.2839)*	2674.428 *** *(253.0671)*	2495.956 *** *(350.0966)*
	Home School	344.3385 *(207.1887)*	364.6684 *(207.2035)*	360.9707 ~ *(209.006)*	564.8295 * *(269.3646)*
Sex	*Female*	--	--	--	--
	Male	-183.2952 *** *(44.60671)*	-186.8006 *** *(44.67856)*	-183.9763 *** *(43.85902)*	-139.5635 * *(58.50651)*
Parents Expectation		43.59667 *** *(12.6192)*	44.00304 *** *(13.00129)*	43.11781 ** *(13.84302)*	49.24179 ** *(14.36877)*
Total Number of Child in the Family Unit		-40.69459 *(25.78938)*	-42.42262 *(25.40613)*	-41.38627 ~ *(24.21397)*	-54.26597 *(32.72046)*
Family Income in 2002		-0.0000453 *(0.0016135)*	-0.00067 *(0.0016133)*	-0.000065 *(0.0016246)*	0.0003644 *(0.0011999)*
Variation in Family Income		0.0063866 *(0.0056582)*	0.0064533 *(0.0056449)*	0.0064301 *(0.0056555)*	0.0033919 ~ *(0.0019642)*
Constant		-26.81352 *(223.2439)*	-212.9463 *(167.3623)*	-706.6435 ** *(264.2616)*	-685.2139 ~ *(350.591)*
R- Square		0.4312	0.431	0.4306	0.469
Significance Code		<0.001 ***	<0.01 **	<0.05 *	<0.10 ~

Regress Total Expenditures on the Following Variables

Variables		Model 5	Model 6	Model 7	Model 8
			Coefficient		
			(Standard Error)		
			(n=2225)		(n=1026)
Age of Child			145.8869 ***	141.1892 ***	148.6898 ***
			(26.42357)	*(25.25051)*	*(28.85262)*
	Below 6	--	--	--	--
	Grade School	295.6888			
		(380.9621)			
	Junior High School	1073.865 **			
		(415.9627)			
	High School and/or above	1530.279 ***			
		(447.1719)			
WJR Score in 2002				5.312196	-0.559727
				(6.218285)	*(5.830592)*
	WJR Score 1st Quartile	--	--	--	--
	WJR Score 2nd Quartile	212.9142	225.3122		
		(218.8742)	*(222.8312)*		
	WJR Score 3rd Quartile	420.8271	439.4327		
		(276.3415)	*(282.4165)*		
	WJR Score 4th Quartile	340.912	380.1206		
		(297.2214)	*(309.1549)*		
Siblings' WJR Score					5.55262
					(4.873667)
School Type	*Not in School*	274.9067	246.4188	259.108	192.5449
		(442.4889)	*(422.0927)*	*(410.0803)*	*(490.5923)*
	Public School	--	--	--	--
	Private	3145.401 ***	3128.813 ***	3130.163 ***	2506.998 ***
		(481.4013)	*(480.4861)*	*(485.5233)*	*(581.4391)*
	Home School	-260.8708	-267.4019	-270.7169	143.9722
		(541.1088)	*(546.7067)*	*(541.1279)*	*(580.3902)*
Sex	*Female*	--	--	--	--
	Male	-161.8117	-185.5202	-189.0184	-86.54715
		(158.3938)	*(157.3537)*	*(156.8966)*	*(188.0362)*
Parents Expectation		177.9035 **	180.6454 ***	196.423 ***	197.1875 **
		(65.24019)	*(65.42784)*	*(61.41707)*	*(51.88842)*
Total Number of Child in the Family Unit		-414.7028 **	-432.3491 ***	-442.8258 ***	-381.1845 ***
		(80.14935)	*(80.31523)*	*(79.06617)*	*(70.47456)*
Family Income in 2002		0.0060927 ~	0.0059364 ~	0.006046 ~	0.0050333
		(0.003027)	*(0.0030025)*	*(0.0030101)*	*(0.0031831)*
Variation in Family Income		0.0131303 ~	0.0134189 ~	0.0134393~	0.0218737 **
		(0.0073088)	*(0.0072562)*	*(0.0072378)*	*(0.0039342)*
Constant		2155.091 ***	1236.726 **	934.8108	756.263
		(543.5116)	*(493.4482)*	*(768.0117)*	*(916.4733)*
R- Square		0.2714	0.269	0.2674	0.2812
Significance Code		<0.001 ***	<0.01 **	<0.05 *	<0.10 ~

Here are some interesting results observed from the regression models:

Result 1

Comparing Model 3 and Model 7, WJR score has a significant effect on the school related expenditure but not the total expenditure. It can be inferred from Model 3 that every 1 unit increase in the standardized WJR score will raise the expected value of the school related expenditure by almost 6 dollars with a standard error of 1.77 dollars. However, a 1 unit increase in standardized WJR score has no significant effect on the expected total expenditure as shown in Model 7.

Result 2

In Model 1, 2 and 3 where school related expenditures were used as the response variable, it can be concluded that holding all others constant, a boy has expected school related expenditure that is about 180 dollars than that for a girl. Similar relationship cannot be observed from Model 5, 6, or 7. This can be put as holding all others equal; girls have a much higher mean in school related expenditure but not necessarily in total expenditure.

Result 3

Comparing Model 1, 2, 3 with Model 5, 6, 7; It can be concluded that with every one more child in the family unit, the expected value of school related expenditure will not decrease significantly; however, the expected total expenditure on one child will decrease by more than 400 dollars as shown in Model 5, 6, 7. That is to say, it is the contrary to common belief that more children will bring down the amount of money invested in one child.

Result 4

From Model 6, and Model 7, every one year increase in age will raise the expected total expenditure by more than 140 dollars; however, from Model 2 and Model 3, a one year increase in age has no significant effect in the expected school related expenditure.

Result 5

When both the child's WJR score and the sibling's WJR score are taken into consideration, a unit increase in WJR score has no significant effect in neither school related expenditure nor total expenditure. However, there is some effect (significant at 0.10 level) in

school related expenditure by the sibling's WJR score—meaning that a unit rise in sibling's WJR score, i.e. ability will raise the child's expected school related expenditure by about 4.8 dollars at a significance level of 0.10.

Result 1 concluded on page 22 that unit increase in WJR score will raise the expected school related expenditure but not the expected total expenditure. This goes in line with the bivariate analysis result that the means of school related expenditures for more capable children are significantly higher than that of not so capable children; but the pattern does not hold for total expenditure. The result shows that there is a strong association between child's WJR score-ability and the amount of school related expenditure, but no significant association with total expenditure. One possible reason is that total expenditure is less elastic than that of school related expenditure. As mentioned in the Data section, total expenditure was constructed by sum up school related expenditure and expenditures on food, clothes, and medical care, dental care and etc. Food, clothes, medical care and dental care are necessities in life and they are not sensitive to the influence of other factors. Thus, total expenditure is a combination of school related expenditure and a more stable component, and the component tends to reduce the effect by other factors in the analysis.

A more stable total expenditure also explains result 3 and 4 that increase in total number of children will reduce the amount of total expenditure, and rising age of children will raise total expenditure. It is hard to conclude a causal relationship between WJR score and school related expenditure either way. The association can be explained by heavy investment in child's intellectual development—through high school related expenditure "produces" high WJR score; or the other way that parents invest more in more capable children in the hope of higher returns of the investment—more capable children can make more from school related expenditures then the less capable one. Further studies are needed to figure the direction of the causation. One possibility is to make use of the panel feature of the PSID CDS data and to follow the children's ability development and the intellectual investments. This analysis might require more waves of CDS data to realize, but since the CDS has only two waves data that are available to the public, other longitudinal child surveys can be used.

However, result 2 and result 5 both reflects parents' intention of having a more equal development among children. Result 2 shows that holding all others constant; girls have an expected school related expenditure that is about 180 dollars higher than that of a boy. This

result can be viewed as parents' intention to compensate girls' disadvantage in employment by investing more intensively in girls' education.

Result 5 shows a child's expected school related expenditure will increase if his/her sibling has high WJR score. This can be argued as evidence that parents tend to allocate more money resource on the less capable child in order to achieve a more balanced development between multiple children at home. The most obvious problem with Model 4 and Model 8, is that within a family unit, siblings tend to have strongly correlated ability—both due to the similar genetic building and same environment that they exposed to.

Another short coming of the models formed is that it did not take into consideration about parents' fertility decision—given more informed and better educated parents today, they might weigh the choice of having more children and allocate money resource according to their ability, or allocate more money resource to the less capable ones in order to achieve a balanced development; or they would rather choose to have less children.

Besides factors that may affect the amount of money resource allocated to children within family, it is also of great interest to examine the factors that may affect the amount of money resource allocated to children by school, a.k.a. per pupil expenditure. However, as mentioned before, a significant amount of variation in money resource allocation by school comes from the difference in price level across the states. Thus, student teacher ratio, as an estimator of real teaching resource allocated to children will be used as the response variable. However, none of the predictors appear to be significant in the new set of models estimated. The regression models estimated gives poor fitting of R-squared value that is less than 0.02. This implies that the factors having a significant effect on money resource allocated to children in family do not affect the overall amount of money resource allocation significantly.

4. Conclusion

Based on the analysis in this project, disparity in money resource allocated to children exists across the United States. The in family money resource allocation depends on the several demographic factors—age and sex of child, the total number of siblings in the family unit; another important factor is the quality of child. Children with higher ability are likely to get more money spent on them than the others. On another hand, the quality of other children in

the family unit has a significant effect on money resource allocation, too. Children with less capable siblings are usually disadvantaged as more resources are directed to the poorly endowed siblings—this is because parents usually prefer a more balanced development of children rather than having one child stands out from all others.

School also allocates resources to children in the form of classes, teachers' attention and etc. The CCD data provides a way to quantify the money resource allocation as per pupil expenditure on child. Despite the great amount of discrepancies in distribution of per pupil expenditure, a large amount of variation came from the difference in price level—high price level drives up teachers' wage and in turn boost the per pupil expenditure. When convert to student teacher ratio as an estimate of real teaching resource allocated to children, the variations are unexplained by the set of predictors that fits the family wise money resource allocation well.

Although there is clear evidence of inequity in the distribution of money resource allocation to children, the effect of money resource allocation will take years to surface. The PSID being a longitudinal study has the advantage that can follow the same group of respondents for many years. Currently the oldest respondent of CDS should be about 22 years old. With more waves of interviews carried out, it is possible to track both their academic achievement and career attainment. Then the effect of money resource allocation to children can be fully revealed.

5. Reference

1) Byron W. Brown & Daniel H Saks (1975) "***The Production and Distribution of Cognitive Skills within Schools***", The Journal of Political Economy, Vol. 83 No. 3, (Jun., 1975), pp. 571-593

2) George E. Johnson and Frank P. Stafford "***On the Rate of Return to Schooling Quality***" *The*
Review of Economics and Statistics, Vol. 78, No. 4 (Nov., 1996), pp. 686-691

3) Barnett, W. Steven, ***"New Wine in Old Bottles: Increasing the Coherence of Early Childhood Care and Education Policy,"*** Early Childhood Re- search Quarterly 8 (1993), 519-588.

4) Martha Hill "***The Panel Study of Income Dynamics: A User's Guide***", *volume 2. 1991*

5) Berry, Chip and Cohen, Crecilla. U.S. Department of Education, National Center for Education Statistics. ***Documentation for the NCES Common Core of Data, School District Finance Survey*** *(F-33), School Year 1997–98, Fiscal Year (FY) 1998,* NCES 2003-336R. Washington DC: 2004.

6) Berry, Chip and Cohen, Crecilla. U.S. Department of Education, National Center for Education Statistics. ***Documentation for the NCES Common Core of Data, School District Finance Survey*** *(F-33), School Year 2001–02, Fiscal Year (FY) 2002,* NCES 2005–360. Washington DC: 2004.

7) Bureau of Labor Statistics, ***U.S. Department of Labor, Occupational Outlook Handbook***, 2008-09 Edition, [date accessed] [http://www.bls.gov/oco/]."

8) R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

[i] Missing value due to 1) child in non-public schools whose information was not recorded in the NCES CCD data; 2) absent value in the NCES CCD data; 3) No match from PSID CDS-II to NCES CCD