MEASUREMENT OF EARNINGS DIFFERENTIALS BETWEEN THE SEXES

Joseph L. Gastwirth, The George Washington University

Introduction

Anyone looking at earnings data is immediately struck by the difference in the average earnings of men and women. Since the principles of our country state that pay should be based on merit and skill factors and that everyone doing the same job should be paid the same wages, explaining the observed wage differences is a high priority social problem.

Our research began when I served as a Visiting Faculty Advisor to OMB and participated in a task force under the direction of the Deputy Assistant Attorney General for Civil Rights. While a substantial literature has been devoted to discovering the factors contributing to high earnings power (e.g., education, experience) and measuring (via regression models) their relative importance, it appeared that one needed a simple, yet statistically sound measure that would 1) enable one to detect areas of the labor market in which women are furthest behind men 2) be applicable to regularly issued statistical series so that progress could be followed over time.

2. Measures of Differentials

The problem of comparing male and female earnings distributions can be regarded as a two-sample problem. We consider the wages of women and men to come from theoretical distributions F(x) and G(x), respectively (where F(x) denotes the fraction of women earning less than x). The Census Bureau often uses the ratio of the medians to compare earnings and income distributions. Recently, they [9] have considered a new Overlap measure. In this section we review the Overlap measure and introduce a Probability measure based on the Wilcoxon test which we feel is superior for the current purpose.

The overlap measure (OVL) is best described in terms of the density functions f(x) and g(x) corresponding to F(x) and G(x). It is defined as the area under both g(x) and g(x), i.e.,

$$OVL = \int_0^\infty \min(f,g)\,dx, \qquad (2.1)$$

and is the shaded area in Figure 1.

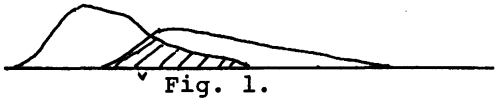

Fig. 1.

While the overlap measure has a nice pictorial representation it has several drawbacks: 1) Two widely different pairs of distributions can have the same value. Moreover, the cause of the disparity in each pair of distributions can be different. This can be illustrated by 2 pairs of distributions with OVL = 0.

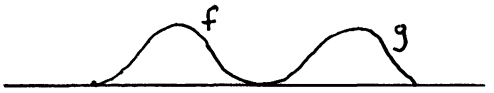

Fig. 2.a.

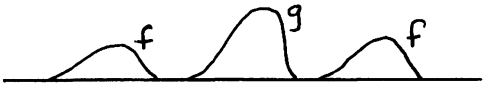

Fig. 2.b.

In Figure 2a, every member of the G population earns more than any member of the F group. In Figure 2b both populations have the same mean but the G population is concentrated near this mean value while the F distribution is really composed of two separate groups. 2) The OVL measure places undo-emphasis on the point, v. (see Fig. 1), where the density functions intersect. In particular, if a female earning more than v obtains a pay raise (while everyone else in both populations remains at the same level) the value of OVL remains unchanged.

As the second objection to the overlap measure also applies to the ratio of medians, we propose to use the probability that a randomly selected woman earns at least as much as a (randomly chosen) man. In order to rank industries, however, general social phenomena which depress women's wages relative to men tend to operate "across the board" so that the PROB measure should serve our purpose. More importantly, the PROB measure will detect any advancement of women relative to men so that it can be used to "monitor" upgrading programs.

Mathematically, the probability that a woman earns at least as much as a man can be expressed in terms of the distributions F(x) and G(x) as

$$PROB = \int_0^\infty [1-F(x)]g(x)\,dx. \qquad (2.2)$$

Because the PROB measure is related to the Mann-Whitney form of the Wilcoxon test, its standard deviation is known [4].

3. Analysis of the Longitudinal Social Security Data

The finding that women do not receive the same economic rewards as men for continuous labor force participation hardly needs extensive statistical documentation. Our task, however, is to show that the PROB measure reveals this fact and also detects small changes. Thus, one can use it to track the status of women over time.

The data base consisted of earnings data for two time periods, e.g. 1965-1970. Workers in the last period (1970) are split into those who worked in 1965 and

those who did not. (This group of workers consists of new and re-entrants to the labor force.) Similarly, the workers in the earlier period are grouped into those who had earnings in 1970 and those who did not. In order to avoid confounding differences based on sex with those due to race we discuss the results for the white population.

In Table 3.1 we present the values of the various measures for the total U.S. for the 1965-70 and 1962-67 periods. The earnings used were reported in the first quarter (multiplied by four).

While all three measures show a great disparity in the male-female earnings differentials between workers who entered (or re-entered) the work force in the 5 year period before 1970 (or 1967) and those who had worked five years earlier, it is more enlightening to look at men and women who worked at both times. When we contrast the differentials based on 1970 earnings to those derived from the 1965 earnings both the PROB and OVL measures DECREASED while the ratio of median incomes INCREASED. The same phenomenon also held during the 1962-67.

4. The Relative Status of Women in Various Industries

In order to demonstrate the utility of the PROB measure we apply it to Social Security data (by industry) for 1966.

Before presenting the results, some technical limitations of the data should be noted. The data is based on a 1% sample of earning records of people who worked in all four quarters. The data cannot distinguish between full and part-time employment, however, so that the over-all status of women may be biased downwards. The advantage of the Social Security data is that the sample is extremely large (390,000).

In Table 4.1 we present our results. Several major features emerge: 1) In almost all industries, black women fare better relative to black men than white women do to white men. 2) All the measures show that, across all industries, women to not fare well, however, the relative rankings do not agree. In particular, the Probability and Overlap measures give low scores to Communications, Public Utilities and Manufacturing (which, unfortunately, is highly aggregated here) while the Ratio of Medians yields a low rank for Retail Trade, Manufacturing, Communications and Services and a high one for Public Utilities and Transportation. 3) The PROB measure generally is relatively further away from its "ideal" value, 1/2, than the other measures are to theirs. This is a desirable property for tracking purposes.

5. Analysis of Occupational Data within an Industry

The low values of all the measures of earnings differentials in the aggregate industry data presented in section 4 can result from a) paying women less than men to for the same job and/or b) excluding women from jobs on a career ladder leading to promotions etc. thereby clustering women in the relatively low paying positions. The longitudinal study in section 3 supports the second explanation. The only available data enabling us to shed some light on this question is the BLS area Wage Survey which is collected regularly from firms on wages for precisely defined jobs.

In Table 5.1 the PROB value is given for a variety of Professional and Office jobs. The two sets of numbers tell different stories. Generally, earnings are nearer "equality" in the Professional categories than in Office jobs. Since women dominate (numerically) the Office jobs surveyed these results cannot be explained by excuses such as lack of qualified applicants with relevant work experience etc. Moreover, the higher the proportion of women in an occupation, the lower is their probability of equal pay. For example, females outnumber male Billers by 8:1 and PROB = .17,[1] they outnumber male order clerks 5:2 and PROB = .19. In the occupations with skill categories, e.g., Accounting Clerks - at the highest level women outnumber men 5:2 and PROB = .25 while at the lower level (B) women outnumber men by about 6:1 and in all industries (except Manufacturing) the PROB value is lower.

The data for Tabulating Machine Operators illustrates the relationship between employment segregation and lower pay for women. For the highest skill level (A), where men outnumber women 5:2, PROB = .35 and the ratio of medians equals .90. For the lower skill levels (B and C) where men still outnumber women the results are similar. Only in the Public Utilities (level B) workers where women outnumber men 2:1 does PROB fall to .26 and the ratio to .80.

Looking back at Table 5.1, in this light, one wonders whether the relatively high values of all our measures give most Professional and Technical occupations results from the scarcity of women in them. The only occupation in an industry category where women are employed in nearly the same numbers as men was computer operators in the Public Utilities[2] - which received the lowest PROB score in the Professional class.

6. Summary

The purpose of our paper was to illustrate how a simple measure of earnings differentials can be used to rank industries (or occupations) and to monitor progress over time. By analyzing several

U.S. government data series we showed that

1) Women do not receive the same economic return for continuous work as men. Indeed, they fall further behind as time (in the labor force) passes.

2) The relative status of the sexes is nearer equality in occupations in which men are employed in substantial numbers. Low values of our measure of equality occurred where women dominate (numerically) the job.

2) In order to obtain a complete picture one should study employment as well as earnings data. A low score can result from a variety of factors, e.g. new hiring as well as placing women primarily in low paying jobs.

It is important to remember that broad statistical series cannot prove that discrimination exists, however, the tools developed can aid in the process of monitoring progress. Moreover, the data used in the section 5 is available to most large companies. If the Public Utilities data had been for one firm, management could immediately spot that something might be amiss in their computer-operator division.

In addition, I hope this paper will also stimulate professional statisticians to work with various government data series and point out to other social scientists who use the data which data sets are most appropriate for various types of analysis. Finally, I would like to thank the Women's Caucus of ASA for honoring me with the invitation to prepare this paper.

Acknowledgement: It is a pleasure to thank David Melcovsky for his assistance with the computer aspects of the paper and Carol Fey for typing several versions. Finally, the partial support of this research by an NSF institutional grant to George Washington University is gratefully acknowledged.

FOOTNOTES

[1]Notice, however, in the Public Utilities where the number of female and male Billers are nearly equal the PROB = .35.

[2]In the Area Wage Survey Public Utilities includes Utility Companies and the Communications and Transportation industries.

REFERENCES

[1] Buckley, J. E. "Pay Differences between men and women in the same job" Monthly Labor Review Nov. 1971 pp. 26-39.

[2] Fuchs, V. R. "Differences in hourly earnings between men and women" Monthly Labor Review May 1971 pp. 9-15.

[3] Gastwirth, J. L. A New Index of Income Inequaltiy. Proc. I.S.I. meeting in Vienna, 1973.

[4] Gibbons, J. D. Nonparametric Statistical Inference, McGraw Hill: New York, 1970.

[5] Henle, P. "Exploring the Distribution of Earned Income," Monthly Labor Review Dec. 1972 p. 16-27.

[6] Kreps, J. Sex in the Marketplace: American Women at Work Johns Hopkins Press. Baltimore, Maryland, 1971.

[7] McNulty, D. J. "Differences in Pay Between Men and Women Workers" Monthly Labor Review Dec. 1967 p. 40-43.

[8] Strasser, A. Differentials and Overlaps in Annual Earning of Blacks and Whites Monthly Labor Review Dec. 1971 p. 16-26.

[9] Weitzman, M. Measures of Overlap of Income Distributions of White and Negro Families in the U.S. Tech. Paper 22 Bureau of the Census, 1970.

Table 3.1 Measures of Male-Female Differences for the United States White Population

	PROB	OVL	RATIO	MEDIAN MALE	MEDIAN FEMALE	MEAN MALE	MEAN FEMALE
1970							
All Workers	0.255	0.569	0.505	7454	3764	8027	4024
Worked in 1965	0.217	0.530	0.540	8424	4548	9224	4850
New Workers	0.427	0.819	0.839	3369	2825	4370	3062
1965							
All Workers	0.243	0.560	0.507	5623	2853	5988	3020
Worked in 1970	0.234	0.547	0.527	5879	3097	6266	3261
Dropouts	0.309	0.658	0.536	4380	2349	4880	2561
1967							
All Workers	0.275	0.606	0.417	5590	2331	6070	2721
Worked in 1962	0.222	0.536	0.498	6761	3370	7380	3532
New Workers	0.446	0.867	0.795	1492	1186	2408	1733
1962							
All Workers	0.266	0.604	0.427	4599	1965	5074	2255
Worked in 1967	0.264	0.604	0.478	4866	2325	5313	2520
Dropouts	0.330	0.711	0.424	2831	1200	3834	1669

Source: Bureau of Economic Analysis (Regional Econ. Div.) Department of Commerce

Table 4.1 Measures of the Relative Status of Women in Various Industries. Derived from the 1966 Social Security Data for 4-Quarter Workers

	White			Black		
Ind.	PROB	OVL	RATIO	PROB	OVL	RATIO
Total	.185	.480	.508	.252	.620	.570
Construction	.217	.545	.570	.312	.648	.717
Mining	.211	.567	.663*	.162	.204	.407*
Manuf.	.136	.404	.526	.226	.567	.606
Trans.	.205	.513	.710	.335	.674	.870*
Commun.	.105	.293	.523	.342	.520	.816
Pub. Util.	.150	.434	.637	.337	.431	.861
Whol. Tr.	.178	.472	.574	.275	.631	.691
Ret. Tr.	.229	.521	.480	.322	.693	.709
Finance etc.	.176	.426	.564	.410	.830	.914*
Service	.259	.612	.534	.308	.697	.599

*These values are based on a very small sample.

Table 5.1 The PROB Measure of Earnings Differentials Evaluated on Weekly Wage Data (1970-1971)

Prof. & Tech. Occupations	Industry	PROB	MEDIAN FEMALE	MEDIAN FEMALE	TOTAL FEMALE	TOTAL MALE
Comp Op. A	All	0.400	158.39	166.72	950	10834
	Manufacturing	0.464	165.57	168.58	383	4830
	Finance	0.356	146.05	158.39	287	2738
Comp Op. B	All	0.299	125.35	144.10	3984	18801
	Manufacturing	0.360	136.16	148.38	1176	6996
	Public Utilities	0.093	115.22	165.21	982	1212
	Wholesale Trade	0.300	126.82	147.29	424	2002
	Finance	0.313	120.40	134.43	976	5494
Comp Op. C	All	0.371	110.87	120.58	2634	7957
	Manufacturing	0.394	117.32	126.50	766	2439
	Public Utilities	0.243	105.21	130.62	795	498
	Finance	0.378	104.20	113.92	577	3058
Comp Prog A	All	0.424	217.88	226.72	1991	11131
	Manufacturing	0.411	219.62	230.42	558	4836
	Public Utilities	0.340	210.68	233.42	269	974
	Finance	0.464	210.39	215.56	607	2682
	Services	0.437	219.86	224.84	283	1303
Comp Prog B	All	0.430	182.39	189.89	3901	13987
	Manufacturing	0.439	187.48	195.02	1136	5575
	Public Utilities	0.412	195.59	204.29	481	1503
	Finance	0.453	175.84	179.22	1655	4220
Comp Prog C	All	0.443	156.61	161.58	2323	5747
	Manufacturing	0.413	159.76	170.52	632	2003
	Public Utilities	0.470	169.73	172.05	308	637
	Finance	0.474	150.97	152.83	1037	2190
Office Occupations						
Billers, Machine	All	0.173	100.64	153.80	8695	1159
	Public Utilities	0.350	143.40	162.01	1207	940
Clerks, Acct., A	All	0.255	127.66	152.19	54544	19228
	Manufacturing	0.258	132.51	156.20	20641	9326
	Public Utilities	0.246	130.56	162.70	7314	3159
	Wholesale Trade	0.268	128.95	150.04	6343	2389
	Retail Trade	0.295	118.94	138.52	6561	804
	Finances	0.265	117.29	138.24	9044	2295
	Services	0.355	130.35	142.73	4654	799
Clerks, Acct., B	All	0.218	100.42	127.92	103110	10211
	Manufacturing	0.253	105.62	127.86	30861	3399
	Public Utilities	0.191	105.83	142.73	16292	2413
	Wholesale Trade	0.203	103.14	133.00	13813	1950
	Retail Trade	0.363	94.48	105.75	18130	443
	Finance	0.201	93.01	114.03	17285	1467
	Services	0.312	101.96	114.21	6733	546
Secretaries, C	All	0.248	138.02	162.50	98895	368
Tab.Mach. Op. A	All	0.345	142.17	157.30	1002	2583
Tab. Mach. Op. B	All	0.355	117.46	130.60	3199	3911
	Public Utilities	0.261	113.81	143.61	986	484
Tab. Mach. Op. C	All	0.352	99.85	111.54	1805	1861