Earnings Functions and the Measurement of the Determinants of Wage Dispersion:
Extending Oaxaca's Approach

by

Joseph Deutsch

and

Jacques Silber

Department of Economics
Bar-Ilan University
52900 Ramat-Gan, Israel
email: jdeutsch@mail.biu.ac.il or silberj@mail.biu.ac.il

May 2002

1 The authors would like to thank Adi Gabbay, Yves Guillotin, participants at the joint meeting of CEPR
and the TSER Project (European Union) on Labour Demand, Education and the Dynamics of Social
Exclusion, held in Florence, Italy, in May 2001, and participants to the Fifth Euro-Mediterranean Meeting
held at the University of Nice-Sophia-Antipolis, France, in October 2001, for their very helpful comments.
I) Introduction

In a pathbreaking paper Oaxaca (1993) proposed a technique that allowed to decompose the relative wage gap between two population subgroups into two components, a first one measuring differences between the groups in human capital characteristics, a second one, labelled "discrimination", taking into account the impact of differences between the groups in the rates of return on these human capital characteristics. The goal of the present study is to extend Oaxaca's approach. While Oaxaca (1993) looked at the determinants of the wage gap between two groups, this paper not only extends the analysis to any number of groups, by showing that their approach amounts to analyzing the between groups wage differences, but also proposes a decomposition technique that permits to analyze the determinants of the overall wage dispersion. The approach presented here combines two techniques. The first one is popular in the field of income inequality measurement and concerns the breakdown of inequality by population subgroups. The second one, very common in the labor economics literature, uses Mincerian earnings functions to derive a decomposition of wage differences into components measuring respectively the impact of human capital, discrimination (differences in rates on return on human capital) and eventually unobservable characteristics. This methodological novelty allows one to determine the exact impact of each of these three elements on the overall wage dispersion, on the dispersion within and between groups and on the degree of overlap between the wage distributions of the various groups. This methodology is then applied to data obtained from income surveys conducted in Israel in 1982, 1990 and 1998, special emphasis being put on the comparison between the earnings of new immigrants and those of natives or older immigrants.

The paper is organized as follows. Section II reviews very briefly the literature on the determinants of wage inequality, the causes of the wage gap between natives and immigrants and the specificity of the immigration to Israel. Section III defines the mean difference of the logarithms of wages, indicates how it may be decomposed into between and within groups inequality and an overlapping component. Section IV explains how these decomposition techniques may be applied to Mincerian earnings functions to determine the respective contributions of human capital characteristics, rates of return on these characteristics and unobservable characteristics to the overall wage dispersion. Section V gives information on the data sets analyzed and presents the results of the estimation of earnings functions in Israel in 1982, 1990 and 1998. In section VI the results of the decomposition techniques suggested earlier are applied to the three income surveys previously mentioned while concluding comments are given in section VII.

II) On Wage Inequality and Immigration

A) The Determinants of Wage Inequality

---

2 A look at the Social Sciences Citation Index indicates that Oaxaca's paper has been cited more than 500 times. It should however be mentioned that in the same year as Oaxaca's paper was published, Blinder (1993) proposed quite a similar framework of analysis.
A vast literature has appeared in recent years dealing with the determinants of the increasing wage dispersion that has been observed in several Western countries during the past twenty years. Among the causes of this increasing inequality a distinction has usually been made between factors that affect the demand for labor, those that have an impact on the supply side and institutional changes that are likely to have also played a role.

There is a general agreement among economists that during the last quarter of the twentieth century there has been an important (positive) shift in the demand for high-skilled labor. The literature has offered two main explanations for this rise in the relative demand for skilled labor. The first argument stresses the role of increased trade openness that has been observed throughout the world during the 1980s and 1990s. This trend towards “globalization” that is usually explained by a decrease in transportation and communication costs and technology transfers implies that goods may be imported at a lower price. Since many of these goods are produced by low skilled labor, the increased degree of trade openness will, in developed countries, lead to a weaker demand for unskilled labor and hence a rise in the relative demand for skilled workers (see, for example Freeman, 1995, or Wood, 1995). Another type of explanation has emphasized the role of skill-biased technological change, a distinction being sometimes made between intensive and extensive skilled-biased technological change (see, Johnson, 1997, and Krueger, 1993).

One may also think of several factors that may affect the inequality of earnings via the supply side. The immigration of low skilled immigrants that has been observed in many Western countries is a first element to be taken into account. In most countries however the flow of immigrants does not represent an important addition to aggregate labor supply but the effect on local labor markets may still be important if immigrants tend to stay in specific areas (see, Topel, 1997, for some illustrations). As a whole, however, the net effect of immigration seems to be small, an additional reason being that the geographic mobility of natives tends to offset the impact of immigration on local labor markets.

Variations in the size of cohorts are another type of change that may have an effect on the supply side. A baby boom may thus lead, a generation later, to an important increase in the share of young cohorts in the labor force and since younger workers have lower wages than experienced workers, this may lead to an increase in overall inequality (see, for example, Welch, 1979, and Berger, 1985). Changes in the female labor force participation rates may also play an important role since younger female cohorts have low experience. They may however have a higher level of education. Another modification on the supply side that should indeed be mentioned is the continuous upgrading of the educational composition of the labor force in the Western world, a factor which leads to a decrease in the relative wage of educated workers and hence probably to a decrease in wage inequality.

Although a demand and supply framework can explain the rise in educational wage differentials by assuming that the rise in the relative demand for educated workers was stronger than that of their supply, some other factors of a more institutional nature should be taken into account. There may be laws that determine the minimum wage or overtime premia and thus affect wage inequality (see, Fortin and Lemieux, 1997). The extent of collective bargaining or the relative size of the public sector are other institutional elements which may play a role. The study of Goldin and Margo (1992) has indeed
clearly shown the impact of institutional change on wage inequality in the United States between 1935 and 1945. This short survey of the arguments put forth to explain the increase in wage dispersion observed in several Western countries during the past decades indicates that immigration could theoretically be an important factor, acting through the supply side but that the empirical evidence of a significant impact is not too abundant.

B) The Analysis of Wage Differences between Natives and Immigrants

Following Chiswick’s (1978) pioneering work, many studies tried to analyze how immigrants’ skills adapted to the host country’s labor market. For a long time the consensus was that at the time of their arrival immigrants earn less than natives because they lack the specific skills rewarded in the host country’s labor market. However as these skills are acquired, the human capital stock of immigrants grows relative to that of natives and immigrants experience faster wage growth (see, Borjas, 1994). There may even be a stage where immigrants have accumulated more human capital than natives, the argument being that there is a self selection process in so far as immigrants are “more able and more highly motivated” than natives (Chiswick, 1978). One has however to take into account the impact of changes in the wage structure since the latter is not likely to have similar effects on natives and immigrants. Thus in periods where rates of return to skills increase, the relative wage of immigrants may fall even if their skills remain constant (see, Levy and Murnane, 1992). Most of the studies looking at the earnings of immigrants refer however to countries where the annual flow of immigrants represents a small addition to the existing labor force. The case of Israel is different because at least twice during the past fifty years there have been periods of massive immigration, first during the late 1940s, then during the early 1990s.

C) On Immigration in Israel

In her comprehensive survey of the research conducted on immigration in Israel, Neuman (2000) indicates that between May 1948 and August 1951 the monthly number of immigrants was 15,000-20,000 so that in about three years the Jewish population which included 649,500 individuals when the State of Israel was created, doubled. Forty years later there was a massive influx of immigrants from the former Soviet Union. Thus in 1990 there were 199,516 immigrants and in 1991 176,000. Between the beginning of 1990 and the end of 1998, 879,486 immigrants were added to the Israeli population of 4.56 million, which corresponds to a growth rate of 19.3%. It should be stressed that these immigrants, most of them from the former Soviet Union, had an exceptionally high level of education. More than half of them had academic and managerial positions before immigration. The degree and speed of assimilation in the Israeli labor market of previous immigration waves has been analyzed in several studies (see, for example, Ofer, Vinokur and Bar-Chaim, 1980; Amir, 1993; Beenstock, 1993; Friedberg, 1995; Chiswick, 1997) while Eckstein and Weiss (1997) analyzed occupational convergence and wage growth of the recent large wave of immigrants from the former Soviet Union. Using panel data they found that upon arrival immigrants receive no significant return on imported human capital but with more time spent in Israel these returns increase, a gap remaining however
between the returns received by immigrants and natives. Ultimately immigrants receive
the same returns on experience but convergence is slow as is occupational convergence.
There have also been in the 1970s and 1980s studies looking at the wage differentials
between immigrants from various countries, a distinction being usually made between
Westerners (immigrants from Europe, America or Australia) and Easterners (immigrants
from North Africa or the Near East). While Weiss, Fishelson and Mark (1979) attributed
the decrease in the wage gap between Westerners and Easterners, that was observed in
the early 1970s, to a decrease in human capital differences, Amir (1980) argued that the
decrease in discrimination played a larger role.

D) Implications for the analysis of the link between immigration and wage
dispersion in Israel

The quick survey of the literature that has just been conducted should lead us to the
following predictions, assuming individuals are grouped by country or continent of
origin. First, given the size of the immigration waves during the past thirty years, we may
expect the kind of supply side effects mentioned earlier to be important. Second these
immigrants came mainly from the former Soviet Union and they had a relatively high
level of education. Therefore we do not expect to face the kind of issues mentioned by
Borjas (1985) concerning the eventual variation over time in the skills of the cohorts of
immigrants. Third it is likely that the skills of these immigrants, who came from a
country with a centralized economy, did not fit very well the requirements of the Israeli
labor market. This should imply that, if human capital (education or experience) is
measured only in years, the rate of return on this capital should be lower for the new
immigrants. For all these reasons (high level of human capital among immigrants, but
low rate of return on it) one may expect between groups gaps in wage to be more related
to differences in rates of return than in human capital. These gaps are also likely to have
grown over time as the share of new immigrants in the Israeli labor force became more
important.

It should be remembered however that the Israeli economy is a developed economy so
that the demand side effects mentioned earlier and affecting most Western countries must
have played also a role in Israel, this being particularly true for skilled-biased
technological change. Moreover it is well known that the Israeli economy was much
more opened to international trade in the 1980s and 1990s than in earlier periods. Both
factors lead us to predict that wage dispersion as a whole must have increased in Israel
between the early 1980s and the late 1990s. This trend is also likely to have taken place
within population sub-groups, if the latter are defined for example on the basis of the
country or continent of origin, and not as a function of human capital characteristics.

Finally, given that we expect that both the between and the within groups dispersion
increased over time, we cannot predict a priori what will happen to the degree of overlap
between the wage distributions of the various population sub-groups.

Despite the relatively important number of studies dealing with the impact of
immigration on the Israeli labor market, there has been no research attempting to quantify
the effect of immigration on income inequality in general, wage dispersion in particular.
The next sections will provide first a methodological framework allowing to better
estimate the impact of immigration on wage dispersion, second an empirical illustration
based on income surveys that have been conducted in 1982, 1990 and 1998.
III) The Decomposition of the Mean Difference of the Logarithms of Incomes by Population Subgroups:

Although the standard deviation is the most popular measure of the dispersion of a distribution, there exists another index of dispersion, called the mean difference MD, that is related to Gini’s famous concentration coefficient (see, Gini, 1912) and is defined (see, Kendall and Stuart, 1989) as

\[
MD = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j|
\]

where \(y_i\) and \(y_j\) are the incomes of individuals \(i\) and \(j\) and \(n\) the number of individuals in the population.

Such an index may also be used when the observations are the logarithms of incomes rather than the incomes themselves, in which case the mean difference, that will be denoted here as \(\Delta\), will be defined as

\[
\Delta = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |\ln y_i - \ln y_j|
\]

Expression (2) indicates in fact that \(\Delta\) measures the expected income gap, in percentage terms, between two individuals drawn at random (with repetition) from the sample of individuals on whom information on their income was collected.

Let now \(m\) represent the number of population subgroups. Expression (2) may then be decomposed into the sum of two terms, \(\Delta_A\) and \(\Delta_W\) where \(\Delta_A\) refers to what may be called the “across-groups inequality” (see, Dagum, 1960 and 1997) while \(\Delta_W\) measures the “within-groups inequality”, with

\[
\Delta_W = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{k \neq h} \sum_{j \in k} |\ln y_{ih} - \ln y_{jk}|
\]

and

\[
\Delta_A = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{k \neq h} \sum_{j \in k} |\ln y_{ih} - \ln y_{jk}|
\]

the second sub-index (\(h\) or \(k\)) in (3) and (4) referring to the group to which the individual belongs.

Let us now assume that the groups are ranked by decreasing values of the average of the logarithms of income in each group so that \(\ln y_{gh}\), the mean logarithm of incomes in group \(h\), is higher than \(\ln y_{g,h+1}\), the mean logarithm of incomes in group \((h+1)\).

Expression (4) may then be written as

\[
\Delta_A = \Delta_d + \Delta_p
\]

with

\[
\Delta_d = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{k \neq h} \sum_{j \in k} (\ln y_{ih} - \ln y_{jk}) \text{ with } \ln y_{ih} > \ln y_{jk}
\]

and

\[
\Delta_p = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{k \neq h} \sum_{j \in k} (\ln y_{jk} - \ln y_{ih}) \text{ with } \ln y_{ih} < \ln y_{jk}
\]

Combining (6) and (7) we derive that

\[
\Delta_d - \Delta_p = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{k \neq h} \sum_{j \in k} (\ln y_{ih} - \ln y_{jk})
\]

\[
\leftrightarrow \Delta_d - \Delta_p = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{k \neq h} \left[ \sum_{i \in h} (n_{k} \ln y_{ih}) - \sum_{j \in k} (n_{h} \ln y_{jk}) \right]
\]

\[
\leftrightarrow \Delta_d - \Delta_p = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{k \neq h} \left[ n_{k} n_{h} (\ln y_{gh} - \ln y_{gk}) \right]
\]

where \(n_h\) and \(n_k\) represent respectively the number of individuals in groups \(h\) and \(k\).

Since the between groups mean difference \(\Delta_B\) is obtained by giving each individual the average value of the logarithms of the incomes of the group to which he belongs, we may, using (2), define an index \(\Delta_B\) as

\[
\Delta_B = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{k \neq h} n_{h} n_{k} |\ln y_{gh} - \ln y_{gk}|
\]

\(^3\) \(y_{gh}\) is evidently the geometric mean of the incomes in group \(h\).
and it may be observed, when comparing (10) and (11), that
\[ \Delta_B = (\Delta_d - \Delta_p) \]  
(12)

Since expression (5) indicates that
\[ \Delta_A = (\Delta_d - \Delta_p) + (2 \Delta_p) \]  
(13)
we conclude, using (1), (2), (3), (4), (5), (12) and (13), that
\[ \Delta = \Delta_w + \Delta_B + (2 \Delta_p) \]  
(14)

One should note that expression (7) indicates that (2\Delta_p), the residual which is obtained in the traditional decomposition of the mean difference by population subgroups, is expressed as a simple function of the “transvariations” which exist between all pairs of population subgroups.

In the next section, the various decompositions that have been previously defined will be combined with Oaxaca’s (1973) traditional breakdown of income differences into a human capital and a discrimination component. This will allow us to analyze the impact of these two elements (the human capital characteristics and the rates of return on them) not only on the difference between the average logarithms of incomes in two population subgroups, but also on the dispersion of these incomes in each group and on the degree of overlapping between two distributions of incomes.

We will see that in some of these decompositions there will be another component that will represent the impact of unmeasured human capital characteristics.

IV) Estimating the contributions of human capital, rates of return (discrimination) and unobservable characteristics to the overall wage dispersion

The human capital model relates the earnings of an individual to the amount of his investment in human capital and using the framework of analysis originally put forth by Mincer (1972) we may write that
\[ \ln y_{ih} = \sum_{l=1}^{L} \beta_{lh} x_{lih} + u_{ih} \]  
(15)
where the subscripts i,h and l refer respectively to the individual (i), the group to which he belongs (h) and the human capital characteristic (l), the coefficients \( \beta_{lh} \) to the rate of return on human capital characteristic \( l \) in group \( h \) while \( x_{lih} \) represents the amount of that characteristic \( l \) possessed by individual \( i \) who belongs to group \( h \). Finally \( u_{ih} \) which is the residual of the regression refers evidently to factors that have not been taken into account.

Combining now expressions (3) and (15) we derive first the within groups inequality
\[ \Delta_w = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{j \in h} \left\{ (\sum_{l=1}^{L} \beta_{lh} x_{lih} + u_{ih}) - (\sum_{l=1}^{L} \beta_{lk} x_{ljh} + u_{jk}) \right\} \]
\[ = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{j \in h} \left\{ (\sum_{l=1}^{L} (\beta_{lh} x_{lih} - \beta_{lk} x_{ljh}) + (u_{ih} - u_{jh}) \right\} \]  

\[ ^4 \text{Following Gini (1959) we may say that there exists a “Transvariation” between two distributions \{x_i \} and \{y_j \} with respect to their (arithmetic, geometric,..) means } m_x \text{ and } m_y \text{ when among the } n_x \text{ and } n_y \text{ possible differences } (x_i - y_j), \text{ the sign of at least one of them is different from that of the expression } (m_x - m_y), \text{ and } n_x \text{ and } n_y \text{ being the number of observations in these two distributions. The importance of such a “Transvariation” may be measured in several ways (see, Deutsch and Silber, 1998). The reference here is to the moment } \mu_1 \text{ of order 1 which is defined as: } \mu_1 = \int_{-\infty}^{+\infty} g(y) \int_{-\infty}^{y} (y-x) f(x) \, dx \text{ where } g(y) \text{ and } f(x) \text{ are the densities of the variable } y \text{ and } x. \]

\[ ^5 \text{It may be observed that Oaxaca’s (1973) decomposition corresponds to what was defined earlier as the between groups mean difference } \Delta_B \text{ as it is defined in (11), for the specific case where there are only two groups } k \text{ and } h, \text{ assuming their size } n_k \text{ and } n_h \text{ are equal. (Oaxaca’s decomposition is in fact then equal to half the mean difference } \Delta_B \text{).} \]
\[
= \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{j \in h} 2 \left[ (\sum_{l=1}^{L} (\beta_{lh} x_{lih} - \beta_{lh} x_{ijh}) + (u_{ih} - u_{jh})) \right] \\
\text{with } \ln y_{ih} > \ln y_{jh} \text{ so that finally}
\]

\[\Delta_W = A + B\]  \hspace{1cm} (16)

where

\[A = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{j \in h} 2 \left[ (\sum_{l=1}^{L} (\beta_{lh} x_{lih} - \beta_{lh} x_{ijh})) \right] \text{with } \ln y_{ih} > \ln y_{jh}\]  \hspace{1cm} (17)

and

\[B = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{j \in h} 2 \left( u_{ih} - u_{jh} \right) \text{ with } \ln y_{ih} > \ln y_{jh}\]  \hspace{1cm} (18)

Expressions (16) to (18) indicate clearly that the within groups inequality of (the logarithms of) incomes is the sum of two elements: a first one (A) that is the consequence of differences between individuals belonging to the same group in the human capital characteristics they possess and a second one (B) that derives from differences between individuals belonging to the same group in unmeasured characteristics.

We may in a similar way derive an expression for the between groups inequality \(\Delta_B\) since by combining (11) and (15) we may write that

\[\Delta_B = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{k=1}^{m} \sum_{n_h n_k} \left[ (\sum_{l=1}^{L} (\beta_{lh} x_{lh} - \beta_{lk} x_{lk})) \right] \text{with } \ln y_{gh} > \ln y_{gk}\]  \hspace{1cm} (19)

where \(x_{lh}\) and \(x_{lk}\) are the arithmetic means of characteristic \(l\) in groups \(h\) and \(k\) respectively.

Remembering then that \((ab-cd)=((a+c)/2)(b-d) + ((b+d)/2)(a-c)\), we finally derive

\[\Delta_B = C + D\]  \hspace{1cm} (20)

where

\[C = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{k=1}^{m} \sum_{n_h n_k} \left[ (\sum_{l=1}^{L} (\beta_{lh} + \beta_{lk}) (x_{lh} - x_{lk})) \right] \]  \hspace{1cm} (21)

with \(\ln y_{gh} > \ln y_{gk}\)

and

\[D = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{k=1}^{m} \sum_{n_h n_k} \left[ (\sum_{l=1}^{L} (x_{lh} + x_{lk}) (\beta_{lh} - \beta_{lk})) \right] \]  \hspace{1cm} (22)

with \(\ln y_{gh} > \ln y_{gk}\)

Expressions (20) to (22) indicate that the between groups inequality of (the logarithms of) incomes is the sum of two elements: a first one (C) that is the consequence of differences between the groups in the average levels of human capital characteristics and a second one (D) that is explained by differences between the groups in the rates of return on these various human capital characteristics.\(^6\) Note that in (21) and (22) no reference is made to unmeasured characteristics since in measuring the between groups inequality we assume that each individual in a group receives the average (logarithm of the) income of the group and that, by the definition of a regression, this average does not include a residual.

The third element of the decomposition of inequality of (the logarithms of) incomes measures the degree of overlap between the distributions of the various groups and, combining (7) and (15), \(\Delta_p\) may be expressed as

\[\Delta_p = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{j \in k} \left[ (\sum_{l=1}^{L} (\beta_{lh} x_{lih} + u_{ih}) - (\sum_{l=1}^{L} (\beta_{lh} x_{lih} + u_{ih})) \right] \text{with } \ln y_{ih} < \ln y_{jk}\]  \hspace{1cm} (23)

Using similar decomposition rules as before, we derive that

\[\Delta_p = E + F + G\]  \hspace{1cm} (24)

where

\[E = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{j \in k} \left[ (\sum_{l=1}^{L} ((\beta_{lh} + \beta_{lh})/2) (x_{lj} - x_{lih})) \right] \]  \hspace{1cm} (25)

\(^6\) The decomposition given in expressions (20) to (22) corresponds to that proposed by Reimers (1983).
with \( \ln y_{ih} < \ln y_{jk} \)

\[
F = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{j \in k} \left[ \sum_{l=1}^{L} \left( \frac{x_{ljk} + x_{lih}}{2} \right) (\beta_{lk} - \beta_{lh}) \right] \tag{26}
\]

and \( G = \frac{1}{n^2} \sum_{h=1}^{m} \sum_{i \in h} \sum_{j \in k} \left[ (u_{jk} - u_{ih}) \right] \tag{27} \)

Expressions (24) to (27) indicate that the degree of overlapping between the distributions corresponding to the various groups is a function of three elements: a first component \( E \) that reflects differences in the human capital characteristics possessed by the individuals affected by the overlapping, a second element \( F \) that is explained by differences between the groups in the rates of return on these human capital characteristics and a third expression \( G \) that is due to unmeasured characteristics among the individuals affected by the overlapping.

Combining expressions (16), (20) and (24) we conclude that \( (A + C + E) \), \( (D + F) \) and \( (B + G) \) represent respectively the contributions of differences in human capital characteristics, in rates of return on these characteristics and in unobservable characteristics to the overall wage dispersion.

We now turn to the results of the empirical investigation.

V) Male Earnings Functions in Israel in 1982, 1990 and 1998:

The empirical illustration that will be presented here is based on the Income Surveys that are conducted each year in Israel. We have chosen to limit our analysis to three surveys: those of the years 1982, 1990 and 1998. Since one of the aims of this research is to look at the impact on earnings of the country of origin and of the period of immigration of the immigrants, we have limited our analysis to the Jewish male population and divided it in four groups: those born in Israel (group IL), those born in Asia or Africa (group AA) and those born in Europe or America. However in order to take into account what happened to the most recent immigrants we have divided the last group into two subgroups: those who immigrated to Israel before 1972 (group EA) and those who immigrated after 1971 (group NIM). It is clear that, specially for the last two surveys analyzed (1990 and 1998), most of the members of the last subgroup came from the former Soviet Union so that we will be able to focus on the earnings of this important population of immigrants.

Let us now take a look at the general characteristics of the population analyzed.

The two first columns of Tables 1-A to 1-C give for each year the means and standard deviations of the various variables that have been introduced in the regressions. The results are given each time for the whole sample. It appears that the proportion of married individuals declined over the years from 87% in 1982 to 75% in 1998. The proportion of singles on the contrary increased during the same period from 10.7% to 20.9%. The other categories of marital status (divorced, widow or separated) increased slightly from 2.3% in 1982 to 4.1% in 1998. The average number of years of schooling increased from 10.7 in 1982 to 12.6 in 1998 while the average number of years of

---

7 The corresponding statistics for each of the four subgroups that have been distinguished (IL, AA, EA and NIM) and for each of the years 1982, 1990 and 1998 may be obtained upon request from the authors.
experience correspondingly decreased from 26.2 years in 1982 to 21.7 in 1998. The proportion of the males having attended a Talmudic school (Yeshiva), a factor likely to have a downward effect on earnings, decreased from 2.6% in 1982 to 1.5% in 1998. It is difficult to compare the means obtained for the various population subgroups since the younger people are more likely to be individuals born in Israel so that one would expect to observe, for these men born in Israel, a smaller proportion of married individuals and less years of schooling or experience. This is actually the case, for each year. Such differences do not prevent us however to compare regression results because then the age is kept constant, since we have defined experience in the traditional way, that is as age minus six minus the number of years of schooling.

The last two columns of Tables 1-A to 1-C give, for the whole sample, the results of the Mincerian earnings functions that have been estimated for each of the three periods analyzed. It appears for example that the rate of return on schooling increased throughout the period, being equal to 6.8% in 1982, 7.5% in 1990 and 8.9% in 1998. The rate of return on experience at the beginning of the career showed a different pattern since it rose from 2.7% to 3.9% between 1982 and 1990 but was equal to 2.9% in 1998. Individuals who were married earned on average 20.8% more than those who were divorced, separated or widows in 1982, 10.7% more in 1990 and 15.5% more in 1998. Single individuals on the contrary earned 10.7% in 1982, 10.1% in 1990 and 4.1% less than those who were neither married nor singles. These data indicate therefore that the gap between married and single men decreased significantly between 1982 and 1998.

Similar regressions have been estimated for each of the four population subgroups that have been distinguished. It appears for example that in 1982 the rate of return on schooling was much higher for those born in Israel (8.8%) than for those born in Asia or Africa (4.9%) or Europe or America (6.0% for those who came before 1972 and 5.0% for those who came after 1971). Similarly in 1998 the rate of return on schooling was 12.5% for those born in Israel, 7.3% for those born in Asia or Africa, 11.0% for those born in Europe who immigrated before 1972 and 6.5% for those born in Europe who arrived in Israel after 1971.

VI) The components of the overall wage dispersion:

Table 2 gives for each of the three years 1982, 1990 and 1998 the decomposition of the overall wage dispersion into the three components mentioned in Section III: the between and within groups dispersions and the overlapping term. The number that appears in the line labeled “Total” gives for each year the overall mean difference of the logarithms of income, that is the expected income difference in percentage terms between two individuals chosen randomly (with repetition) in the sample. Whereas this mean difference only slightly increased between 1982 to 1990 (from 63.6% to 64.4%), the change was very important between 1990 and 1998 since the mean difference reached the

---

8 This implies that the years spent in the army are considered as part of the working experience.
9 Since this study is based on cross-sections, it is in fact impossible to make a distinction between the impact of experience and that of the business cycle and one has to be careful in interpreting some of the changes observed for example between 1982 and 1990.
10 The results of these estimations may be obtained upon request from the authors.
value of 73.0% in 1998. What are the reasons for such an important increase in the overall dispersion observed during the decade 1990-1998. This is a period where in several Western countries wage dispersion increased for reasons related to technological change, the increasing openness to trade and institutional change such as the weakening of the trade unions (see the short survey of the literature in section II). It should however be remembered that during the 1990-1998 period 880,000 individuals immigrated to Israel, mostly from the former Soviet Union. Since one of the population subgroups includes only those who migrated from Europe or America after 1971, the analysis presented in this section enables one to determine the impact of this immigration on the overall wage dispersion. However, the decomposition techniques presented in Section III to V give also the specific impact on the overall dispersion of incomes, and on its three components, the between and within groups dispersion and the overlapping element, of human capital characteristics, the rates of return on these characteristics and unobserved characteristics. All these results will now be presented and analyzed.

1. The relative importance of the between and within groups dispersion and the contribution of the overlapping component

Table 2 indicates that in absolute terms the contribution of the between groups dispersion to the overall dispersion decreased from 11.5% to 6.4% between 1982 and 1990 but it increased between 1990 and 1998 to reach 14.7% in 1998. In percentage terms the picture is similar since the contribution of the between groups dispersion decreased from 18.1% to 10.0% between 1982 and 1990 but was equal to 20.1% in 1998.

The within groups dispersion increased in absolute terms during both sub-periods. It was equal, in absolute terms, to 17.9% in 1982, 21.7% in 1990 and 26.5% in 1998. The picture is quite similar if one looks at the relative contribution of the within groups to the overall wage dispersion since this contribution rose from 28.1% in 1982 to 33.8% in 1990 and 36.3% in 1998.

For the overlapping term the pattern is as follows: in absolute terms it increased from 34.2% in 1982 to 36.2% in 1990 but fell down to a level of 31.8% in 1998. In relative terms the contribution of the overlapping term rose form 53.8% in 1982 to 56.2% in 1990 to go back to 43.6% in 1998.

The picture during the 1982-1990 period is hence very different form the one observed during the years 1990-1998. During the first period the between groups dispersion decreased while the within groups dispersion rose, the overlapping term increasing only slightly. These conclusions are true in absolute and relative terms. During the second sub-period, on the contrary the between as well as the within groups dispersion rose while the amount of overlapping decreased, this being again true in absolute and relative terms. Two factors at least may explain these patterns. First there was at that time an increase in wage dispersion in several Western countries and this is probably also true for the within groups dispersion. At the same time there is a specific Israeli story: the massive immigration of Jews from the former Soviet Union has increased, at least in a first stage, the degree of stratification in the Israeli society, leading thus to an increase in the between groups dispersion. This latter effect was more important than the increase in the within groups dispersion, that was just mentioned, since the degree of overlapping decreased during this period.
To better understand these changes we now take a look at the respective role played by human capital characteristics, the rates of return on them and by unobserved characteristics.

2. The contribution of human capital characteristics, the rates of return on these characteristics (“discrimination”) and unobserved characteristics to the wage dispersion

Table 3 indicates that in 1982 out of a total wage dispersion of 63.6%, human capital characteristics contributed in absolute terms 16.8%, rates of return on these characteristics 2.7% and unobserved characteristics 44.1%. The corresponding figures for 1990 when the overall dispersion was 64.4%, were 19.9%, 0.3% and 44.2%. In 1998 the mean difference of the logarithms of wages was equal to 73.0% while the three contributions previously mentioned were respectively equal to 21.4%, 4.8% and 46.7%. It appears therefore that over time the contribution of human capital increased in absolute value. The contribution of unobserved characteristics on the contrary did not vary very much over time while that of the rates of return on human capital characteristics was low and unstable.

The figures are somehow different in percentage terms (see again Table 3). It appears that over time there was also an increase in percentage terms in the contribution of human capital characteristics, at least between 1982 where it was equal to 26.4% and 1990 when it reached 30.9%. There was no important change during the 1990-1998 period. The relative contribution of unobserved characteristics decreased over time, mainly during the second sub-period (from 69.4% in 1982 to 68.6% in 1990 and 64.0% in 1998). Finally the relative contribution of subgroup differences in the rates of return on the human capital characteristics varied over time since it was equal to 4.2% in 1982, 0.5% in 1990 and 6.6% in 1998.

A similar analysis may be conducted at the level of each of the three components of the overall wage dispersion: the between and within groups dispersion and the overlapping term. The results are presented in table 4. For the between groups dispersion, as was mentioned previously, only human capital characteristics and the rates of return on them play a role. It appears that the relative role of human capital varied strongly over time: it was equal to 35.4% in 1982, 57.0% in 1990 but only 11.8% in 1998. The picture is evidently the opposite for the relative contribution of the rates of return on human capital characteristics, what is usually labeled discrimination. The very important role of the latter in 1998 indicates clearly that, ceteris paribus, in particular at comparable levels of schooling and experience, new immigrants receive a much lower rate of return on their human capital characteristics. This is confirmed by Appendix 1-C which indicates an important difference between two categories of individuals. The first group includes the new immigrants from Europe or America (mostly immigrants from the former Soviet Union) and those born in Asia or Africa and they had rates of return on schooling respectively equal to 6.5% and 7.3%. The second group is composed of individuals who came from Europe or America before 1972 or were born in Israel and they had rates of return on schooling respectively equal to 11.0% and 12.5%.

For the within groups dispersion only two factors (see Section IV) play a role: human capital and unobserved characteristics. Table 4 indicates that, in relative terms, the
contribution of human capital steadily rose over time, from 27.6% in 1982, to 32.3% in 1990 and 36.0% in 1998. The trend, in relative terms, is evidently opposite for unobservable characteristics.
For the overlapping component, as was mentioned in Section IV, each of the three factors (human capital and unobserved characteristics and rates of returns on human capital characteristics) plays a role. It is first interesting to note that the data indicate that the component measuring the impact of the rates of return on human capital had, each year, a negative contribution to the degree of overlap. This implies that if there had been no differences between the individuals involved in the overlap in the amount of their human capital or in their unobservable characteristics, the between groups differences in the rates of return on human capital would have led to a smaller amount of overlap.
As far as the two other components are concerned, it appears that the relative importance of human capital increased over time (from 22.8% in 1982 to 25.3% in 1990 and 31.9% in 1998). The relative contribution of unobservable characteristics on the contrary was rather unstable (91.1% in 1982, 81.5% in 1990 and 93.4% in 1998).

3. Summarizing the empirical results

The various observations that have just been made could be summarized as follows.
First during the two sub-periods that have been analyzed, the between groups dispersion first decreased, then increased; since the same pattern has been observed, in percentage terms, for the rates of return component of this between dispersion, which contributes most to this dispersion, we may fairly assume that rates of return on human capital characteristics played a central role here.
Second, the within groups dispersion increased in both periods, a pattern that is observed also, in percentage terms, for the human capital component of this type of dispersion. Although this component never represents more than a third of the within groups dispersion, it is likely that its variation over time explains the increasing importance of the within groups dispersion.
Third, the overlapping component first increased, then decreased. This is also the pattern observed for the rates of return, although their contribution remains negative throughout the period. We may therefore conjecture that the story of the overlap is mainly that of the rates of return and if the share of the overlap in the overall dispersion decreased drastically between 1990 and 1998, it seems to be a consequence of the fact that the sharp decrease in the rates of return on human capital observed among new immigrants led to a reduction in the amount of overlap between the income distributions of the four population subgroups.

VII) Concluding Comments

This paper proposes a new methodology for analyzing the respective impact of human capital characteristics, rates of return on these characteristics and unobservable characteristics on the overall wage dispersion. It also shows how to determine the role played by the between and within groups dispersion as well as by the degree of overlap between the groups’ wage distribution, when the individuals are also characterized by the groups to which they belong. An illustration based on income surveys conducted in Israel
in 1982, 1990 and 1998 indicated that the approach proposed here sheds some interesting light on the evolution of the wage dispersion over time. In particular it allowed us to check that most of the predictions made on the basis of either economic theory or institutional-historical knowledge were confirmed. We thus found that the overall wage dispersion as well as the between and within groups dispersion increased over time, that changes over time in the between groups dispersion and in the amount of overlap were mostly related to variations that occurred in the rates of return on human capital observed among the different subgroups while the within groups dispersion reflected more changes in the human capital characteristics.

Bibliography


Mincer, J., Schooling, Experience and Earnings, 1974, N.B.E.R.


Table 1-A
Descriptive statistics and regression results for the 1982 Income Survey
(whole population)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Regression Coefficients</th>
<th>t-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logarithm of wage per hour</td>
<td>4.0184</td>
<td>0.5723</td>
<td>0.2077</td>
<td>3.46</td>
</tr>
<tr>
<td>Married</td>
<td>0.8672</td>
<td>0.3394</td>
<td>0.0681</td>
<td>24.03</td>
</tr>
<tr>
<td>Single</td>
<td>0.1068</td>
<td>0.3088</td>
<td>-0.0179</td>
<td>-1.55</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>10.6754</td>
<td>3.7835</td>
<td>0.1079</td>
<td>6.66</td>
</tr>
<tr>
<td>Years of Experience</td>
<td>26.2035</td>
<td>14.7535</td>
<td>0.0024</td>
<td>9.66</td>
</tr>
<tr>
<td>Square of years of experience</td>
<td>904.2872</td>
<td>880.5513</td>
<td>-0.0004</td>
<td>-8.57</td>
</tr>
<tr>
<td>Attended Talmudic School</td>
<td>0.0257</td>
<td>0.1582</td>
<td>-0.3546</td>
<td>-5.85</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td>2.7680</td>
<td>33.76</td>
</tr>
<tr>
<td>R²</td>
<td>0.2564</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>2725</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1-B
Descriptive statistics and regression results for the 1990 Income Survey
(whole population)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Regression Coefficients</th>
<th>t-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logarithm of wage per hour</td>
<td>2.5072</td>
<td>0.5742</td>
<td>0.1070</td>
<td>2.20</td>
</tr>
<tr>
<td>Married</td>
<td>0.8140</td>
<td>0.3891</td>
<td>-0.0109</td>
<td>-1.78</td>
</tr>
<tr>
<td>Single</td>
<td>0.1523</td>
<td>0.3593</td>
<td>-0.0095</td>
<td>-1.42</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>11.6693</td>
<td>3.2367</td>
<td>0.0749</td>
<td>24.83</td>
</tr>
<tr>
<td>Years of Experience</td>
<td>23.3972</td>
<td>13.8600</td>
<td>0.0385</td>
<td>14.52</td>
</tr>
<tr>
<td>Square of years of experience</td>
<td>739.5288</td>
<td>787.2934</td>
<td>-0.0005</td>
<td>-12.07</td>
</tr>
<tr>
<td>Attended Talmudic School</td>
<td>0.0202</td>
<td>0.1408</td>
<td>-0.2695</td>
<td>-4.30</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td>1.0609</td>
<td>14.71</td>
</tr>
<tr>
<td>R²</td>
<td>0.2810</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>3113</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1-C
Descriptive statistics and regression results for the 1998 Income Survey
(whole population)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Regression Coefficients</th>
<th>t-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logarithm of wage per hour</td>
<td>3.4487</td>
<td>0.6553</td>
<td>0.1554</td>
<td>4.29</td>
</tr>
<tr>
<td>Married</td>
<td>0.7460</td>
<td>0.4353</td>
<td>-0.0415</td>
<td>-0.97</td>
</tr>
<tr>
<td>Single</td>
<td>0.2109</td>
<td>0.4080</td>
<td>0.0886</td>
<td>32.23</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>12.6383</td>
<td>2.8172</td>
<td>0.0289</td>
<td>12.59</td>
</tr>
<tr>
<td>Years of Experience</td>
<td>21.6922</td>
<td>13.0829</td>
<td>0.0004</td>
<td>-9.33</td>
</tr>
<tr>
<td>Square of years of experience</td>
<td>641.7122</td>
<td>679.5137</td>
<td>-0.0004</td>
<td>-2.63</td>
</tr>
<tr>
<td>Attended Talmudic School</td>
<td>0.0153</td>
<td>0.1229</td>
<td>-0.1567</td>
<td>-31.01</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td>1.8412</td>
<td>31.01</td>
</tr>
<tr>
<td>R²</td>
<td>0.2360</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>6197</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2  
Decomposition of the Wage Dispersion into a between groups dispersion, a within groups dispersion and an overlapping component

<table>
<thead>
<tr>
<th>Actual results</th>
<th>1982</th>
<th>1990</th>
<th>1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>11.5</td>
<td>6.46</td>
<td>14.65</td>
</tr>
<tr>
<td>Within groups</td>
<td>17.89</td>
<td>21.74</td>
<td>26.48</td>
</tr>
<tr>
<td>Overlap</td>
<td>34.19</td>
<td>36.2</td>
<td>31.85</td>
</tr>
<tr>
<td>Total</td>
<td>63.58</td>
<td>64.4</td>
<td>72.98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>In percentage terms</th>
<th>1982</th>
<th>1990</th>
<th>1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>18.08</td>
<td>10.02</td>
<td>20.07</td>
</tr>
<tr>
<td>Within groups</td>
<td>28.14</td>
<td>33.76</td>
<td>36.29</td>
</tr>
<tr>
<td>Overlap</td>
<td>53.78</td>
<td>56.21</td>
<td>43.64</td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Table 3
Decomposition of the Wage Dispersion into a human capital, a rates of return on human capital and an unobservable characteristics components

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human Capital</td>
<td>16.80</td>
<td>19.87</td>
<td>21.44</td>
</tr>
<tr>
<td>Rates of return</td>
<td>2.69</td>
<td>0.33</td>
<td>4.83</td>
</tr>
<tr>
<td>Unobservable Characteristics</td>
<td>44.10</td>
<td>44.20</td>
<td>46.72</td>
</tr>
<tr>
<td>Total</td>
<td>63.58</td>
<td>64.40</td>
<td>72.98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>In percentage terms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human Capital</td>
<td>26.42</td>
<td>30.85</td>
<td>29.37</td>
</tr>
<tr>
<td>Rates of return</td>
<td>4.23</td>
<td>0.52</td>
<td>6.62</td>
</tr>
<tr>
<td>Unobservable Characteristics</td>
<td>69.36</td>
<td>68.63</td>
<td>64.01</td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Table 4
Decomposition for each year of the between groups, within groups and overlapping components into human capital, rates of return and unobservable characteristics elements

<table>
<thead>
<tr>
<th>Elements of the decomposition</th>
<th>1982</th>
<th>1990</th>
<th>1998</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between groups dispersion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differences in human capital</td>
<td>35.41</td>
<td>56.98</td>
<td>11.82</td>
</tr>
<tr>
<td>Differences in rates of return</td>
<td>64.59</td>
<td>43.02</td>
<td>88.18</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Within groups dispersion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differences in human capital</td>
<td>27.61</td>
<td>32.35</td>
<td>36.00</td>
</tr>
<tr>
<td>Differences in unobservable characteristics</td>
<td>72.39</td>
<td>67.65</td>
<td>64.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100.000</td>
<td>100.000</td>
<td>100.000</td>
</tr>
<tr>
<td><strong>Overlap</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differences in human capital</td>
<td>22.77</td>
<td>25.29</td>
<td>31.93</td>
</tr>
<tr>
<td>Differences in rates of return</td>
<td>-13.86</td>
<td>-6.75</td>
<td>-25.39</td>
</tr>
<tr>
<td>Differences in unobservable characteristics</td>
<td>91.09</td>
<td>81.46</td>
<td>93.45</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>