INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
Garen, John Edward

THE EFFECT OF FIRM SIZE ON WAGE RATES

The Ohio State University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106
THE EFFECT OF FIRM SIZE
ON WAGE RATES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

John Edward Garen, B.A., M.A.

* * * * *

The Ohio State University

1982

Reading Committee:                 Approved By
Belton M. Fleisher                   Donald O. Parsons
Howard P. Marvel                    Adviser
Donald O. Parsons                   Department of Economics
To Bev
The author gratefully acknowledges the help of his dissertation committee consisting of Professors Belton M. Fleisher, Howard P. Marvel, and Donald O. Parsons. He is especially grateful to Professors Parsons and Marvel for their valuable suggestions and guidance at various stages of writing this thesis.

The author is also grateful for the financial support of the National Institute on Aging and The Ohio State University's Presidential Fellowship program.

For computational assistance, the author wishes to thank James Brown, Donald Dickerson, and Bruce Holloway.
VITA

November 24, 1953. . . . . . . . Born - Indianapolis, Indiana

1976 . . . . . . . . . . . . . . . . . . . . . . . B.A., University of Washington, Seattle, Washington

1978 . . . . . . . . . . . . . . . . . . . . . . . M.A., The Ohio State University, Columbus, Ohio

1976-1980. . . . . . . . . . . . . . . Teaching Associate, Department of Economics, The Ohio State University, Columbus, Ohio

1980-1981. . . . . . . . . . . . . . . Graduate Fellow, The Ohio State University, Columbus, Ohio

1981-. . . . . . . . . . . . . . . . . . . . . . . Assistant Professor of Economics, Wayne State University, Detroit, Michigan

FIELDS OF STUDY

Major Field: Economics

Studies in Labor Economics. Professors Donald O. Parsons and Belton M. Fleisher.


TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................ iii
VITA ................................................... iv
LIST OF TABLES ........................................ vi
INTRODUCTION .......................................... 1

Chapter

I. LITERATURE REVIEW .................................. 3

II. SOME HYPOTHESES .................................. 10

The Data ................................................. 10
Hypothesis Tests ........................................ 12

III. THE MODEL .......................................... 29

The Wage Offer Function ................................ 31
The Evaluation Cost Function .......................... 38
Labor Cost Minimization ................................ 42
Profit Maximization .................................... 43
Labor Market Equilibrium .............................. 44

IV. ADDITIONAL EVIDENCE .............................. 46

The Ability Gradient .................................... 46
Wage Dispersion ........................................ 51
Other Worker Characteristics ........................ 56
Worker Productivity .................................... 59

V. CONCLUSION .......................................... 61

APPENDIX

Properties of the Wage Offer Function .................. 63

FOOTNOTES ............................................... 67

LIST OF REFERENCES .................................... 71
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Means, Standard Deviations of Important Variables, Production Workers in Manufacturing, NLS Older Men and Young Men</td>
<td>13</td>
</tr>
<tr>
<td>2.</td>
<td>Coefficient Estimates of Logarithmic Wage Equations, Production Workers in Manufacturing, NLS Older Men and Young Men</td>
<td>16</td>
</tr>
<tr>
<td>4.</td>
<td>Coefficient Estimates of Logarithmic Earnings Equations, Production Workers in Manufacturing, NLS Older Men and Young Men</td>
<td>21</td>
</tr>
<tr>
<td>5.</td>
<td>Probit Estimates of Probability of Being Covered by a Pension Plan, Production Workers in Manufacturing, NLS Older Men</td>
<td>22</td>
</tr>
<tr>
<td>6.</td>
<td>Coefficient Estimates of Logarithmic Wage Equations, Production Workers in Manufacturing, NLS Young Men, Measures of Ability Included</td>
<td>24</td>
</tr>
<tr>
<td>7.</td>
<td>Coefficient Estimates of Logarithmic Wage Equations, Production Workers in Manufacturing, NLS Young Men, by Union Status</td>
<td>26</td>
</tr>
<tr>
<td>8.</td>
<td>Coefficient Estimates of Logarithmic Wage Equations, Production Workers in Manufacturing, NLS Older Men, by Union Status</td>
<td>27</td>
</tr>
<tr>
<td>9.</td>
<td>Coefficient Estimates of Logarithmic Wage Equations, Production Workers in Manufacturing, NLS Young Men, Ability Interactions</td>
<td>50</td>
</tr>
</tbody>
</table>
10. F-tests for Differences in Standard Deviations in Wages and Standard Errors of Wage Regressions, Production Workers in Manufacturing, CPS, NLS Young Men and Older Men ............................. 53

11. Coefficient Estimates of the Determinants of Schooling, Production Workers in Manufacturing, CPS, NLS Young Men and Older Men ............................. 58
INTRODUCTION

A positive association between firm size and wage rates has been observed frequently by economists. This relationship remains even after worker quality is accounted for. The finding is somewhat perplexing since it is contrary to standard labor market theory. In the simple model of competitive labor markets, all firms pay the same wage for workers of a given skill. Although a larger firm's marginal productivity of labor schedule may lie above that of a smaller firm, the larger firm simply hires more workers. The larger firm continues to hire additional labor until its marginal product equals the market wage rate. The market wage is the same for all firms, large and small. Obviously, this simple model is not sufficient to explain the phenomenon in question.

The first objective of this study is to examine carefully the empirical relationship between firm size and wage rates. Chapter I surveys the substantial amount of literature concerned with this issue. Some studies have directly addressed the issue, while for others, it is tangential to the main thrust of the research. All are empirically oriented. Few provide any explanation of the findings other than conjecture. The main conclusion that can be drawn from this body of literature is
the relationship between firm size and wage rates is quite robust.

The second objective is to provide a theoretical explanation for this phenomenon. The explanation should imply not only that large firms pay higher wages, but also should contain a richer set of implications which furnish a means of further testing its validity. Chapter II develops and carries out statistical tests of several simple hypotheses regarding the firm size, wage rate association. None are accepted as satisfactory explanations.

Chapter III presents a model of wage rate determination which yields the implication that large firms pay higher wages. It is based on the difficulties large firms have in screening their workers. In the absence of efficient screening, large firms may pay a premium to insure themselves of a high quality labor force.

This model has further implications which, among other things, yields inferences about other aspects of wage structure in large and small firms. These and other predictions are tested and supported in Chapter IV. The conclusion is presented in Chapter V.
The amount of literature concerning the relationship between firm size and wage rates is substantial. For the most part, it deals with this issue in a strictly empirical manner, i.e., it offers no theory to explain the firm size, wage rate correlation. There is also a related literature which predates this body of research. The interest of these studies is in the correlation between wage rates and product market concentration. This is related to the firm size, wage rate relationship because concentration and firm size are strongly correlated. Due to this strong correlation, any relationship between wages and concentration may carry over to wages and firm size.

The focus of the older literature is on the monopoly wage hypothesis. This hypothesis maintains that concentrated industries pay higher wages because they can afford to be extravagant in wage settlements. If this is true, wages and firm size are likely to be correlated because wages and concentration will be correlated.

Several authors attempt to test the monopoly wage thesis by relating wage changes to concentration. Ross and Goldner
(1950), Garbarino (1950), and Bowen (1960) correlate industry wage changes to concentration and unionism. Each finds a positive association between concentration and wage increases. Reder (1962) and Lewis (1963) both call into question the results and methodology of these studies. They demonstrate the relationship between wage changes and concentration is unstable and sensitive to the time period chosen. More importantly, each points out a correct test of this hypothesis should relate the level of wages with concentration. The theory implies monopolists pay higher wages, but nothing in it suggests they will increase wages faster than other firms.

Segal (1964) extends the theory one additional step. He claims concentrated industries will pay higher wages only when faced with a strong union. Otherwise, the argument states, there is no means for workers to extract a part of the monopoly profit. This view fell out of acceptance after the publication of an article by Weiss (1966). Using the wage rates of individual workers as the dependent variable, his regression analysis shows concentration has a stronger effect when unions are weaker. Furthermore, any influence of concentration disappears once individuals' characteristics are entered in the regression equation. Later empirical studies lend little support to the monopoly wage theory. Most indicate concentration has no significant impact on wages.

This older literature does present a potential explanation for the firm size effect on wages, but the evidence suggests it
is lacking. Perhaps this is not surprising. One expects both monopolists and competitive firms to minimize costs which implies paying a wage equal to that prevailing in the market.

Articles of the more recent literature directly correlate firm size and wages. Several studies use industry averages of wage rates and establishment size. Lester (1967) calculates the simple correlation coefficient between the two and finds it is positive. Very little can be concluded from this because the skill of workers in different industries is not controlled for. This correlation could result if large firms simply hire a more skilled labor force. Masters (1969) regresses average industry wage on the extent of industry unionism, region and race variables, and average plant size. The plant size coefficient is positive and significant. However, this study suffers from the same criticisms as Lester's; worker skills are not accounted for. Of the two, only Masters ventures an explanation. He conjectures that the results may be due to two factors: a compensating differential because of presumably less pleasant working conditions in large firms, and/or a premium only large firms pay to discourage unionism. One might also add it could result from large firms hiring better workers. Masters tests neither of his hypotheses.

Haworth and Rasmussen (1971), in an extension of Masters' paper, make an attempt to standardize for worker quality. They introduce to the analysis a measure of the skill mix and educational attainment of workers in each industry. These variables are included as regressors in the wage equations. The results show
firm size still has a positive and significant effect on wages. The authors do not attempt to pursue the analysis further, offering neither additional empirical work nor a theory to explain their results. Their findings are, however, much stronger than Lester's or Masters'. They indicate large firms pay workers of the same skill more. This is consistent with the findings of Rosen (1969, 1970). Rosen uses very similar data to investigate union/nonunion wage differentials. He includes a firm size variable as a regressor in his wage regressions, along with measures of the age, education, and extent of unionism of each industry's labor force. The firm size coefficient in these equations is positive and significant. Rosen did not explore this result, more than likely because it is tangential to the issue discussed in his papers.

These findings have been substantiated on disaggregate data sets. Bailey, King, and Schwenk (1970) and Bailey and Schwenk (1980) utilize establishment-based data in their work. They regress the average wage paid by an establishment on the establishment's size, unionism, region and race variables, and on a measure of the occupation mix within the plant. They find establishment size is positively related to average establishment wage. In both papers, the authors continue by introducing into the analysis measures of worker productivity: value added per employee and value of shipments per employee. They find these measures of worker productivity are correlated with wages and with firm size, but do not reduce the effect of firm size on wages to zero. They conclude that higher productivity may be one reason large firms pay more. This inference
is questionable, however. Occupation may not adequately control for worker skill and if large firms hire more educated, experienced people in each occupation, one expects them to be more productive. Furthermore, even if occupation adequately controls for worker skill, the finding yields no theoretical insights. If workers of the same skill are more productive in large firms, large firms should hire more until labor's marginal productivity is equal to its opportunity wage.

Freeman and Medoff (1978), using establishment data, estimate logarithmic wage equations with similar regressors to those Bailey, King, and Schwenk use, but also include education and experience per worker. Even with better control for worker quality, the average wage paid by an establishment rises with its size. The authors, in another part of their study, estimate wage equations using data on individual workers. They find a positive and significant relationship between the firm size variable and wage rates after controlling for schooling, experience, unionism, and other worker traits such as occupation and race. Their study is concerned primarily with the estimation of the effect of the percent of union workers in an industry on wages, so they do not develop or test hypotheses about the effect of firm size. They merely state the result they find is consistent with the work of others.

Utgoff's (1980) work corroborates this result on another data file of individual workers. The focus of the paper is on the effect of firm size on quits. Wages are an important determinant of quits, thus Utgoff incorporates a wage determination equation
in the analysis with firm size as a regressor. She finds firm size has a positive and significant impact on wages. Because the analysis is more concerned with quits than wage rates, no explanation for this result is suggested.

Miller (1980) is the only author whose results are contrary to the others. Ironically, he is also the only author who develops a formal, theoretical model of firm size and wage rates. His model implies only managers should earn higher wages in large firms. Managers are assumed to choose the place in the firm's hierarchy for which they are best suited. Those with superior management talent choose to supervise more individuals, so choose a higher place in the hierarchical structure. Large firms have more developed hierarchies, so they offer positions higher in the hierarchy than small firms do. Superior management is therefore attracted to large firms where they earn higher wages than the less talented managers of small firms. His empirical results indicate managers and the more educated workers earn more in large firms. However, for production workers, any correlation between firm size and wages disappears once unionism is accounted for. This is puzzling because all other research in this area show a positive effect of firm size on wages, even after controlling for unionism. Miller's results seem to be anomalous.

Taken together, the papers surveyed establish a strong, empirical regularity. The positive effect of firm size on wages is a robust result, having been verified on a wide variety of data sets covering different time periods. The only exception is Miller.
Although this phenomenon is well established, there has been little attempt to explore it carefully or to develop models explaining its existence.
Chapter II

SOME HYPOTHESES

The intent of this chapter is to develop and execute statistical tests of the conjectures and other straightforward explanations of the firm size effect on wage rates. A discussion of the data used is in order before proceeding.

The Data

Previous authors have used a variety of data sets: industrial, establishment, and individual data. A priori, establishment and individual data seem superior to industry data since there is less potential for aggregation bias. The aggregation problem does not disappear, however. With establishment data, one must aggregate over all workers in the establishment to obtain wages and worker characteristics variables. The advantage is no aggregation is required to obtain establishment size. Using individual data, the opposite is true. Worker characteristics are readily available, but the industry of employment is all that is known about the workplace. To obtain an establishment size measure, one must aggregate over all establishments in the industry. Individual-based data sets were chosen for use in this study. In general, they provide a much richer set of information regarding workers. In addition, all
of the variables regarding establishments can be obtained from industry data, although in aggregate form. These data are then merged with the data on individual workers.

The data on individual workers used in this study are primarily the 1969 National Longitudinal Survey (NLS) of Older Men and the 1969 National Longitudinal Survey of Young Men. Some use is made of the March 1971 Annual Demographic File of the Current Population Survey. All three data sets provide information on many personal characteristics of each individual. The NLS data include some important variables which the CPS does not have, thus the reason for its primary use. In particular, the CPS has no job tenure variable and contains no measure of innate ability, such as IQ. Both are important determinants of wages and the ability measure plays an important part in this investigation. The study is limited to male production workers who were employed in manufacturing at the time they were surveyed. This makes the sample size for each NLS data set roughly seven hundred and for the CPS, approximately five thousand. Only manufacturing industries are included since establishment size data are generally limited to these industries. Nonproduction workers are excluded because the focus of the earlier work is on production workers only.

The establishment data are collected from the 1967 Census of Manufactures. Each industry's establishment size distribution is characterized by a single number, and this number is assigned to each individual employed in that industry. This serves as a measure of firm size. The number used to characterize an industry's
size distribution is the percent of an industry's workers employed in establishments of greater than five hundred employees. This is conceptually comparable to the measure used commonly in the literature. It can be interpreted as one hundred times the probability of a worker in a given industry being employed in a plant of at least five hundred employees.

Table 1 presents some summary statistics for the NLS Young Men and the Older Men data sets. Means and standard deviations of the variables of interest are shown. These two data sets will be used almost exclusively.

Hypothesis Tests

The approach used here and by others, to test for a firm size effect on wages, is to estimate logarithmic wage equations with firm size included as a regressor. A firm size effect is indicated by a significant coefficient on the firm size variable. Column 1 of Table 2 reports the estimates of such an equation for the NLS Young Men data set and column 2 does likewise for the Older Men survey. Included as regressors are the human capital variables, schooling, experience, and tenure, along with unionism variables. All have the expected sign. Several dummy variables are also included to standardize for locational, occupational, and racial differences. The coefficient of particular interest, that of firm size (PCT500), is positive and highly significant in both equations. This corroborates the findings of earlier investigators.

Some conclusions concerning the firm size effect can already be drawn. First, it appears that the firm size effect is not an
Table 1

Means, Standard Deviations of Important Variables, 
Production Workers in Manufacturing, 
NLS Older Men and Young Men 
(standard deviations in parentheses)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Young Men (1)</th>
<th>Older Men (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schooling</td>
<td>10.95</td>
<td>8.67</td>
</tr>
<tr>
<td></td>
<td>(1.85)</td>
<td>(3.09)</td>
</tr>
<tr>
<td>Experience</td>
<td>4.96</td>
<td>39.14</td>
</tr>
<tr>
<td></td>
<td>(3.75)</td>
<td>(5.53)</td>
</tr>
<tr>
<td>Tenure</td>
<td>1.49</td>
<td>16.56</td>
</tr>
<tr>
<td></td>
<td>(1.90)</td>
<td>(10.71)</td>
</tr>
<tr>
<td>Union Member</td>
<td>.422</td>
<td>.604</td>
</tr>
<tr>
<td></td>
<td>(.494)</td>
<td>(.489)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>50.64</td>
<td>53.8</td>
</tr>
<tr>
<td></td>
<td>(18.03)</td>
<td>(17.31)</td>
</tr>
<tr>
<td>PCT500</td>
<td>45.88</td>
<td>49.58</td>
</tr>
<tr>
<td></td>
<td>(25.59)</td>
<td>(26.32)</td>
</tr>
<tr>
<td>Wage Rate</td>
<td>2.90</td>
<td>3.43</td>
</tr>
<tr>
<td></td>
<td>(.975)</td>
<td>(1.14)</td>
</tr>
<tr>
<td>Sample Size</td>
<td>693</td>
<td>768</td>
</tr>
</tbody>
</table>

The following sets of dummy variables are also used. 
Location: South, rural residence 
Race: Nonwhite 
Occupation: Craftsmen, operative 

Source: 1969 National Longitudinal Surveys
Table 1 (continued)

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Years of schooling completed.</td>
</tr>
<tr>
<td>b</td>
<td>Age minus schooling minus five.</td>
</tr>
<tr>
<td>c</td>
<td>Years with current employer.</td>
</tr>
<tr>
<td>d</td>
<td>Dummy variable equal to one if union member, zero otherwise.</td>
</tr>
<tr>
<td>e</td>
<td>Percent of production workers in the respondent's industry that are union members. Source is Freeman and Medoff (1979).</td>
</tr>
<tr>
<td>f</td>
<td>Percent of workers in the respondent's industry that are employed in establishments of five hundred or more workers.</td>
</tr>
<tr>
<td>g</td>
<td>Wage rate earned in dollars per hour.</td>
</tr>
</tbody>
</table>
illusion created by different compensation schemes over the life-cycle of workers. For example, if large firms steepen the wage profile by paying young workers less and older workers more, then a wage equation is misspecified unless firm size is interacted with experience. If this is not done, the firm size coefficient may appear to be positive due to specification error. Table 2 reveals that this is not the case. The firm size coefficient is positive for both age groups. Furthermore, the coefficient in the Older Men equation is not significantly different from the coefficient in the Young Men equation \( (t=0.98) \). The effect of firm size varies little across age groups.

This result may be somewhat suspect since there are no observations on workers in the middle of their careers. In order to rectify this shortcoming, wage equations are estimated using the CPS data, which includes workers of all ages. Table 3 reports these findings. Column 1 is for all age groups; column 2, for those with ten or fewer years of experience; column 3, for workers who have between ten and thirty years of experience, and column 4, for those with thirty or more years of experience. The firm size coefficient is positive for all workers collectively and for each experience group individually. It is also highly significant in each case, with the exception of the lowest experience group. In addition, for all three experience groups, none of the coefficients on PCT500 are significantly different from the others. The conclusion drawn is the effect of firm size on wages lasts throughout a
Table 2

Coefficient Estimates of Logarithmic Wage Equations,
Production Workers in Manufacturing

NLS Older Men and Young Men\textsuperscript{a}

(Absolute value of t-ratios in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>Young Men (1)</th>
<th>Older Men (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.2306</td>
<td>.8830</td>
</tr>
<tr>
<td></td>
<td>(2.64)</td>
<td>(6.65)</td>
</tr>
<tr>
<td>Schooling</td>
<td>.0432</td>
<td>.0167</td>
</tr>
<tr>
<td></td>
<td>(6.69)</td>
<td>(3.69)</td>
</tr>
<tr>
<td>Experience</td>
<td>.0208</td>
<td>-.0030</td>
</tr>
<tr>
<td></td>
<td>(6.34)</td>
<td>(1.29)</td>
</tr>
<tr>
<td>Tenure</td>
<td>.0146</td>
<td>.0055</td>
</tr>
<tr>
<td></td>
<td>(2.70)</td>
<td>(6.09)</td>
</tr>
<tr>
<td>Union Member</td>
<td>.1375</td>
<td>.0520</td>
</tr>
<tr>
<td></td>
<td>(6.78)</td>
<td>(2.44)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>.0020</td>
<td>.0003</td>
</tr>
<tr>
<td></td>
<td>(3.20)</td>
<td>(0.39)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.0014</td>
<td>.0020</td>
</tr>
<tr>
<td></td>
<td>(3.47)</td>
<td>(4.42)</td>
</tr>
<tr>
<td>Dummy Variables</td>
<td>included\textsuperscript{b}</td>
<td>included</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.491</td>
<td>.470</td>
</tr>
<tr>
<td>Sample Size</td>
<td>693</td>
<td>768</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Variables defined in Table 1

\textsuperscript{b} Indicates the dummy variables listed in Table 1 are included in the regression equation.

Source: 1969 National Longitudinal Surveys
worker's career and is not a life-cycle phenomenon.

A second conclusion drawn from Table 2 is the firm size effect is not simply a return to a low turnover rate among workers in large firms. The firm size coefficient is still positive even after standardizing for tenure, a measure related to turnover.11

Table 4 demonstrates that not only do employees in large firms have higher wages, but their yearly earnings are also higher. It shows the results of estimating logarithmic earnings equations with the same independent variables as before. Firm size has a positive and significant impact on yearly earnings for both the Young Men and Older Men surveys. The impact of firm size on earnings has approximately the same magnitude as its impact on wage rates. Apparently, the higher wages of workers in large firms is directly translated into higher earnings.

Wages are not the only form of compensation workers receive. Fringe benefits, to some degree, can be substituted for earnings. Hence, it is possible that small firms pay their workers lower wages because they receive more in fringe benefits. To test this, a probit equation is estimated of the probability of being covered by a pension plan. The findings are shown in Table 5. The table reveals that the likelihood of a pension plan rises with firm size. If this is indicative of other fringe benefits, then the wage differential understates the effect of firm size on compensation. Thus, this hypothesis can be rejected.

Workers are likely to differ in ways that are not captured by the variables used in Tables 1 through 5. One way workers
Table 3

Coefficient Estimates of Logarithmic Wage Equations,
Male Production Workers in Manufacturing,
Current Population Survey

<table>
<thead>
<tr>
<th>Variable</th>
<th>All (1)</th>
<th>Experience ≤ 10 (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.2681</td>
<td>-.3076</td>
</tr>
<tr>
<td></td>
<td>(4.94)</td>
<td>(1.80)</td>
</tr>
<tr>
<td>Schooling</td>
<td>.0297</td>
<td>.0571</td>
</tr>
<tr>
<td></td>
<td>(10.09)</td>
<td>(5.40)</td>
</tr>
<tr>
<td>Experience</td>
<td>.0315</td>
<td>.0636</td>
</tr>
<tr>
<td></td>
<td>(16.49)</td>
<td>(11.12)</td>
</tr>
<tr>
<td>Experience Squared</td>
<td>-.00047</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>(12.01)</td>
<td></td>
</tr>
<tr>
<td>Union Member</td>
<td>.1131</td>
<td>.2286</td>
</tr>
<tr>
<td></td>
<td>(7.79)</td>
<td>(6.02)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>.0017</td>
<td>.0017</td>
</tr>
<tr>
<td></td>
<td>(3.81)</td>
<td>(1.49)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.0014</td>
<td>.0010</td>
</tr>
<tr>
<td></td>
<td>(4.78)</td>
<td>(1.33)</td>
</tr>
<tr>
<td>Dummy Variables</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>.234</td>
<td>.204</td>
</tr>
<tr>
<td>Sample Size</td>
<td>5270</td>
<td>1339</td>
</tr>
</tbody>
</table>

\(^a\) Variables defined in Table 1

Source: March 1971 Current Population Survey
Table 3 (continued)

<table>
<thead>
<tr>
<th></th>
<th>10 &lt; Experience &lt; 30</th>
<th>Experience ≥ 30</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Constant</td>
<td>.6311</td>
<td>.7362</td>
</tr>
<tr>
<td></td>
<td>(7.74)</td>
<td>(7.07)</td>
</tr>
<tr>
<td>Schooling</td>
<td>.0346</td>
<td>.0235</td>
</tr>
<tr>
<td></td>
<td>(8.01)</td>
<td>(6.57)</td>
</tr>
<tr>
<td>Experience</td>
<td>.0039</td>
<td>.0027</td>
</tr>
<tr>
<td></td>
<td>(2.29)</td>
<td>(1.53)</td>
</tr>
<tr>
<td>Experience Squared</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Union Member</td>
<td>.0410</td>
<td>.0940</td>
</tr>
<tr>
<td></td>
<td>(1.99)</td>
<td>(4.67)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>.0013</td>
<td>.0019</td>
</tr>
<tr>
<td></td>
<td>(1.83)</td>
<td>(3.18)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.0017</td>
<td>.0015</td>
</tr>
<tr>
<td></td>
<td>(3.72)</td>
<td>(3.69)</td>
</tr>
<tr>
<td>Dummy Variables</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.171</td>
<td>.268</td>
</tr>
<tr>
<td>Sample Size</td>
<td>2151</td>
<td>1780</td>
</tr>
</tbody>
</table>
Table 4

Coefficient Estimates of Logarithmic Earnings Equations, Production Workers in Manufacturing, NLS Older Men and Young Men\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>Young Men (1)</th>
<th>Older Men (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>6.263</td>
<td>8.712</td>
</tr>
<tr>
<td></td>
<td>(29.09)</td>
<td>(64.47)</td>
</tr>
<tr>
<td>Schooling</td>
<td>.1084</td>
<td>.0144</td>
</tr>
<tr>
<td></td>
<td>(6.25)</td>
<td>(3.11)</td>
</tr>
<tr>
<td>Experience</td>
<td>.0904</td>
<td>-.0060</td>
</tr>
<tr>
<td></td>
<td>(10.76)</td>
<td>(2.50)</td>
</tr>
<tr>
<td>Tenure</td>
<td>.0405</td>
<td>.0051</td>
</tr>
<tr>
<td></td>
<td>(3.05)</td>
<td>(5.48)</td>
</tr>
<tr>
<td>Union Member</td>
<td>.2739</td>
<td>.0288</td>
</tr>
<tr>
<td></td>
<td>(5.51)</td>
<td>(1.33)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>-.0014</td>
<td>.0007</td>
</tr>
<tr>
<td></td>
<td>(0.89)</td>
<td>(0.96)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.0026</td>
<td>.0015</td>
</tr>
<tr>
<td></td>
<td>(2.53)</td>
<td>(3.14)</td>
</tr>
<tr>
<td>Dummy Variables</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>(R^2)</td>
<td>.398</td>
<td>.453</td>
</tr>
<tr>
<td>Sample Size</td>
<td>678</td>
<td>768</td>
</tr>
</tbody>
</table>

\(^a\) Variables defined in Table 1

Source: 1969 National Longitudinal Surveys
Table 5

Probit Estimates of Probability of Being Covered by a Pension Plan, Production Workers in Manufacturing, NLS Older Men\textsuperscript{a, b}

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Absolute Value of Asymptotic t-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.395</td>
<td>(1.61)</td>
</tr>
<tr>
<td>Schooling</td>
<td>.0549</td>
<td>(1.94)</td>
</tr>
<tr>
<td>Experience</td>
<td>-.0181</td>
<td>(1.14)</td>
</tr>
<tr>
<td>Tenure</td>
<td>.0356</td>
<td>(5.95)</td>
</tr>
<tr>
<td>Union Member</td>
<td>.4439</td>
<td>(3.36)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>.0197</td>
<td>(3.83)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.0078</td>
<td>(2.49)</td>
</tr>
</tbody>
</table>

Dummy Variables:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occupation (2)</td>
<td>included</td>
</tr>
</tbody>
</table>

Log Likelihood: -359.62

Sample Size: 664

\textsuperscript{a} Absolute value of asymptotic t-ratios in parentheses.

\textsuperscript{b} Variables defined in Table 1

Source: 1969 National Longitudinal Surveys
differ, which is seemingly relevant to wage determination, is some have more innate ability than others. These workers should earn higher wages, since they are more productive. If big firms hire a disproportionately large share of high ability individuals, they would appear to be paying high wages, when in fact it is actually a return to the innate talent of their workers.

Fortunately, the NLS Young Men survey provides a possible way to test this. There are two measures of ability in this data file. One is the IQ test score. This variable is incorporated into the analysis by estimating a wage equation with IQ included as a regressor. Column 1 of Table 6 reports the results. The coefficients on the other variables are largely unchanged, with the exception of that on schooling. This is also true of the firm size coefficient. It is still positive and significant, indicating IQ does not account for the positive influence of firm size on wages. However, there may be some difficulties with this evidence. Over one third of the respondents in the sample have no IQ score reported for them. This trims the sample to 428. If these IQ scores are missing non-randomly, the results may be biased. This may account for the insignificant effect of IQ on the wage rate. This problem can be addressed using the methods of Griliches, Hall, and Hausman (1978). They apply statistical procedures to estimate IQ's of individuals for whom none was reported. This allows them to use the entire sample. A simpler and less expensive method to avoid this problem is using the second measure of ability on the NLS. This is the approach taken here.
This second measure is the score on the "knowledge of the world of work" questionnaire. Each respondent was given a quiz testing his knowledge of the duties and training required of various occupations. The score on this quiz is reported for nearly every respondent, therefore the sample selection problem is avoided. More importantly, IQ and the score on this quiz (KWW) are quite highly correlated. Also, Griliches (1976) finds that KWW performs as well as IQ in the determination of wage rates.

A wage equation is estimated using KWW, rather than IQ, as the ability measure. The use of KWW does not require the elimination of any of the observations previously used. The findings are reported in column 2 of Table 6. Most coefficients are largely unchanged, except for schooling. The KWW coefficient is positive and significant. The firm size effect is positive, significant, and virtually the same as estimated before. This does not support the hypothesis that large firms pay more simply because they hire workers of greater ability.

There may be some additional difficulties with this evidence, however. Neither IQ nor KWW are perfect measures of innate ability, i.e., they are measured with error. This errors-in-variables problem can cause biased coefficient estimates to remain even after the inclusion of the imperfectly measured variable. In spite of this, inclusion of the proxy variable is likely to reduce the bias in the other coefficients. The fact that the firm size coefficient did not change suggests the initial bias due to the
Table 6

Coefficient Estimates of Logarithmic Wage Equations,
Production Workers in Manufacturing,
NLS Young Men,
Measures of Ability Included

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.4156</td>
<td>.2038</td>
</tr>
<tr>
<td></td>
<td>(2.53)</td>
<td>(2.33)</td>
</tr>
<tr>
<td>Schooling</td>
<td>.0263</td>
<td>.0345</td>
</tr>
<tr>
<td></td>
<td>(2.13)</td>
<td>(4.90)</td>
</tr>
<tr>
<td>Experience</td>
<td>.0269</td>
<td>.0179</td>
</tr>
<tr>
<td></td>
<td>(6.17)</td>
<td>(5.26)</td>
</tr>
<tr>
<td>Tenure</td>
<td>.0047</td>
<td>.0137</td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(2.77)</td>
</tr>
<tr>
<td>IQ(^b)</td>
<td>-.00007</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td></td>
</tr>
<tr>
<td>KWW(^c)</td>
<td>---</td>
<td>.0041</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.08)</td>
</tr>
<tr>
<td>Union Member</td>
<td>.1527</td>
<td>.1355</td>
</tr>
<tr>
<td></td>
<td>(5.99)</td>
<td>(6.73)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>.0019</td>
<td>.0020</td>
</tr>
<tr>
<td></td>
<td>(2.33)</td>
<td>(3.20)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.0013</td>
<td>.0015</td>
</tr>
<tr>
<td></td>
<td>(2.38)</td>
<td>(3.58)</td>
</tr>
<tr>
<td>Dummy Variable</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>(^2) R</td>
<td>.412</td>
<td>.497</td>
</tr>
<tr>
<td>Sample Size</td>
<td>428</td>
<td>693</td>
</tr>
</tbody>
</table>

\(^a\) Variables defined in Table 1 and below

\(^b\) IQ score recorded on NLS survey. Its mean is 94.22 with a standard deviation of 14.37.

\(^c\) Score on the "knowledge of the world of work" questionnaire. The mean value is 31.48 with a standard deviation of 8.14.

Source: 1969 National Longitudinal Surveys
exclusion of ability is unimportant.

A hypothesis, which earlier authors mentioned, is that large firms pay more because the working conditions they offer are less pleasant. The premium paid is merely a compensating differential. With most data sources, this is a very difficult supposition to test directly. The NLS is no exception. There are variables available on the NLS which provide the respondent's evaluation of the pleasantness of the job, but this does not reveal the market's evaluation. It is the market's evaluation that is relevant to the theory of compensating differentials. This information is not easy to obtain, which makes any direct test prohibitively costly. However, through indirect means, it will be shown that this hypothesis is inconsistent with some evidence presented in Chapter IV.

The monopoly wage hypothesis was discussed earlier and rejected as an explanation of the firm size effect. An offshoot of the monopoly wage theory, which others have put forth as a reason for large firms paying higher wages, is the union threat effect model. The basis of this model is that firms might be willing to pay their nonunion workers a premium to discourage them from becoming unionized. Workers in large firms are more likely to join unions because it is probably cheaper for unions to organize them. Therefore, large firms will be more likely to pay such a "discouraging" premium. However, this implies that larger firms pay this wage premium only to nonunion workers. There is no need to pay union workers an additional
premium since they are already members of a union.

Tables 7 and 8 show estimates of wage equations for union
and nonunion workers on both the NLS Young Men and Older Men data
files. Column 1 of both Tables 7 and 8, are for union workers
of the Young Men and Older Men surveys respectively. Column 2
of Tables 7 and 8 are for nonunion workers. The evidence tends to
support the theory for the Young Men data set. The firm size
coefficient is positive and significant for nonunion workers,
but it is not significantly different from zero for union members.
This is not the case with the Older Men survey, however. Firm
size has a positive and significant effect on the wages of both
union and nonunion workers. The coefficient is larger for the
nonunion sample, but this is not entirely unforseen. There is
evidence in the literature which shows unions tend to lower the
absolute size of many regression coefficients in wage equations.
The firm size variable may be influenced by this effect. In any
case, the threat model does not explain why older union workers
receive wage premiums for working in large firms, but young union
workers do not. Perhaps a more complete version of this model can,
but this is not pursued here.

Each of the hypotheses considered in this chapter seem to
be inadequate explanations of why large firms pay wage premiums.
None is supported by further tests of their implications. The
evidence at this point suggests all workers, of all age groups,
are subject to the effect of firm size.
Table 7

Coefficient Estimates of Logarithmic Wage Equations,
Production Workers in Manufacturing,
NLS Young Men, by Union Status\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th>Union</th>
<th>Nonunion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Constant</td>
<td>.3230</td>
<td>.2217</td>
</tr>
<tr>
<td></td>
<td>(2.37)</td>
<td>(1.90)</td>
</tr>
<tr>
<td>Schooling</td>
<td>.0427</td>
<td>.0450</td>
</tr>
<tr>
<td></td>
<td>(4.22)</td>
<td>(5.29)</td>
</tr>
<tr>
<td>Experience</td>
<td>.0158</td>
<td>.0249</td>
</tr>
<tr>
<td></td>
<td>(3.21)</td>
<td>(5.65)</td>
</tr>
<tr>
<td>Tenure</td>
<td>.0076</td>
<td>.0205</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>(2.76)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>.0034</td>
<td>.0018</td>
</tr>
<tr>
<td></td>
<td>(3.25)</td>
<td>(2.22)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.00026</td>
<td>.0020</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(3.56)</td>
</tr>
<tr>
<td>Dummy Variables</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.320</td>
<td>.416</td>
</tr>
<tr>
<td>Sample Size</td>
<td>293</td>
<td>400</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Variables defined in Table 1

Source: 1969 National Longitudinal Surveys
Table 8

Coefficient Estimates of Logarithmic Wage Equations,
Production Workers in Manufacturing,
NLS Older Men, by Union Status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Union (1)</th>
<th>Nonunion (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.090</td>
<td>.7180</td>
</tr>
<tr>
<td></td>
<td>(6.72)</td>
<td>(3.34)</td>
</tr>
<tr>
<td>Schooling</td>
<td>.0082</td>
<td>.0210</td>
</tr>
<tr>
<td></td>
<td>(1.42)</td>
<td>(3.12)</td>
</tr>
<tr>
<td>Experience</td>
<td>-.0027</td>
<td>-.0017</td>
</tr>
<tr>
<td></td>
<td>(0.95)</td>
<td>(0.46)</td>
</tr>
<tr>
<td>Tenure</td>
<td>.0031</td>
<td>.0075</td>
</tr>
<tr>
<td></td>
<td>(2.63)</td>
<td>(5.52)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>-.0003</td>
<td>.0011</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.0013</td>
<td>.0028</td>
</tr>
<tr>
<td></td>
<td>(2.34)</td>
<td>(3.84)</td>
</tr>
<tr>
<td>Dummy Variables</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.184</td>
<td>.654</td>
</tr>
<tr>
<td>Sample Size</td>
<td>464</td>
<td>304</td>
</tr>
</tbody>
</table>

*Variables defined in Table 1

Source: National Longitudinal Surveys
Chapter III

THE MODEL

This chapter presents the details of a model which is consistent with the empirical evidence examined thus far. It is based on the desire firms have to evaluate the abilities of their workers and the difficulties large firms have in doing so. In the absence of an inexpensive means of evaluating workers, large firms may pay a premium to insure themselves of a high quality labor force.

The recent signaling and screening literature deals with the problem of determining who the most able are. Schooling is focused on as a possible "signal" or "screen." The signaling scenario is as follows: if it is less costly for high ability individuals to obtain schooling, firms can structure wages to give incentive so that only high ability people will acquire schooling. Schooling becomes a signal of high ability. The screening model is quite similar. The costs of schooling are the same for all, but ability is revealed after the completion of schooling and the worker is paid accordingly. Therefore, only the most able choose to acquire schooling. In both models, self-selection occurs. Incentives are such that workers voluntarily sort themselves by ability.
Guasch and Weiss (1981) model these type of self-selection devices in another context. They consider the profitability of offering workers the opportunity to have their ability tested. A fee is charged to each worker who takes the test. Those who pass receive a higher wage than those who do not pass. Workers who believe they have a high probability of passing choose to take the test and the others decline the opportunity. Guasch and Weiss determine the conditions under which this self-selection mechanism will accurately sort workers by ability. They find that it has serious limitations in a world where workers are risk averse, face different discount rates, have differing amounts of wealth, and where tests are imperfect indicators of ability. The same limitations apply to schooling as a sorting device.

In view of the above considerations, a different approach is taken here. It is assumed that self-selection devices do not sort workers perfectly, thus a substantial variation in ability remains for a given set of observable characteristics, such as schooling, experience, etc. The firm accomplishes further evaluation of its workers through direct monitoring. Workers are not directly charged for being monitored, nor does monitoring serve as a self-selection mechanism. If the firm chooses not to evaluate its labor force further, it can insure the retention of its highest ability workers, as well as all workers of lower ability, by paying everyone the opportunity wage of the most able. Alternatively, it can save on its wage bill by attempting to evaluate workers' abilities and paying each his opportunity wage. It is
shown below that the premium required to maintain the ability level of the firm's labor force falls, the more accurate the information about workers is. Information is costly, however. It is argued that large firms, due to hierarchical "loss of control", incur higher costs of evaluating workers. Minimization of the sum of wage and evaluation costs implies large firms rely less on accurate evaluation and more on wage premiums to maintain labor force quality. The resulting cost function of this minimization is placed in a model of profit maximization. This allows us to see the conditions necessary for large firms to be viable, even with higher labor costs. The details of this approach are spelled out in the sections below.

**The Wage Offer Function**

Suppose the firm faces a group of observationally equivalent workers, in terms of education, experience, etc., whose differences in ability are not readily discernable. Let this ability be denoted $y$ and its density function $f(y)$. Assume $f(y)$ is a normal distribution with mean $\bar{y}$ and variance $\sigma^2_y$. Each individual has an opportunity wage, $w$, which is an increasing function of $y$. Those with more ability have a higher opportunity wage. It is assumed that individuals have the opportunity to work in a job where ability is readily revealed. This sort of alternative job could be of several types: self-employment, working for piece rates, or employment in a small firm where productivity is easily observed by the owner who works side-by-side with the employees. Potential earnings in this type of employment are likely to play an important
role in determining the wage which will be accepted elsewhere. Thus, it is an alternative which must be met to attract workers. For simplicity, let the opportunity wage function be \( w = y \). Assume firms know the distribution of \( y \), and therefore of \( w \).

Workers enter firms, their ability is evaluated, and the firm makes a wage offer based on that evaluation. Denote this evaluation by \( q \), where \( q = y + u \), and \( u \sim N(0, \sigma_u^2) \), \( \text{cov}(y, u) = 0 \). \( q \) is an unbiased indicator of \( y \) in the sense that \( E(q|y) = y \). It follows that \( q \sim h(q) \), where \( h(q) \) is a normal distribution with mean \( \bar{y} \) and variance \( \sigma_q^2 = \sigma_y^2 + \sigma_u^2 \). The conditional distribution of \( y \), given \( q \), is as follows: \(^1^9\)

\[
y|q \sim N(\bar{y} + \gamma(q - \bar{y}), (1 - \gamma)\sigma_y^2), \quad \text{where} \quad \gamma = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_u^2}.
\]  

(1)

If \( q \) is a perfect indicator of ability, \( \sigma_u = 0 \), \( \gamma = 1 \), and \( q = y \). The distribution of \( y|q \) is degenerate at \( y \). Above, it is assumed that employment exists where this occurs. Cost minimization by these types of firms implies paying workers \( y \). This sets the opportunity wage. If \( q \) is useless, \( \sigma_u = \infty \), \( \gamma = 0 \), and \( f(y|q) = f(y) \). In this case, \( q \) yields no information about \( y \). This is the pure adverse selection case. It is discussed in a labor market setting by Weiss (1980) and, in another context, by Leland (1979). The remuneration scheme that results is very different from that of the perfect information case.

There are likely to be intermediate cases where information is not useless, but not perfect either. In many instances, acquiring perfect information is prohibitively costly, but it may be worthwhile to obtain some information. The object here is to characterize the
cost minimizing wage offer function for each information level. The
index of information will be \( \gamma \), since it is zero if the evaluation
is useless \( (\sigma_u = \infty) \) and one if \( q \) is without error \( (\sigma_u = 0) \). To
carry out this minimization, we must find an expression for the
wage function. Also, because the optimization occurs for each \( \gamma \) and
also for a given level of ability of the firm's labor force, we
need an expression for the ability function.

Firms observe \( q \), not \( y \), so they must base the wage offer on \( q \).
Thus, the wage offer will be a function dependent on \( q \), \( w(q) \).
Workers will accept the job offer only if \( y \leq w(q) \). From (1),
there is a distribution of \( y \) for each \( q \). Consequently, for each \( q \),
the probability a worker accepts the job is:

\[
\Pr \{ y \leq w(q) | q \} = \int_{L}^{w(q)} f(y|q) dy = F(w(q)|q),
\]

(2)

where \( \Pr \{ : \} \) denotes probability, and \( L = \) the lower bound of the
normal distribution = \(-\infty\). The average ability of those who accept
jobs, for each \( q \), is:

\[
E(y|q, y \leq w(q)) = \frac{\int_{L}^{w(q)} y f(y|q) dy}{\int_{L}^{w(q)} f(y|q) dy}
= \bar{y} + \gamma(q - \bar{y}) - (1 - \gamma) \sigma^2 \frac{f(w(q)|q)}{\int_{L}^{w(q)} f(y|q) dy}
\]

(3)

where \( E \) is the expected value operator. Also, for each \( q \), the
average wage is simply \( w(q) \).

The average ability and average wage for the firm's entire
labor force is found by aggregating over \( q \). Average ability is:

\[
A = \frac{\int_{L}^{T} \left( \bar{y} + \gamma(q - \bar{y}) - (1 - \gamma) \sigma^2 \frac{f(w(q)|q)}{\int_{L}^{w(q)} f(y|q) dy} \right) Fh(q) dq}{\int_{L}^{T} Fh(q) dq},
\]

(4)
where $T$ = the upper bound of the normal distribution = $\infty$. The average wage for the whole labor force is:

\[ W = \frac{\int_{-T}^{T} w(q) F h(q) dq}{\int_{-T}^{T} F h(q) dq}. \]  

(5)

The firm's problem, at this stage, is to minimize the wage per worker, $W$, for each $\gamma$ and subject to the average ability constraint.

Formally, the firm will minimize (5) subject to (4).

Before proceeding, note that $A$ can be rewritten as $\int_{-T}^{T} A e^{\frac{q}{t}} dq = \frac{A}{t} e^{t} = A$ and $\lim_{t \to T} \int_{-T}^{T} A e^{\frac{q}{t}} dq = 0$. This implies the Hamiltonian for the problem can be written as:

\[ H = w(q) - \lambda (\bar{y} + \gamma (q - \bar{y}) - (1 - \gamma) \sigma_{y}^{2} \frac{f}{F}), \]  

(6)

where $\lambda$ is the multiplier on $A$.

A necessary condition for the minimization to hold is:

\[ H_{w} = \frac{\partial H}{\partial w} = 1 - \lambda (-1 - \gamma) \sigma_{y}^{2} \frac{\partial (f/F)}{\partial w} = 0. \]  

(7)

In the appendix, it is shown that:

\[ \frac{\partial (f/F)}{\partial w} = \frac{-f}{F} \frac{1}{(1 - \gamma) \sigma_{y}^{2}} [w - \bar{y} - \gamma (q - \bar{y}) + (1 - \gamma) \sigma_{y}^{2} \frac{f}{F}]. \]  

(8)

From the moments of the truncated normal, 22

\[ E(y|q, y \leq w) = \bar{y} + \gamma (q - \bar{y}) - (1 - \gamma) \sigma_{y}^{2} \frac{f}{F}, \]  

(9)

Therefore,

\[ H_{w} = 1 - \lambda \frac{f}{F} (w - \bar{y} - \gamma (q - \bar{y}) + (1 - \gamma) \sigma_{y}^{2} \frac{f}{F}) = \]
Note that \( w - E(y|q, y \leq w) > 0 \) because \( w \) is the upper truncation point.

Now consider the term:

\[
H_{ww} = \frac{\partial^2 H}{\partial w^2} = -\lambda \left[ \frac{f}{F}(1 + (1 - \gamma)\sigma_y^2 \frac{\partial(f/F)}{\partial w}) + (w - y - \gamma(q - y) + (1 - \gamma)\sigma_y^2) \frac{\partial(f/F)}{\partial w} \right].
\]

Further results in the appendix show that the term in brackets simplifies to:

\[
\frac{f}{F} \frac{1}{(1 - \gamma)^2 \sigma_y^2} \left[ \text{Var}(y|q, y \leq w) - (w - E(y|q, y \leq w))^2 \right],
\]

where \( \text{Var} \) denoted variance. This term is negative because \( \text{Var}(y|q, y \leq w) \) is the average of squared deviations from the mean, while \( (w - E(y|q, y \leq w))^2 \) is the largest squared deviation from the positive side of the mean, since \( w \) is the upper truncation point. Hence, \( \text{Var}(y|q, y \leq w) < (w - E(y|q, y \leq w))^2 \). This implies \( H_{ww} > 0 \), as it should be for a minimum.

The optimal wage can be characterized by further manipulation of equation (10). Consider the "trajectory" of \( w \), or how \( w \) moves with \( q \). Using the implicit function theorem on (10),

\[
\frac{\partial w}{\partial q} = \frac{-H_{wq}}{H_{ww}} = \lambda \left[ \frac{f}{F}(-\gamma + (1 - \gamma)\sigma_y^2 \frac{\partial(f/F)}{\partial q}) + (w - E(y|q, y \leq w))\frac{\partial(f/F)}{\partial q} \right] \]

\[
-\lambda \left[ \frac{f}{F}(1 + (1 - \gamma)\sigma_y^2 \frac{\partial(f/F)}{\partial w}) + (w - E(y|q, y \leq w))\frac{\partial(f/F)}{\partial w} \right].
\]
Again relegating the proof to the appendix, we find:

\[
\frac{\partial(f/F)}{\partial q} = -\gamma \frac{\partial(f/F)}{\partial w}.
\]

(14)

Using this, we get \( H_{wq} = -\gamma H_{ww} \). Therefore,

\[
\frac{\partial w}{\partial q} = \frac{\gamma H_{ww}}{H_{ww}} = \gamma.
\]

(15)

This implies the optimal wage offer function is linear in \( q \): \( w = \alpha + \gamma q \), where \( \alpha \) may depend on various parameters. Wage offers move with \( q \) only at the rate of \( \gamma \). As the information firms have about workers worsens, \( \gamma \) decreases, therefore the rate of change of \( w \) with \( q \) is lower. This makes sense, because the less accurate the firm's evaluation is, the less it should rely on it. This result is used in Chapter IV.

A crucial conceptual experiment to consider is how \( w \) moves with \( \gamma \). Again using the implicit function theorem on (10):

\[
\frac{\partial w}{\partial \gamma} = \frac{-H_{w\gamma}}{H_{ww}} = \frac{\lambda[(y - \sigma_y^2 \frac{f}{F} + (1 - \gamma)\sigma_y^2 \frac{\partial(f/F)}{\partial \gamma} \frac{f}{F} + \frac{\partial(f/F)}{\partial \gamma}(w - E(y|q, y \leq w))]}{H_{ww}}.
\]

(16)

The appendix shows:

\[
\frac{\partial(f/F)}{\partial \gamma} = \frac{f}{F} \frac{1}{2(1 - \gamma)}\left[1 - \frac{\alpha + (1 - \gamma)\gamma}{(1 - \gamma)\sigma_y^2}(w - E(y|q, y \leq w))\right].
\]

(17)

Substituting this into the expression for \( H_{w\gamma} \) and simplifying, we get:
which is positive, because the term in brackets is negative, as shown above. Therefore:

\[
\frac{\partial w}{\partial y} = \frac{-H_{wy}}{H_{ww}} < 0. 
\]

This is a very important result. It states as information improves \((y \text{ increases})\), the wage offer to each individual, for all \(q\), falls. In other words the expected wage offer each worker receives will be smaller, the larger is \(y\). This also means that as the information the firm has worsens, the larger the premium needs to be to insure a given level of ability in the labor force. In terms of the wage function, this implies \(\alpha\) is a function of \(y\):

\[
w = \alpha(y) + yq \quad \text{and} \quad \frac{\partial w}{\partial y} = \frac{\partial \alpha}{\partial y} + q < 0. 
\]

So \(\partial \alpha / \partial y\) must be negative and larger in absolute value than \(q\).

Before closing this section, one final result can be discussed. We may want to know what happens to the wage schedule when the desired ability of the firm's labor force changes. To address this problem, consider the constraint equation (4). Due to the linearity of the wage offer function in \(q\), both \(f(w(q)|q)\) and \(F(w(q)|q)\) are independent of \(q\). Hence, the constraint can be simplified.

\[
A = \frac{F^T_L (\bar{y} + y(q - \bar{y}) - (1 - y)\sigma^2 y \frac{f}{F} h(q) dq}{F^T_L h(q) dq} = \bar{y} - (1 - y)\sigma^2 y \frac{f}{F}. 
\]
Define the function, \( J \), as:

\[
J = y - (1 - \gamma)\sigma^2_y \frac{f}{F} - A = 0
\]  

(22)

Therefore:

\[
\frac{\partial w}{\partial A} = \frac{-J_A}{J_w} = \frac{1}{-(1 - \gamma)\sigma^2_y \frac{\partial (f/F)}{\partial w}} = \frac{1}{f/F(w - E(y|q, y \leq w))} > 0. 
\]  

(23)

The wage offer increases with the desired level of ability in the labor force. The wage offer function must be of the form:

\[
w = \alpha(A, \gamma) + \gamma q, \text{ with } \frac{\partial \alpha}{\partial A} > 0.
\]  

(24)

This section has characterized the optimal wage offer function for a given amount of information the firm has. It is shown that the amount of information is important in determining the appropriate wage offer. Now, we would like to know which firms are likely to have poor information, and which will have good information.

The Evaluation Cost Function

Increasing the accuracy, \( \gamma \), of the evaluation is not costless. Here, the cost of evaluating workers is specified in the context of a hierarchical model of the firm. A substantial amount of work is found in the literature regarding these types of models. Williamson (1967) models the firm's choice of the number of hierarchical levels. He assumes the supervisee/supervisor ratio (span of control) is constant and there is a loss of productivity at each level. Calvo and Wellisz (1978) present a model in which the effort of workers at each level depends on the span of control, the wage, and the
effort at the level above. The firm chooses the span of control and the wage.

In both papers, supervision is utilized to prevent workers from shirking. However, it is plausible that supervision is also used to evaluate workers' abilities. Thus, a simple hierarchical model will be applied to this notion.

The nature of these models is that there is a residual income claimant, or president of the firm, who is ultimately responsible for all decisions. Thus, the president is responsible for the evaluation and remuneration of all members of the work force, including the production level workers. The information he obtains about each person is less accurate, as the number of individuals he personally supervises increases. Less time can be spent with each one and workers can perform well only when being monitored. This suggests accuracy is a decreasing function of the number supervised.

Because of this, for enterprises of much size, it is likely to be too ineffective for the president to personally monitor and evaluate each employee. Hence, the president will hire people to supervise for him. The president then supervises only the supervisors. Of course, the same limitations apply to the supervisors. The more individuals they attempt to monitor, the less accurate will be the information they obtain. In addition, theories of hierarchy postulate another source of information loss when hired managers replace the residual income claimant in part of the supervision. Because the president is further removed from direct evaluation of production workers, less accurate information may be
generated about them for a given span of control. This may be due to poor communication between the president and the managers about certain goals of the operation, or simply management error that the president does not correct. Also, it may be due to manager malfeasance. The more levels there are in the supervisory hierarchy, the further removed many managers are from the president and the more important these arguments are. Managers can more easily pursue their own goals at the expense of the firm's. This, along with the other considerations, implies the accuracy of information that reaches the president is a negative function of the number of levels in the hierarchy.

Formally, the accuracy of the evaluation, \( \gamma \), is a decreasing function of both: \( s = \text{supervisee/supervisor ratio} = \text{span of control} > 1 \) and \( n = \text{number of levels in the hierarchy} \). Assuming \( \gamma \) is a function only of \( s \) and \( n \), this can be written as:

\[
\gamma = M(s, n), \quad M_s < 0, \quad M_n < 0. \tag{25}
\]

Let \( i \) be an index of the level in the organization. It takes values from zero to \( n \), zero being the presidential level and \( n \) the production worker level. The number of employees at each level is \( s^i \), with \( s^n \) the number of production workers, assuming \( s \) is constant for all \( i \). The cost per worker, of maintaining a given level of precision, depends on the number of managers hired. The total number of managers hired is

\[
\sum_{i=0}^{n-1} s^i = \frac{s^n - 1}{s - 1}.
\]

The cost per worker of evaluation is an increasing function of

\[
\frac{s^n - 1}{s - 1}.
\]

Call this:

\[
V = V\left(\frac{s^n - 1}{s - 1}\right), \quad V' > 0 \tag{26}
\]
The firm will minimize evaluation costs, $V$, for each $\gamma$ and each organization size. The problem is to minimize, with respect to $s$ and $n$:

$$V = V\left(\frac{s^n - 1}{s - 1}\right)$$  \hspace{1cm} (27)

subject to $\gamma = M(n, s)$ and the organization size constraint, $x = s^n$. Solving the second constraint for $n$, one finds $n = \frac{\ln x}{\ln s}$.

The relevant Lagrangian is:

$$\min \ L = V\left(\frac{s^n - 1}{s - 1}\right) + \mu (\gamma - M(s, \frac{\ln x}{\ln s})),$$ \hspace{1cm} (28)

where $\mu$ is a Lagrange multiplier.

The interest lies in how the cost per worker of maintaining a given level of $\gamma$ responds to a change in the size of operation, $x$. Define $V^* = V\left(\frac{s^n - 1}{s - 1}\right) = V\left(\frac{x - 1}{s - 1}\right)$, where $s$ and $n$ are the optimal levels of $s$ and $n$ from the minimization problem. By the envelope theorem:

$$\frac{\partial V^*}{\partial x} = \frac{\partial L}{\partial x} = v' \cdot \left(\frac{1}{s - 1}\right) - \mu M \frac{1}{\ln s} > 0$$  \hspace{1cm} (29)

The cost per worker of maintaining a given level of $\gamma$ is higher for larger organizations.

This is not a surprising result; it is the nature of "loss of control." Loss of information increases as it is passed through more levels. To offset this loss, it will cost more, the more levels there are.

The general form of the evaluation cost function is $V = V(\gamma, x)$, with $V_\gamma > 0, V_x > 0$. For simplicity, let the
functional form be:
\[ V = \gamma V(x), \text{ with } V > 0, V_x > 0 \]  
(30)

**Labor Cost Minimization**

The above two elements of cost are brought together in this section which considers the cost minimizing choice of \( \gamma \). The firm's desire is to choose \( \gamma \) to minimize labor costs per worker for a given organization size and ability level of its labor force. Labor costs consist of wage and evaluation costs.

From equation (5), wage costs per worker are:
\[
W = \frac{\int_L^T w(q)h(q)dq}{\int_L^T h(q)dq} = \left( \alpha(A, \gamma) + \gamma q \right)h(q)dq = \\
\alpha(A, \gamma) + \gamma y, \tag{31}
\]

with \( \frac{\partial W}{\partial y} < 0 \) and \( \frac{\partial W}{\partial A} > 0 \). Evaluation costs per worker are given in equation (30). The firm will choose \( \gamma \) to:

Minimize \( C = W(\gamma, A) + \gamma V(x) \)  
(32)

for each \( x \) and each \( A \). Substituting \( x \) and \( A \) into the objective function, we find the first order condition is:
\[
\frac{\partial C}{\partial \gamma} = \frac{\partial W}{\partial \gamma} + V = 0 \tag{33}
\]

The second order condition is:
\[
\frac{\partial^2 C}{\partial \gamma^2} = \frac{\partial^2 W}{\partial \gamma^2} > 0 \tag{34}
\]

Differentiating (34) with respect to \( x \), we find:
\[
\frac{\partial y}{\partial x} = \frac{-V}{\frac{\partial^2 C}{\partial \gamma^2}} < 0 \tag{35}
\]
The larger the organization size, the less accurate evaluation it will undertake. Larger firms evaluate with less precision because it is more costly for them to do so. Furthermore, for each \( w(q) \),

\[
\frac{\partial w}{\partial x} = \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial x} > 0. 
\]  

The larger firm substitutes wage premiums for precise evaluation to maintain labor force quality. This implies the larger the enterprise, the larger the expected premium paid to each worker. In addition, it is consistent with the evidence from Chapter II.

**Profit Maximization**

The last section implies large firms will incur higher labor costs per worker. To see this differentiate the labor cost function, \( C \), with respect to firm size, \( x \). By the envelope theorem:

\[ \frac{\partial C}{\partial x} = \gamma \nu_x > 0. \]

One might ask how large firms can exist in competition with small firms if they have these higher costs. The answer is that there must be some offsetting benefit to large size that makes incurring these costs worthwhile. To illustrate, consider the following model.

Let \( g(x, A) \) be the firm's production function, where \( g_x > 0 \), \( g_A > 0 \). For each firm, profit can be expressed as:

\[ \Pi = g(x, A) - C(x, A)x, \]

where \( C \) is the optimally determined per unit labor cost and the price of output is normalized to one. Maximizing (38) with respect to \( x \) and \( A \), we find the following conditions:
\( \Pi_x = g_x - C_x X - C = 0 \)

\( \Pi_A = g_A - C_A X = 0 \)

(39)

and:

\[
D = \begin{vmatrix}
    g_{xx} - C_{xx} X - 2C_x & g_{xA} - C_{xA} X - C_A \\
    g_{Ax} - C_{Ax} X - C_A & g_{AA} - C_{AA} X
\end{vmatrix} > 0,
\]

\[
g_{xx} - C_{xx} X - 2C_x < 0, g_{AA} - C_{AA} X < 0.
\]

(40)

Suppose there is a parameter, \( \Delta \), that shifts the marginal product of labor outward such that \( g_{x\Delta} > 0 \). It could represent technological factors that account for greater scale economies in some industries, or a greater importance of team production that makes it more viable to bring together a large number of workers. The standard comparative statics exercise with respect to \( \Delta \) yields:

\[
\frac{\partial x}{\partial \Delta} = \frac{-g_{x\Delta} (g_{AA} - C_{AA} X)}{D} > 0
\]

\[
\frac{\partial A}{\partial \Delta} = \frac{-g_{x\Delta} (g_{Ax} - C_{Ax} X - C_A)}{D} < 0
\]

(41)

The more important these factors which \( \Delta \) represents are, the larger the firm. The larger firm bears higher labor costs, but it is made worthwhile by the economies it embraces. No prediction can be made as to how the average ability of the labor force moves with \( \Delta \).

Labor Market Equilibrium

The expected return to labor must be equal in all sectors for equilibrium to result. In the scenario of this chapter, larger
firms paying higher wages will attract a queue of workers attempting to secure high wage jobs. Equilibrium will occur when the queue is long enough so that the wage times the probability of employment is just equal, in utility terms, to the opportunity wage.

The queue will not be the same length for all workers. Consider the expected wage offer, given $y$:

$$ E(w|y) = E(\alpha + \gamma q|y) = \alpha + \gamma E(q|y) = \alpha + \gamma y. \quad (42) $$

The expected wage offer minus the opportunity wage, $y$, is a measure of the potential gain, $G$, to be had from employment in a large firm.

$$ G = E(w|y) - y = \alpha + (\gamma - 1)y \quad (43) $$

Note that:

$$ \frac{\partial G}{\partial y} = (\gamma - 1) \leq 0, \quad (44) $$

since $\gamma \leq 1$. High ability people have less to gain from employment in a large firm so they are willing to wait in a queue for a shorter period of time. If the firm did nothing to renew its pool of applicants, over time it would become heavily dominated by those of low ability. This is because workers of high ability will not wait as long for these jobs as low ability workers will. The firm is likely to take some action to guard against this, such as periodically discarding all previous applications (or requiring previous applicants to re-apply) and seeking a new set of applicants. This will insure the firm's pool of prospective employees is a sample of workers of all ability levels.
Chapter IV

ADDITIONAL EVIDENCE

The model presented in Chapter III is compatible with the simple observation that large firms pay higher wages. However, this is also true with the other explanations of the firm size effect which were not substantiated by the empirical examination in Chapter II. Thus, it is important to provide additional tests of the hypothesis presented here. Further implications of the model are derived and each is tested.

The Ability Gradient

An interesting result from Chapter III is given in equation (15). It shows that the rate of increase of the wage with the firm's evaluation, $q$, is equal to $\gamma$. For large firms, $\gamma$ is smaller, so this rate of increase should be smaller.

To state this formally, let:

$\ln(w_i) =$ the natural logarithm of the $i$th worker's wage,

$q_i =$ the firm's evaluation of the $i$th worker, and

$v_i =$ a vector of other characteristics of the $i$th worker.

Let the relationship between these variables be:

$$\ln(w_i) = a v_i + b q_i + \epsilon_i,$$  \hspace{1cm} (45)
where $\varepsilon_i$ is a random disturbance, uncorrelated with $v_i$ and $q_i$.

Suppose this equation is estimated for a sample of large firms. The estimate of $b$ should be smaller than the corresponding estimate from a sample of small firms, reflecting the fact that $q$ is a less precise indicator for large firms.\(^{27}\)

However, the estimation of equation (45) is not possible because we do not have the variable $q_i$, each firm's evaluation of their workers' abilities. Two measures of worker ability are available which are independent of the firm the worker is employed in: IQ and KWW. Let $z_i$ represent IQ or KWW for individual $i$. Consider the following regression equation for either of the aforementioned large firm or small firm samples:

$$\ln(w_i) = cv_i + dz_i + \varepsilon_i.$$  \hspace{1cm} (46)

The equation is misspecified since $z_i$ replaces $q_i$. The analysis of specification error shows the ordinary least squares estimate of $d$ is related to $b$ as follows:\(^{28}\)

$$E(\hat{d}) = \delta_{qz,v} b,$$  \hspace{1cm} (47)

where $\hat{d}$ is the estimate of $d$ and $\delta_{qz,v}$ is the regression coefficient of $q$ on $z$, holding constant $v$. We know that $b$ is smaller for large firms, but it is not immediately obvious how $\delta_{qz,v}$ varies with firm size.

The expression for such a regression coefficient is well known. Here, it is shown for the case where $v$ is a row vector, but can easily be generalized to the case where $v$ is a larger matrix. The expression is:
\[
\delta_{qz.v} = \frac{\sigma_y^2 Cov(q, z) - Cov(q, v)Cov(v, z)}{\sigma_y^2 \sigma_z^2 - (Cov(z, v))^2}.
\] (48)

Suppose \( z \) (IQ or KWW) is related to \( y \) as follows: \( z = y + e \), where \( e \sim N(0, \sigma_e^2) \) and \( Cov(e, y) = 0 = Cov(e, u) \). Thus,

\[
z \sim N(y, \sigma_y^2 + \sigma_e^2).
\]

We obtain the following, assuming \( e \) and \( u \) are also uncorrelated with \( v \):

\[
Cov(q, z) = E(q - y)(z - y) = \sigma_y^2
\]

\[
Cov(q, v) = E(q - y)(v - v) = Cov(y, v) \quad (49)
\]

\[
Cov(v, z) = E(v - v)(z - y) = Cov(v, y).
\]

Substituting relations (49) into (48), we find:

\[
\delta_{qz.v} = \frac{\sigma_y^2 \sigma_z^2 - (Cov(y, v))^2}{\sigma_v^2 (\sigma_y^2 + \sigma_e^2) - (Cov(y, v))^2}.
\] (50)

To investigate how (50) varies with firm size, recall that

\[
\gamma = \frac{\sigma_y^2}{(\sigma_y^2 + \sigma_u^2)}
\]

is smaller for large firms. When \( \sigma_u^2 \) is larger, \( \gamma \) is smaller. Thus, large firms have a smaller \( \gamma \) only because they have a larger \( \sigma_u^2 \). Notice that \( \delta_{qz.v} \) is independent of \( \sigma_u^2 \), i.e.,

\[
\frac{\partial \delta_{qz.v}}{\partial \sigma_u^2} = 0.
\]

Therefore, \( \delta_{qz.v} \) is independent of firm size.

Thus, under the assumptions that the errors in measurement of \( z \) and \( q \) are uncorrelated with each other, with ability, and with other determinants of wages, \( \hat{d} \) should be smaller for large firms. This is because \( E(\hat{d}) = \delta_{qz.v}b \), and \( b \) is smaller for large firms while \( \delta_{qz.v} \) is not dependent on firm size. This provides a test of the hypothesis of Chapter III, even though \( \hat{d} \) is still a biased estimate of \( b \).
This analysis is unchanged if \( z \) (IQ or KWW) is not simply a function of \( y \) and \( e \). Let \( z_i = k y_i + y_i + e_i \), where \( e_i \) is as defined before. In this case, 
\[
z \sim N(k \bar{y} + \bar{y}, k \sigma_y^2 + \sigma_e^2 + 2k \text{Cov}(y, v) + \sigma_e^2)
\]
and, under the same assumptions regarding errors of measurement:

\[
\begin{align*}
\text{Cov}(q, z) &= E(q - \bar{y})(k(v - \bar{v}) + y - \bar{y} + e) = k \text{Cov}(y, v) + \sigma_y^2, \\
\text{Cov}(q, v) &= E(q - \bar{y})(v - \bar{v}) = \text{Cov}(y, v), \\
\text{Cov}(v, z) &= E(v - \bar{v})(k(v - \bar{v}) + y - \bar{y} + e) = k \sigma_v^2 + \text{Cov}(y, v),
\end{align*}
\]}

(51)

with

\[
\delta_{qz,v} = \frac{\sigma_v^2(k \text{Cov}(y, v) + \sigma_y^2) - \text{Cov}(y, v)(k \sigma_v^2 + \text{Cov}(y, v))}{\sigma_v^2(k \sigma_v^2 + \sigma_y^2 + 2k \text{Cov}(y, v) + \sigma_e^2) - (k \sigma_v^2 + \text{Cov}(y, v))^2}
\]}

(52)

In this case, \( \delta_{qz,v} \) is still not dependent on \( \sigma_u^2 \) and thus is independent of firm size. Therefore, the above analysis holds: the estimated coefficient of \( \ln(w) \) on IQ or KWW should be smaller for large firms.

This is tested using the NLS Young Men data file because it is the only data set which includes IQ and/or KWW. Logarithmic wage equations are estimated with either IQ or KWW entered as a regressor, along with the interaction term of the ability measure times firm size. This analysis implies the coefficient on the interaction term should be negative, i.e., the effect of measured ability on wages should be less in large firms. Table 9 presents the results of the estimation of such a wage equation. Column 1 shows the regression coefficients with KWW as the ability measure. As expected, the interaction term of KWW times firm size is negative.
Table 9

Coefficient Estimates of Logarithmic Wage Equations, Production Workers in Manufacturing, NLS Young Men, Ability Interactions

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.0219</td>
<td>.1386</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.66)</td>
</tr>
<tr>
<td>Schooling</td>
<td>.0338</td>
<td>.0255</td>
</tr>
<tr>
<td></td>
<td>(4.84)</td>
<td>(2.07)</td>
</tr>
<tr>
<td>Experience</td>
<td>.0183</td>
<td>.0269</td>
</tr>
<tr>
<td></td>
<td>(5.42)</td>
<td>(6.19)</td>
</tr>
<tr>
<td>Tenure</td>
<td>.0140</td>
<td>.0047</td>
</tr>
<tr>
<td></td>
<td>(2.61)</td>
<td>(0.71)</td>
</tr>
<tr>
<td>IQ</td>
<td>---</td>
<td>.0030</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.72)</td>
</tr>
<tr>
<td>KWW</td>
<td>.0103</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>(4.36)</td>
<td></td>
</tr>
<tr>
<td>Union Member</td>
<td>.1354</td>
<td>.1507</td>
</tr>
<tr>
<td></td>
<td>(6.76)</td>
<td>(5.93)</td>
</tr>
<tr>
<td>Pct. Union</td>
<td>.0020</td>
<td>.0019</td>
</tr>
<tr>
<td></td>
<td>(3.23)</td>
<td>(2.25)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.0057</td>
<td>.0073</td>
</tr>
<tr>
<td></td>
<td>(4.10)</td>
<td>(2.50)</td>
</tr>
<tr>
<td>(IQ) x (PCT500)</td>
<td>---</td>
<td>-.00006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.09)</td>
</tr>
<tr>
<td>(KWW) x (PCT500)</td>
<td>-.00013</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>(3.16)</td>
<td></td>
</tr>
<tr>
<td>Dummy Variables</td>
<td>included</td>
<td>included</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.503</td>
<td>.416</td>
</tr>
<tr>
<td>Sample Size</td>
<td>693</td>
<td>428</td>
</tr>
</tbody>
</table>

Variables defined in Table 1

Source: 1969 National Longitudinal Surveys
and significant. Also, firm size and KWW are positive and significant by themselves. Column 2 displays the same general results with IQ used to measure ability. The interaction of IQ times firm size is negative and significant. These results are very supportive of the hypothesis presented here.

Wage Dispersion

The model has another ramification regarding the dispersion of wages. The variance of the wage rate each firm pays can be calculated if the wage offer function and the expected value of the wage are known. The former is given in equation (24) and the latter in equation (31). Thus, the variance of the wage is:

\[
\text{Var}(w) = E[(\alpha + \gamma q - \alpha - \gamma \bar{y})^2] = \gamma^2 E[(q - \bar{y})^2] = \gamma \sigma_q^2 = \gamma \left( \sigma_y^2 + \sigma_u^2 \right) = \gamma \sigma_y^2. \tag{53}
\]

Also,

\[
\frac{\partial \text{Var}(w)}{\partial \gamma} = \sigma_y^2 > 0. \tag{54}
\]

Firms with more accurate information about workers (a larger \( \gamma \)) have a larger wage dispersion. This may seem paradoxical, but it actually is not. Where information is poor, workers cannot be differentiated and their wages reflect it. Large firms choose a smaller \( \gamma \) so they should have a smaller wage dispersion. This implies the variation in actual wages around the regression line should be smaller in a wage equation estimated only for workers in large firms, compared to a regression only for workers in small firms. In other words, the standard error of the regression should be smaller for the large firm sample. However,
note that there is evidence which suggests unionism tends to reduce wage dispersion. Since large firms tend to be more unionized than small firms, part of the smaller dispersion of wages we might observe in large firms may be due to unionism.

This possibility is tested by estimating wage equations separately for union and nonunion workers. The standard errors of the regressions are reported in the upper part of Table 10 for the two NLS data sets and the CPS data. The standard error of the union wage regression is significantly smaller than that of the corresponding nonunion regression for all three data sets. Also, the simple standard deviation in wages is smaller in the union sample. Thus, to test hypotheses about wage dispersion in large and small firms, the sample should be divided by union status.

This is done by estimating wage equations for workers in large firms (PCT500 > 50), and for workers in small firms (PCT500 < 50), for union members and likewise for nonunion workers. The resulting standard errors are reported in the lower part of Table 10. In four of the six subsamples, the standard errors are significantly smaller for large firms. In the other two subsamples, the hypothesis of equal standard errors cannot be rejected. Neither is significantly larger than the other. The evidence regarding the simple standard deviation of wages is similar. The model predicts the standard error should be smaller for large firms in the majority of cases, but four of six is not an overwhelming majority. Nevertheless, this does suggest some support for the model.
Table 10

F-tests for Differences in Standard Deviations in Wages and Standard Errors of Wage Regressions, Production Workers in Manufacturing, CPS, NLS Young Men and Older Men

<table>
<thead>
<tr>
<th>Union</th>
<th>q_\text{w}^a</th>
<th>Std. Error^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.03</td>
<td>.2404</td>
<td></td>
</tr>
<tr>
<td>F-value</td>
<td>1.42</td>
<td>1.24*</td>
</tr>
</tbody>
</table>

Union
- Large Firm^c: .732 .1907
- Small Firm^d: .897 .2067
- F-value: 1.50* 1.17*

Nonunion
- Large Firm: .884 .2116
- Small Firm: 1.08 .2551
- F-value: 1.49* 1.45*

* Indicates significance at the 5% level.

a The standard deviation of wages in each subsample.

b The standard error of regression in each subsample.

c For individuals with PCT500 > 50.

d For individuals with PCT500 < 50.

Table 10 (continued)

<table>
<thead>
<tr>
<th></th>
<th>Older Men</th>
<th></th>
<th>CPS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma_w$</td>
<td>Std. Error</td>
<td>$\sigma_w$</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Union</td>
<td>.876</td>
<td>.2243</td>
<td>2.02</td>
<td>.4239</td>
</tr>
<tr>
<td>Nonunion</td>
<td>1.43</td>
<td>.2535</td>
<td>2.40</td>
<td>.5072</td>
</tr>
<tr>
<td>F-value</td>
<td>2.66*</td>
<td>1.27*</td>
<td>1.41*</td>
<td>1.43*</td>
</tr>
<tr>
<td>Union</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Firm</td>
<td>.763</td>
<td>.2013</td>
<td>2.07</td>
<td>.4311</td>
</tr>
<tr>
<td>Small Firm</td>
<td>1.02</td>
<td>.2465</td>
<td>1.95</td>
<td>.4134</td>
</tr>
<tr>
<td>F-value</td>
<td>1.78*</td>
<td>1.49*</td>
<td>.88</td>
<td>.92</td>
</tr>
<tr>
<td>Nonunion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Firm</td>
<td>1.60</td>
<td>.2629</td>
<td>2.29</td>
<td>.4547</td>
</tr>
<tr>
<td>Small Firm</td>
<td>1.17</td>
<td>.2504</td>
<td>2.44</td>
<td>.5385</td>
</tr>
<tr>
<td>F-value</td>
<td>.53</td>
<td>.90</td>
<td>1.14*</td>
<td>1.40*</td>
</tr>
</tbody>
</table>
A somewhat stronger test can be developed by combining the estimated standard errors for all subsamples. All three data sets can be combined to test whether or not the overall variation of wages is the same for large and small firms. One straightforward way of proceeding is to develop a likelihood ratio test. The null hypothesis is large firm and small firm standard errors, \( \sigma \), are the same within each data file and the union or nonunion sector. Formally,

\[
H_0: \sigma_{ijl} = \sigma_{ijs} \quad \text{for } i = i \text{ and } j = j
\]  

where \( i \) = data set (NLS Young Men, NLS Older Men, CPS), \( j \) = union, nonunion, and \( l \) = large firm, \( s \) = small firm. The alternative hypothesis does not restrict the standard errors to equality.

Under the alternative hypothesis, there is a standard error to estimate for each firm size/union status sector for all three data files. This makes a total of twelve. Under the null, the large firm and small firm standard errors are restricted to equality, thus the number of estimates needed drops to six. The likelihood ratio statistic, \( \lambda \), is given by:

\[
\lambda = \frac{L(H_0)}{L(H_1)} = \frac{12}{6} \frac{\hat{\sigma}^N}{\hat{\sigma}^m},
\]

where the \( \hat{\sigma}_k \)'s are the twelve estimated standard errors from the large firm and small firm regressions and the \( \hat{\sigma}_m \)'s are the estimated standard errors from combining the large firm and small firm samples. A representative \( \hat{\sigma}_k \) from the large firm sample is \((\hat{\epsilon}_l' \hat{\epsilon}_l / N_l)^{1/2}\) and
from the small firm sample is \( \left( \frac{\hat{\epsilon}_s' \hat{\epsilon}_s}{N_s} \right)^{\frac{1}{2}} \). \( \hat{\epsilon} \) is the residual from the wage regression. The corresponding \( \hat{\sigma}_m \) is 
\[
\hat{\sigma}_m = \left( \frac{\hat{\epsilon}_s' \hat{\epsilon}_s + \hat{\epsilon}_{\ell}' \hat{\epsilon}_{\ell}}{(N_s + N_{\ell})} \right)^{\frac{1}{2}}.
\]

Thus, \( \lambda \) can be calculated from the regression residuals. Doing so, it is found that \( \lambda = 4.0364 \times 10^{-10} \). A well known result is 
\[
-2\ln\lambda \sim \chi^2(r),
\]
where \( r = \text{degrees of freedom} = \text{the number of parameters restricted by the null hypothesis} \).\(^{34}\) In this case, \( r = 12 \). The one percent critical value for a chi-square distribution with twelve degrees of freedom is 26.2. If 
\[
-2\ln\lambda > 26.2,
\]
we can reject the null with 99% confidence. In this case, 
\[
-2\ln\lambda = 43.26 > 26.2.
\]
Thus, the hypothesis that large and small firm standard errors are equal is soundly rejected. Ideally, we want a statistic which tests whether small firm standard errors are greater than those of large firms, not simply unequal. Nevertheless, this test suggests a somewhat stronger support of the model.

Other Worker Characteristics

Chapter III considered the firm's problem of minimizing the sum of wage and evaluation costs per worker for each level of labor force ability, \( A \). Large firm size causes accurate evaluation to be more costly, therefore large firms substitute away from it. Under general conditions, this increase in the cost of accuracy will also raise the marginal cost of ability. Thus, if there are other qualities of workers which are good substitutes for ability and are easy to discern, large firms may hire disproportionately
more workers with these characteristics. Schooling may be such a quality. If so, large firms will tend to hire workers who have more education.

Similar implications might be drawn from the previously discussed screening models. If large firms have difficulty discerning their workers' abilities, they might rely on a characteristic, such as schooling, to sort workers by ability. This relies on the presumption that schooling and ability are highly correlated. Large firms hire educated workers only because they are of higher ability. Ability, not schooling, is an argument in the firm's production function. This is in contrast to the discussion in the previous paragraph. It presumes that schooling and ability are substitutes in production. Large firms may hire disproportionately more highly educated workers because education is easier to discern than ability. This can occur in the absence of a correlation between schooling and ability.

Both of the above arguments imply schooling and firm size should be positively correlated. Table 11 presents some evidence regarding this issue. In columns 1, 3, and 4, schooling is regressed on firm size. Firm size is positively related to schooling for all three data sets, a finding consistent with either of the above arguments.

The screening hypothesis implies that schooling is related to firm size only because schooling is a surrogate for ability. If this is the case, the correlation between schooling and firm size should vanish once ability is accounted for. This is not true for
Table 11

Coefficient Estimates of the Determinants of Schooling, Production Workers in Manufacturing, CPS, NLS Young Men and Older Men

<table>
<thead>
<tr>
<th></th>
<th>Young Men</th>
<th>Older Men</th>
<th>CPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Constant</td>
<td>10.317</td>
<td>7.806</td>
<td>7.745</td>
</tr>
<tr>
<td></td>
<td>(72.75)</td>
<td>(28.67)</td>
<td>(32.94)</td>
</tr>
<tr>
<td>PCT500</td>
<td>.0139</td>
<td>.0115</td>
<td>.0187</td>
</tr>
<tr>
<td></td>
<td>(5.18)</td>
<td>(4.55)</td>
<td>(4.47)</td>
</tr>
<tr>
<td>KWW</td>
<td>---</td>
<td>.0834</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(10.53)</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.037</td>
<td>.170</td>
<td>.025</td>
</tr>
<tr>
<td>Sample Size</td>
<td>693</td>
<td>693</td>
<td>768</td>
</tr>
</tbody>
</table>

*Variables defined in Table 1

the alternative model presented here. Schooling and firm size are correlated regardless of whether ability is held constant or not.

A simple test of these opposing hypotheses is conducted using the NLS Young Men data set. The KWW variable available on this data file is used as a measure of ability. Schooling is regressed on KWW and firm size. The results are shown in column 2 of Table 11. Firm size is still positively correlated to schooling, even after controlling for ability. This is not consistent with the screening model. It is supportive of the hypothesis here, but should be regarded as a preliminary result. Schooling has not been formally integrated into the model. Also, KWW may be a function of schooling, thus not exogenous.

Worker Productivity

Some authors who have investigated the firm size effect also examined the relationship between firm size and other variables. In particular, Miller (1978) and Bailey and Schwenk (1980) find large firms that pay higher wages also experience higher value added per employee and value of shipments per employee. They conclude that workers in these firms are more productive. Some attempt is made to control for worker skills when arriving at this conclusion. 35

This evidence is consistent with the scenario developed here. Large firms do not pay more because their workers are more productive. The causation is just the opposite. Workers in large firms are more productive because they are paid more. Larger firms encounter higher labor costs than do small firms, so it behooves them to hire less labor than they would have if
there were no evaluation costs. This moves them back up the marginal productivity schedule into the range where productivity is higher. Hence, one is likely to observe this higher productivity of workers in large firms, even if worker skill is adequately controlled for.
Chapter V

CONCLUSION

The empirical association between firm size and wage rates is demonstrated to be quite strong. Numerous authors, using a wide variety of data sets, have verified this relationship. The evidence presented here corroborates these findings. The correlation is shown to remain even after accounting for the standard earnings determinants - schooling, experience, unionism - and for worker ability, as measured by IQ and KWW. Furthermore, the analysis indicates the firm size effect is not due to union threat effects, a difference in payment schemes over the life-cycle, nor is it to compensate workers for lower fringe benefits.

The explanation pursued here is a variant of an adverse selection model. The less information one has, the more one must pay to insure quality. When applied to firms, this implies large firms pay more for workers because it is too costly for them to obtain better information.

This explanation is consistent with a variety of findings, including the fact that larger firms have more productive workers and a more educated labor force. In addition, it
correctly predicts that large firms have a relatively flat wage, ability gradient and a wage dispersion which is smaller.
APPENDIX

Properties of the Wage Offer Function

The proofs of several statements in Chapter III are outlined here. Recall that:

\[ f = \frac{1}{(2\pi (1 - \gamma)\sigma_y^2)^{1/2}} \exp\left(-\frac{1}{2} \frac{(w - \bar{y} - \gamma(q - \bar{y}))^2}{(1 - \gamma)\sigma_y^2}\right), \]

\[ F = \int_{-\infty}^{x} \frac{1}{(2\pi (1 - \gamma)\sigma_y^2)^{1/2}} \exp\left(-\frac{1}{2} \frac{y - \bar{y} - \gamma(q - \bar{y})}{(1 - \gamma)\sigma_y^2}\right) dy. \]

\[ \frac{\partial(f/F)}{\partial w} = \frac{\partial f/\partial w F - \partial F/\partial w f}{f^2} \]

\[ \frac{f}{F} \left[ \frac{-(w - \bar{y} - \gamma(q - \bar{y}))}{(1 - \gamma)\sigma_y^2} \right] fF - f^2 ]/F^2 = \]

\[ \frac{-f}{F} \frac{1}{(1 - \gamma)\sigma_y^2} \left[ w - (y + \gamma(q - \bar{y}) - (1 - \gamma)\sigma_y^2 \frac{f}{F} \right], \]

which proves equation (8).

To prove equation (12), the expression in brackets from equation (11) must be manipulated:

\[ \frac{f}{F} \left(1 + (1 - \gamma)\sigma_y^2 \frac{\partial(f/F)}{\partial w} \right) + (w - \bar{y} - \gamma(q - \bar{y}) + (1 - \gamma)\sigma_y^2 \frac{f}{F} \right) \frac{\partial(f/F)}{\partial w}. \]

Using the expression for \( \partial(f/F)/\partial w \), we obtain:

\[ \frac{f}{F} \frac{1}{(1 - \gamma)\sigma_y^2} \left[ (1 - \gamma)\sigma_y^2 \left(1 - \frac{f}{F}(w - y - \gamma(q - y)) - (1 - \gamma)\sigma_y^2 \frac{f^2}{F^2} \right) - \right. \]

63
\[(w - E(y|q, y \leq w))^2\]

From Johnson and Kotz (1970), the variance of \(y\) given \(q\), truncated at \(w\) is given by the first term in the brackets above. Thus, we have:

\[
\frac{f}{F} \left( \frac{1}{(1 - \gamma) \sigma^2_y} \left[ \text{Var}(y|q, y \leq w) - (w - E(y|q, y \leq w))^2 \right] \right),
\]

as is given in equation (12).

For equation (14), we must find:

\[
\frac{\partial (f/F)}{\partial q} = \frac{\partial f/\partial q \cdot F - \partial F/\partial q \cdot f}{f^2}
\]

Now:

\[
\frac{\partial f}{\partial q} = \gamma \left( w - \bar{y} - \gamma (q - \bar{y}) \right) \frac{f}{(1 - \gamma) \sigma^2_y}
\]

and

\[
\frac{\partial F}{\partial q} = -\gamma f = -\gamma \frac{\partial F}{\partial w}.
\]

Substituting in above, we have:

\[
\frac{\partial (f/F)}{\partial q} = -\gamma \left( \frac{\partial f/\partial w \cdot F - \partial F/\partial w \cdot f}{f^2} \right) = -\gamma \frac{\partial (f/F)}{\partial w},
\]

as shown in equation (14).

Equation (15) shows the wage offer function is \(w = \alpha + \gamma q\).

Using this,

\[
f = \frac{1}{(2\pi (1 - \gamma) \sigma^2_y)^\frac{1}{2}} \exp\left\{ -\frac{1}{2} \frac{(\alpha - (1 - \gamma) \bar{y})^2}{(1 - \gamma) \sigma^2_y} \right\},
\]

\[
F = \int_L^{\alpha + \gamma q} \frac{1}{(2\pi (1 - \gamma) \sigma^2_y)^\frac{1}{2}} \exp\left\{ -\frac{1}{2} \frac{(y - \bar{y} - \gamma (q - \bar{y}))^2}{(1 - \gamma) \sigma^2_y} \right\} dy
\]

It is readily seen that \(f\) is independent of \(q\); \(\partial f/\partial q = 0\). Also,

\[
\frac{\partial F}{\partial q} = \gamma f(\alpha + \gamma q) - \gamma f(\alpha + \gamma q) = 0,
\]
thus $F$ is independent of $q$. This proves the result referred to in footnote 25.

Equation (17) calls for evaluating the expression:

$$\frac{\partial (f/F)}{\partial \gamma} = \frac{\partial f}{\partial \gamma} \frac{F - \partial F/\partial \gamma f}{f^2}.$$

Now:

$$\frac{\partial f}{\partial \gamma} = f\left[\frac{(1 - \gamma)\sigma_y^2 - (\alpha + (1 - \gamma)y)(\alpha - (1 - \gamma)y)}{2(1 - \gamma)\sigma_y^2}\right], \text{ and}$$

$$\frac{\partial F}{\partial \gamma} = f\frac{(\alpha + (1 - \gamma)y)}{2(1 - \gamma)}.$$

Substituting in above, we obtain:

$$\frac{\partial (f/F)}{\partial \gamma} = \frac{\sigma_y^2 (1 - \gamma) - (\alpha - (1 - \gamma)y)(\alpha + (1 - \gamma)y)}{2(1 - \gamma)\sigma_y^2 f} \frac{f}{F}$$

$$= \frac{-(\alpha + (1 - \gamma)y)}{2(1 - \gamma)} \frac{f^2}{F^2} =$$

$$\frac{f}{F} \frac{1}{2(1 - \gamma)} \left[1 - \frac{(\alpha + (1 - \gamma)y)(\alpha - (1 - \gamma)y)}{(1 - \gamma)\sigma_y^2} \frac{f}{F} \right] =$$

$$\frac{f}{F} \frac{1}{2(1 - \gamma)} \left[1 - \frac{(\alpha + (1 - \gamma)y)(w - E(y|q, y \leq w))}{(1 - \gamma)\sigma_y^2} \right],$$

which is the result in equation (17).

Finally, this previous result is substituted into equation (16) to find an expression for $H_{\gamma \gamma}$. Doing so, we have:

$$\frac{f}{F}[y - \frac{\sigma_y^2}{F} f + \frac{\sigma_y^2}{F} f] \frac{1}{2} \left[(1 - \frac{(\alpha + (1 - \gamma)y)(w - E(y|q, y \leq w))}{(1 - \gamma)\sigma_y^2})ight]$$

$$+ \frac{1}{2(1 - \gamma)} \left(1 - \frac{(\alpha + (1 - \gamma)y)(w - E(y|q, y \leq w))}{(1 - \gamma)\sigma_y^2} \right)\left(w - E(y|q, y \leq w)\right)].$$
With some algebra, this becomes:

\[
\frac{f}{F} \frac{1}{2(1 - \gamma)} \frac{(\alpha + (1 - \gamma)\overline{y})}{(1 - \gamma)\sigma^2} \left( (1 - \gamma)\sigma^2 y (1 - \frac{f}{F}(w - E(y|q, y \leq w))) - (w - E(y|q, y \leq w))^2 \right).
\]

Again, from Johnson and Kotz (1970), the first term in brackets is \(\text{Var}(y|q, y \leq w)\). Thus,

\[
\frac{f}{F} \frac{(\alpha + (1 - \gamma)\overline{y})}{2(1 - \gamma)\sigma^2 y} \left[ \text{Var}(y|q, y \leq w) - (w - E(y|q, y \leq w))^2 \right],
\]

which is given in equation (18).
**FOOTNOTES**

1 See Haworth and Rasmussen (1971), Bailey, King, and Schwenk (1970), and Freeman and Medoff (1978). However, Parsons (1980) finds a positive effect of concentration on wages.

2 There is a data file of individual workers which includes the respondent’s estimate of the employment size of his/her workplace. This study was already underway before I became aware of its existence. Also, perhaps it is not superior to other data because workers may not have accurate knowledge of the number of employees at their place of employment.

3 Both NLS surveys were first conducted in 1966 with over 5,000 original respondents. The number declined in the ensuing years. The Older Men survey is for males aged 45 to 59 in 1966, and Young Men survey for males between the ages 14 and 24 in 1966.

4 The CPS is a labor force survey of over 50,000 individuals of all age and sex groups.

5 There are other published individual-based data sets, but the NLS and CPS were readily available to me. In general, the other data sets contain no more detail than the NLS.

6 Others have used the percent of an industry's workers employed in establishments of greater than x employees, where x is 250, 500, or 1,000. Donald Parsons suggested use of this variable to me. In addition, he made available to me this and other variables he obtained from the Survey of Manufactures.

7 Experience appears linearly, rather than as a quadratic, because the variation in experience within each sample is small, hence the square of experience does not increase explanatory power.
The statistic $(b_1 - b_2)/(s_1^2 + s_2^2)^{1/2}$, where $b_1$ is the regression coefficient in one equation, $b_2$ is the corresponding coefficient in the other equation, and $s_1^2$ and $s_2^2$ are the respective estimates of the variances, is distributed approximately as a $t$-distribution. It can be used to test the significance of $b_1 - b_2$. See Hoel (1971), pg. 265.

Tenure is not included as a regressor in Table 3 because it is not available from the CPS data.

For the equations in columns 2 and 3, $t=0.79$. For columns 2 and 4, $t=0.58$, and for columns 3 and 4, $t=0.32$.

Of course, tenure (or turnover) is likely to be jointly determined with wages, which means this inference may be more difficult to make.

See Maddala (1977), pgs. 159 - 161 and pgs. 304 - 305. His analysis shows when the proxy is a function only of the true variable and a random disturbance, the statement in the text holds with certainty. If the proxy is a function of other variables or other variables in the regression are also measured with error, it is possible for the bias not to be reduced.

Parsons' (1980) findings are similar to those of Tables 7 and 8. Note also that the division into union/nonunion subsamples does not effect the result that the effect of firm size is invariant across the two age groups. The hypothesis that the two firms size coefficients are equal cannot be rejected for the nonunion sample ($t=0.87$) of for the union sample ($t=1.18$).

See, for example, Bloch and Kuskin (1978).

Recent work on union/nonunion wage differentials (see Lee (1978)) focuses on the potential selectivity bias induced by a non-random determination of union membership. A model which corrected for this possible bias was estimated, but had no significant effects on the parameters of the wage equations, including the firm size coefficients. Therefore, it is not reported.
See Spence (1973) for one of the earliest discussions of this.

See Stiglitz (1975). Also, Riley (1979) surveys much of the signaling/screening literature and presents his own model.

This is generally in the spirit of a discussion by Stigler (1962).

This follows from the fact that \( y \) and \( q \) are distributed as a bivariate normal. See Hoel (1971), pgs. 153 - 154.

Johnson and Kotz (1970), pgs. 81 - 83, give a full treatment to the moments of the truncated normal distribution.

The Hamiltonian is written only for each \( q \). It simplifies to (6) after manipulating the constraint equation.


See the appendix.

This simplifies further since \( H_{\gamma y} = \frac{\alpha + (1 - \gamma)\gamma}{2(1 - \gamma)} H_{\gamma \gamma} \). Thus,

\[
-\frac{H_{\gamma y}}{H_{\gamma \gamma}} = \frac{-(\alpha + (1 - \gamma)\gamma)}{2(1 - \gamma)} < 0.
\]

This is shown in the appendix.

It is possible large size can result from the entrepreneur being a more efficient monitor, thereby having a lower cost function. This does not necessarily negate the argument that large firms have higher evaluation costs. It is possible that the entrepreneur of a large firm may increase his labor force so that the marginal cost of evaluation is greater than that of a small firm which has an inferior monitor, but fewer workers to supervise.

The coefficient, \( b \), in equation (45) will not equal \( \gamma \) because \( \ln(w) \), not \( w \), is the dependent variable in the equation. However, there is a one-to-one relationship between \( \gamma \) and \( b \), so when \( \gamma \) becomes larger, \( b \) becomes larger and vice-versa.

29. The metric of $z$ and $\sigma_e^2$ will not be the same for IQ and KWW.

30. These results cast doubt on the hypothesis that large firms pay more only to offset their nonpecuniary disadvantages. These findings indicate that, for this to hold, those with less ability require compensation for the unpleasant working conditions of a large firm, while those of higher ability do not. This seems doubtful.

31. Stigler (1962) also discusses wage dispersion in firms of different sizes.

32. See Freeman (1978).

33. Anderson (1958), pgs. 247-249, discusses a likelihood ratio test similar to the one used here.

34. See Maddala (1977), pgs. 179-180.

35. For example, Bailey and Schwenk hold constant occupation.
LIST OF REFERENCES


Lewis, H. Gregg, Unionism and Relative Wages in the United States, Chicago, 1963.


Parsons, Donald O., "Industrial Demand for Older Workers," preliminary report, June 1980.


