SLAC-PUB-177
February 1966

EMPLOYEE CLASSIFICATION IN A RESEARCH LABORATORY, THE STANFORD LINEAR ACCELERATOR CENTER PLAN

by

Omar Snyder
Stanford Linear Accelerator Center, Stanford University, Stanford California

(Submitted to College and University Personnel Association Journal)

1. THE INSTANT ORGANIZATION

A common phenomenon in this age of technological development at explosive rates is the instant organization. To an idea and a multimillion dollar budget add people and stir. Rather surprisingly the result is often quite good, sometimes even better than the old fashioned variety.

The Stanford Linear Accelerator Center is an instant organization. First proposed in 1957 by a small group of Stanford physicists, and authorized by Congress in 1961, it has designed and built the largest scientific instrument in man's history. Now, five years and 114 million dollars later, an organization of over 1100 people; scientists, engineers, administrators, technicians, secretaries, clerks and laborers, is within a few months of putting it to work searching for the elementary particles of the material universe. As the machine itself represents a synthesis of the most advanced technological concepts, so does the building of the organization which built the machine represent an attempt to rationalize and synthesize new concepts of organization. This is the story of one small facet of that attempt.

2. THE HISTORICAL BACKGROUND

Of course instant organizations are not exactly new. Wartime armies have been built this way for centuries, and until quite recently they provided the model, par excellance, of instant organization. The recipe can be stated quite simply. One needs only a small but highly organized cadre, a precise and complete set of procedural definitions, a full strength manning table, and a precise description of every job in the table, in as simple terms as possible. Now add personnel, assign according to the plan, train to the job and educate in the procedures. This month's recruits are next month's cadres and growth is geometrical. The building of the United States Armed Forces in World War II furnishes a nearly perfect example of the extreme efficiency of the method. It works.

But World War II also provided a contrasting example, in fact many contrasting examples. To choose only one, let us consider briefly the Manhattan Engineer District, the military pseudonym for the atomic bomb project. When the President initialled his approval of the Bush Report, on June 17, 1942, the project was born. There were no cadres, no budgets, no procedures, no manning plans, and no one could know what would have to be done in any detail, much less describe any job of the tens of thousands which were to eventually appear. All that was known was, in the essence of the Bush Report, that there was at that time no known technical reason why it could not be done. The military directive establishing the Manhattan Engineer District was issued the following day, June 18th.

Yet on December 2nd of the same year the world's first sustained chain reaction was demonstrated. In January, 1943, the essential design criteria of the plutonium reactors were settled and a site chosen at Hanford, Washington. On April 6, 1943, ground was broken for construction, and in September, 1944, the first Hanford reactor was in operation. Concurrently, construction of the electromagnetic separation plant began on March 1st and of the gaseous diffusion plant on June 1st, 1943, at Oak Ridge. First personnel arrived at the bomb development site in Los Alamos in March, 1943. On July 16, 1945, the first bomb was demonstrated at Alamogordo.

With every ingredient of the traditional recipe missing, the greatest technical organization the world has ever seen came into sudden being, and successfully served its purpose. Whatever the new recipe may have been, there was no doubt of one thing: it worked too. I think that it may be useful to ask how and why.

3. MAN ORIENTED ORGANIZATIONS, THE NEW DIRECTION

Nothing ever happens entirely spontaneously, yet the sudden appearance of such instant organizations as the Atomic Bomb Project, the Cambridge Radiation Laboratory, the Proximity Fuse Project, and numerous other equally successful but less spectacular wartime scientific development groups has almost the appearance of an organizational mutation. With the later establishment of the Atomic Energy Commission's several national laboratories, NASA's various space centers, such university laboratories as JPL and MITRE, the RAND and Aerospace Corporations, and countless nearly autonomous major industrial research organizations, it seems that the viability of the mutation is solidly established.

Although they are endlessly variable, all of these have one common resemblance, they are man oriented and not job oriented. The one common objective is to provide definitions, not live by them, to create an environment where boundaries to the activity of the creative mind do not exist, in short to do a job which, by definition, cannot be defined.

While this orientation may seem peculiarly that of the research laboratory it becomes day by day a visibly less singular attribute. Within fifty miles of my desk are two of the nation's major universities, two of its major national laboratories for basic research in physics, two of its major national applied physics laboratories, a huge plant of the nation's largest supplier of space hardware, a major plant of its largest supplier of computing machinery, a NASA space center, and many more whose names are thoroughly familiar. Supporting these are hundreds of smaller satellite organizations in each of which more people are engaged in solving problems than in producing things. This is not to say that, as yet, the largest activity of this entire area is research and development, but with today's trend this could be true within a very few years. Indeed it may well be true for the entire nation within a very few decades.

It seems evident that any attempt to describe an organization in terms of jobs will shortly be doomed, since very soon any job which can be described in reasonably simple terms will be done by machines, leaving for man only the task of providing the definitions. This is much the more demanding as well as the proper role of man since it has no limits. There is a conceivable surfeit of things but never of ideas.

Thus the primary objective in the development of the Stanford Linear Accelerator Center classification plan was to define organization in terms of people, not of jobs, a simple enough objective but the essence of the whole.

4. EMPLOYEE CLASSIFICATION AS A DESCRIPTIVE SYSTEM

Any organization is a system of coordinated human effort. In order to consider its operation in rational terms, to forecast its output, control its efficiency, cost, and magnitude it must be analyzed in terms of functions and parameters to which we may attach real, measurable values. This is exactly what any classification system, together with a wage and salary plan, always attempts to do. So long as all tasks were simple, outputs standardized, any man exactly replaceable by any other, and task categories relatively few in number, job descriptions no doubt served as well as anything to establish the parameters of the simple linear systems involved. There may even be a few such organizations left today.

What does one do, however, in today's research laboratory, where there are very likely more Ph.D. scientists than laborers, more engineers than machinists; where it may sometimes take months of close scrutiny to discover <u>any</u> visible result of a man's work, and years, often generations, to put a value on it. And these are the normal conditions in any modern research laboratory. To attempt the application of job descriptions in this environment is approximately equivalent to telling a space ship pilot to estimate his velocity on the way to the moon by counting the number of telephone poles he passes in sixty seconds. It may have worked somewhere, sometime, but it does not work here, now.

The plain fact is that, by the older standards of job evaluation, almost nothing about the system is measurable except the number of people who come in the gate each morning. The result is almost inevitable and fully predictable. We finally end up with a job description for each employee. It is a mathematical truism that any system of order n can be exactly described by a function having n parameters, so we have exactly, but not very practically, described our system. An attempt to attach any sort of quantitative evaluation to such a description is entirely redundant. The mere existence of sufficient labels is entirely adequate to describe the system. A name or number will do as well.

Fortunately, there is a truly simple and practical solution. Any single organization is an economic sub-system of a larger system, the labor market. A free labor market will, for any set of variables, provide an explicit solution for the allocation of labor resources. Consequently, if we observe the distribution of salaries in any organization which is reasonably in equilibrium with a free labor market, and relate this distribution to the variables of the system by some appropriate function, the fitted parameters of the function are statistical estimators of the sub-system. This is a fairly exact, though not entirely sufficient, statement of the theory. It is, however, sufficient to meet my present intention, a description of the development of its practical application.

5. RATIONAL ANALYSIS OF THE ORGANIZATION AS AN ECONOMIC SUB-SYSTEM

What, in essence, I have said in the foregoing paragraph is that even though no single organization may have sufficient data to make a rational evaluation of the economic value of a research worker, the labor market as a whole will, even in the absence of sufficient data, make an explicit evaluation. In even simpler terms, there is a going rate for anything, which, of course, everybody knows anyway. How can we describe it more exactly, in easily comprehensible terms?

There are certain key words in the theory which high-light essential features of the practical classification plan. First, it is statistically based. Descriptions come out in the shape of distributions and probabilities rather than in exact and certain values. Second, it assumes the existence of variables which have objective reality and thus are susceptible to reasonably exact measurement. For practical applications these variables must be not too great in number. Third, it assumes the existence of relational functions which, for practical applications, must be not excessively complex.

The theory, additionally, carries an assumption which could finally be proven true or false only by experience. Do all, or any, organizations operate in reasonable equilibrium with a free labor market? Does a free labor market even exist? I will examine both of these questions later, but for the present it will suffice to say that experience has shown that such qualifications as are necessary on affirmative answers to both questions do not pose large problems.

Finally, we must state a critical and not too obvious characteristic which follows from consideration of the theory and the necessary measures which must be taken to implement it in a practical plan. Any such plan must be descriptive, not prescriptive. It cannot tell us what should be done, only what is being done and what has been done. It is, however, fully capable, within the limitations of statistical inference, of providing estimates of what will be done in the future. A full understanding of this critical characteristic is necessary for successful administration of the plan. The only administrative problems of any significance which have ever arisen in the five years of our experience have been the result of occasional failures to fully grasp all of the implications of dealing with a purely descriptive system.

All that is necessary now in order for us to make an explicit model of the system and evaluate its parameters, is to select its significant variables and to formulate appropriate functional relationships.

6. THE SYNTHESIS OF THE ADMINISTRATIVE MODEL

Most workers in this field are familiar with the so-called maturity curve approach, in which salaries in some relatively homogeneous group are expressed as a function of age of individuals. Until quite recently this approach has been limited to quite narrowly defined classes, BS degree level, non-supervisory engineers for example. I know of very few attempts to apply maturity curve descriptions to less sharply defined classes, and none, other than ours, to virtually an entire organization.

The common use of maturity curves is easily recognizable as an application of the statistical technique known as stratified sampling. From this viewpoint the Stanford Linear Accelerator classification plan is nothing more than a system of population stratification, carried just far enough to provide an acceptably small variance in the salary distribution in each stratum.

Our system is based on analysis of a population stratified by occupational classification. In 1961, when the plan was first put into practical operation, we chose to use nine such occupational strata. Subsequent experience, however, showed that two of these, even after a ten-fold increase in size of the organization, still provided samples too small for statistical significance, and further that there was no suggestion that they could ever be differentiated from two other strata. Consequently, at the end of 1965 the number of occupational classifications was reduced to seven, where it seems likely to remain.

The seven occupational classifications, with the titles which we have applied to them, are the following:

	Classification		Number
1.	Executives and Faculty Members		24
2.	Staff Members		272
3.	Staff Associates		174
4.	Technical Staff Assistants		361
5.	Administrative Staff Assistants		31
6.	Clerical		150
7.	Manual		91
		Total	1103

In more detail, each class includes these types of persons:

1. The first classification is self descriptive and is limited to approximately 2% of the total. Because of both the small size of this subgroup and the highly diversified characteristics of persons comprising it, no attempt is made to subject it to administrative statistical analysis or control. Each person is considered on a purely individual basis.

2. Staff members are exempt, professional level employees of all kinds; scientists, engineers, and administrators. It is limited to persons having demonstrated professional competence, which is to say, at least enough experience in positions of responsibility beyond the bachelor's degree level to allow judgment of ability. Higher degrees may be substituted for experience.

3. Staff associates comprise a relatively heterogeneous group of exempt personnel. Included are some less experienced degreed engineers and administrators, supervisors of shops and administrative support groups, and some highly skilled technical specialists. About one-fourth have degrees and another one-third have substantial education beyond high school graduation.

   All personnel in the remaining four groups are non-exempt under wage and hour law standards.

4. Technical staff assistants include all skilled technicians in all occupational categories, for example, electronic technicians, maintenance electricians, machinists, electroplaters and draftsmen.
5. Administrative staff assistants include such occupations as junior accountants, analysts, buyers, expediters, librarians, writers, and storekeepers.
6. Clerical is self descriptive and includes all occupations conventionally classed as such. Its most notably distinguishing characteristic is that it is almost exclusively a female group.
7. Manual is also self descriptive and includes all unskilled and semiskilled employees.

We have now established the basis of the stratification of the organization by occupational class. If we next examine the distribution of salaries in each stratum we find that they cluster about parabolic curves with reasonably small scatter. We describe each such curve by a second degree, or quadratic equation, in which the parameters of the function are fitted to the data by the method of least squares.

Although it is possible to extend the analysis to include more variables and fewer functions, which I will cover in the next section, there is little or nothing to be gained in administrative convenience by doing so. While the model becomes mathematically simpler and more precise, it also becomes administratively more complex. We are now able to describe the organization as an economic subsystem in terms of two continuous variables, age and salary, and six dummy variables, the six occupational classifications, excluding only the first, or top level group.

Measurement of both continuous variables is completely objective and of the six dummy variables quite acceptably so. This model requires six nonlinear, but quite simple functions, each having three parameters. These relate age and salary in each of the occupational classifications. There is no need for any job descriptions and we have none. This, then is the substance of the administrative classification plan. How does it work out in practice?

Since salaries are described as statistical distributions, the average experience is always easily determined from the current age and salary distribution in each of the occupational classifications. Control is exercised over the integral of each distribution, or, to look at it in another way, over the entire class age adjusted average. Thus if it is decided to allow some predetermined increase in the salaries of an entire class over, say one year, this is applied to the average of the class, not to individuals. There are no ranges. Budgetary limitations have been quite easily met in actual experience. It is only necessary to keep in mind the necessity of fair allocation; more than average here must be balanced by less than average somewhere else. Thus the salary review and adjustment process becomes primarily one of allocation based on judgment of peer rankings. Reviewers need not be concerned with technical details in order to make fair allocation decisions.

In practice, salary reviews and adjustments are a nearly continuous process on a completely individual basis. The normal practice here is to make all reviews on the anniversary of employment. The predetermined budgetary controls attempt to forecast the probable changes which will occur in the entire system from all causes during the period of each forecast, normally a year. Thus each adjustment will include perhaps a bit for individual improvement, a bit for age, and a bit for the general rise in values, call it cost of living, productivity, inflation, or what you will. It has never been necessary to grant any general increases for any reason.

Among other things, it has never been necessary to set up a wage and salary administration group. We have so completely integrated the administrative functions of classification, records maintenance, procedures, reports, controls and adjustments that it is virtually impossible to identify any one of them as an individual process, or as an individual responsibility. One aspect of this total concept is, however, of such critical importance to the successful operation of the whole that it deserves special note.

No such system can conceivably operate without an accurate, complete, and current records system. The records system here was designed specifically as an integral part of the whole. It is highly mechanized. All statistical data is kept in punched card files, brought to current status at least weekly, more often when necessary. Errors are virtually non-existent. All major routine reports and many non-routine reports are computer produced as are all routine salary review forms. Administrative responsibility for operation of the system is carried by one part time professional employee. All new employee inductions, all terminations, all salary review paper work, all record changes, reports, surveys, inquiries, and numberless miscellaneous duties are handled by an office staff of three persons.

System design, programming, monitoring, analysis and control have been a part time responsibility of the author. Parenthetically, virtually all of the administrative problems which occur in operation of the system are handled routinely by the administrative group. They have never accounted for more than a completely insignificant fraction of my effort.

7. THE SYNTHESIS OF THE ANALYTICAL MODEL

The analytical model continues the resolution of the administrative model to a mathematically simpler and more concise form, introducing more variables and fewer but more complex functions with a substantial reduction in the number of parameters necessary to describe the system.

The key observation which led to the development of this model came about when the entire set of parabolic curves, representing age and salary relationships in each of the occupational strata, were plotted together on the same scale. Quite surprisingly, this set of curves unmistakably suggested that they were all members of what is called a one parameter family. Since each curve has three parameters, two will serve to describe the common characteristics of all, and only one is necessary to distinguish one from the other. Figure 1 is such a plot, showing the relationship of age-salary curves in the seven occupational strata, each over the actual range of ages found in the data.[1]

The two apparently common characteristics of all of these curves are a single axis of symmetry, and a common point of intersection. The distinguishing characteristic is the order, or curvature. This entire family of curves may thus be described by the simple function

$$Y = R\alpha + \beta + u$$

where Y is salary rate in dollars per unit of time.

R is the ranking parameter which describes the curvature or order in the family.

α is a non-linear quadratic function of age.

β is the point of common intersection.

u is the stochastic error term. [2]

It is interesting to note that at $R = 0$, $Y = \beta$, consequently the parameter, β, expresses the effective minimum wage rate in the organization, (more precisely, $\beta + u_{\beta}$).

The ranking parameter, R, is in itself a complex function of several variables. I have identified three, by multiple regression analysis, as being responsible for nearly all of the observed range of values of R. These are occupation, education, and level of responsibility. While all three are highly intercorrelated their range of variability encountered in individual cases is quite sufficient to allow reliable statistical estimation of their parameters, even in relatively small samples of size 100 or so.

Unfortunately, the effects of some other possible variables cannot be identified because of intercorrelation and excessively limited ranges of variability within a sample of the size so far studied. For example the possible effect of sex cannot be estimated since it is excessively intercorrelated with occupation. There is simply no way to say that the value statistically estimated from the sample for the clerical occupation has its observed value because of the occupation or because that occupation is almost exclusively inhabited by females. Conversely, the number of females in any other occupation is so small compared to numbers of males that any effect due to sex is quite impossible to find, with any reasonable degree of significance.

Several forms of the function which may serve to express R in terms of the variables occupation, education, and level of responsibility have been investigated empirically. The one which I have finally chosen as best suited is the following:

$$R = X_O + \sqrt{X_O X_E} + X_R \quad .$$

In this form all three are dummy variables. It is necessary to introduce the educational variable, X_E, in the form of its geometric mean with the occupational variable, X_O, in order to account for the peculiar way in which the two are intercorrelated. Education is a strong factor in determining compensation only when it is appropriately applied in the occupational field. A Ph.D. is a highly significant attribute of scientists, but it would be entirely insignificant in clerks and laborers.

There appears to be no way to handle the occupational variable except as a dummy variable. It is, however, possible to express the educational variable, X_E, as a function of the continuous variable, years of school completed, and the level of responsibility variable, X_R, as a function of the number of persons supervised. This has been done and in both cases the results are interesting and highly informative.

The education variable, X_E, may be quite exactly expressed as

$$X_E = ae^{it} + B$$

where t is the number of years of school completed,

e is the base of natural logarithms,

and a, i and B are system parameters.

From this function, and the magnitude of the various parameters, it may be observed that the effect of more or less schooling at the level of high school graduation, t = 12, is rather insignificant because of the size of the constant term, B. At doctorate levels of college education however, t = 20, the effect is almost purely exponential.

The function which may be used to express the level of responsibility, X_R, is exceedingly simple,

$$X_R = \sqrt[3]{N}$$

where N is the number of persons supervised. It is evident that responsibility is compensated not in terms of the volume of responsibility, but of its linear measure. Thus, if we speak of the sphere of influence, the measure which determines compensation is its radius. It appears that the important factor is not how many we touch but how far we reach.

In spite of the apparent complexity of the complete model, all of its parameters may be quite easily fitted by computer programmed multiple regression methods, and all parameters of the Stanford Linear Accelerator Center static model have been computed. Successive computations show that the model is dynamic and additional parameters will be necessary to reflect such changes as national level of employment, productivity and other labor market factors. Such a model has not as yet been formulated.

Finally we should comment briefly on what experience has shown us about free labor markets and system/market equilibrium. It is unquestionably true that there is no such thing as the labor market; there are many markets, more or less interrelated, more or less free. Some are almost completely isolated from the others. At one end of the spectrum, for example, is the market, or several markets since they are each isolated from the others, for the highly specialized professionals, medical doctors, attorneys, clergymen, etc. Fairly closely following is the academic market, whose workings have been so aptly described by Caplow and McGee in "The Academic Market Place."

At the other end of the spectrum is the agricultural labor market, and the unionized trades' labor markets. Most of these are subject to some form or forms of restraint; monopoly controls, limited access, etc., so that they are not merely not free but hardly markets. To the extent that employees are drawn from such sources, all that has been said cannot apply without major qualifications.

Between these limits, however, lies the vast market which is outlined in the advertising pages of the daily newspapers and the trade and technical journals. This market is, for all practical purposes, a truly free market and certainly the major part of the working force lies under its control. Geographical and institutional market subsystems within it vary only by what amounts to minor arbitrage, which is easily accounted for. For the most part minimum wage law limitations have little significance since few organizations which draw from this market have any requirements which can be met by employees who would be subject to its action.

Lastly, on reasonable equilibrium with the market, it is a practical impossibility for any organization to operate in any other way. While it is conceivable that scales of pay could be set so high that all interaction occurs on a one way street (one recent well publicized example comes to mind), it is inconceivable that any such a situation could exist undetected for any substantial length of time. Assurance that reasonable equilibrium is in fact the condition prevailing is given by simple observation of turnover rates. If both acquisitions from the market and losses to the market are occurring then equilibrium is an inescapable fact. The notion that any organization can set its own, wholly self determined value system is sheer fantasy.

REFERENCES

1. The observation that the family of age-salary curves produced by sample stratification was clearly a one parameter family would not have been so surprising had I been aware of the existence of an earlier paper. See H. S. Houthakker, "Education and Income," The Review of Economics and Statistics, 41, 24-28, (1959). Houthakker made an analysis of the distribution of money incomes as a function of age, stratified by years of school completed, in eight levels. His data was drawn from the 1950 census of population, which tabulated this information for a 3-1/3 percent sample of the male population ages 14 and over. Aside from the difference in modulus of the salary scale, 1965 being more than double 1949, his plotted data relationships were virtually identical with those which I found in analysis of Stanford Linear Accelerator Center data. In order to facilitate comparison I have plotted Houthakker's Table I, data for the upper seven educational levels on Fig. 2. Compare Fig. 1. The occupational classification stratification scheme is considerably more sensitive than one based on years of schooling. Since occupation and education are highly intercorrelated, however, years of schooling provides entirely acceptable stratification in a sample as large as Houthakker's, on the order of two million by my estimate. Unfortunately, I did not learn of the existence of Houthakker's paper until a few weeks ago. While he did not extend his observations to the development of a system model hypothesis, the way to do so was clearly indicated by his work.

2. I have avoided a discussion in the text of the implications which arise from analysis of the stochastic error term, u, since some technicalities arise which are not of great interest to anyone except theorists. For those interested, see "Econometric Methods," J. Johnston, McGraw-Hill, (1963), 5-8, for an excellent and lucid discussion. In the present context the distribution of u may be taken as

the measure of the imputed variability of individuals in the organization matrix, measured in terms of the variability in actual salaries when all other assumed causes of variation have been taken into account. As such it encompasses not only differences in people but also errors in evaluation of these differences, and the possible effect of neglected variables.

The distribution of u is heteroscedastic in my model but the coefficient of variation of its distribution, defined as the standard deviation divided by the mean, for any small subset of y within the data matrix is very nearly constant with a value of .08. The range of values of y is about ten to one in my data. I consider this value of .08 almost uncomfortably small, hence my emphasis in the text on the application of the system as a descriptive, not a prescriptive model. A constant surveillance of the coefficient of variation provides a useful monitor on the development of possible perversions.

This is of course somewhat a matter of personal preference, some people like highly deterministic systems, others prefer anarchy. Our experience has indicated that a middle course, with coefficients of variation in the range of .1 to .08 provides a generally acceptable compromise.

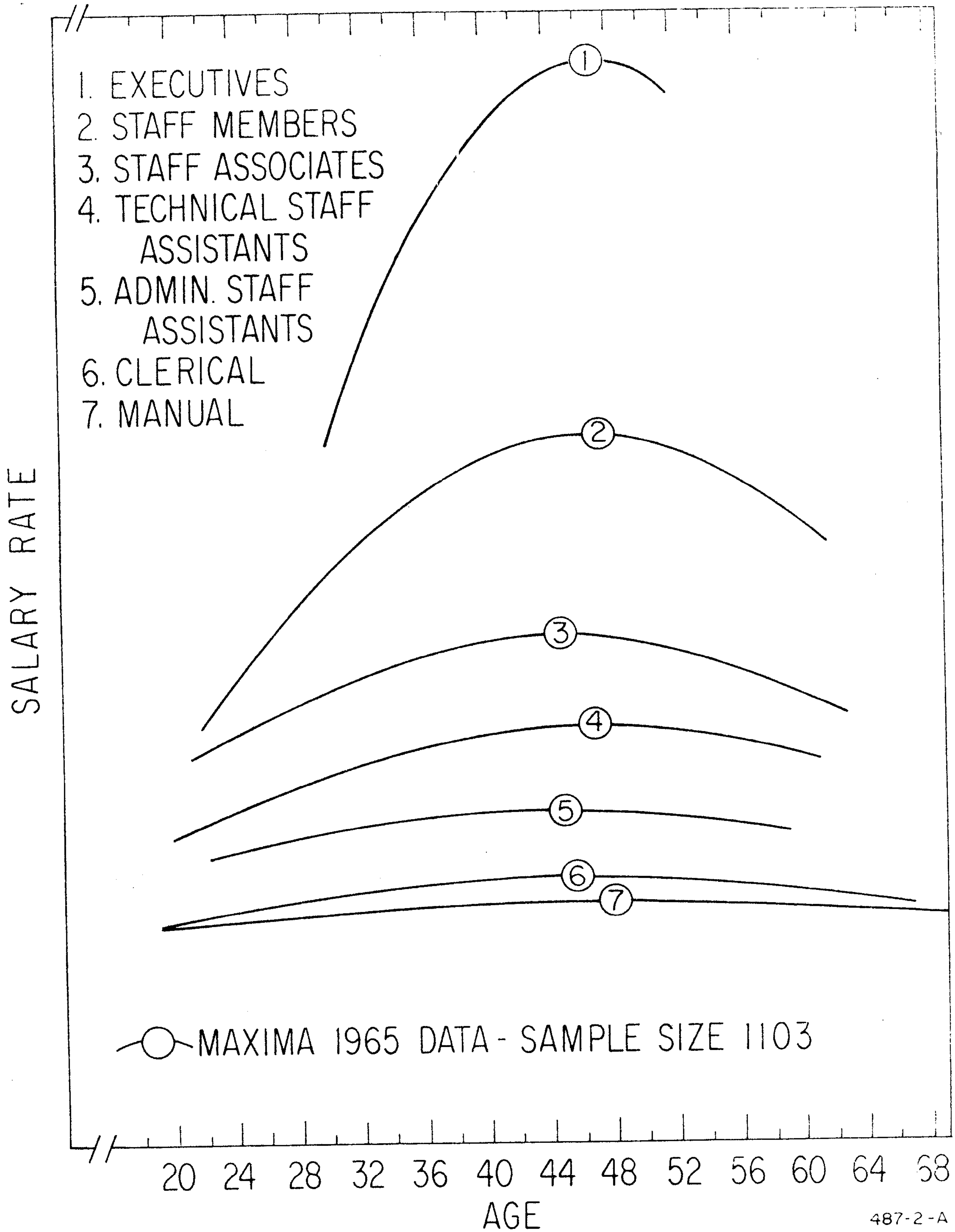

FIGURE 1--LEAST SQUARES FITTED PARABOLAS, AGE VS. SALARY, S.L.A.C.

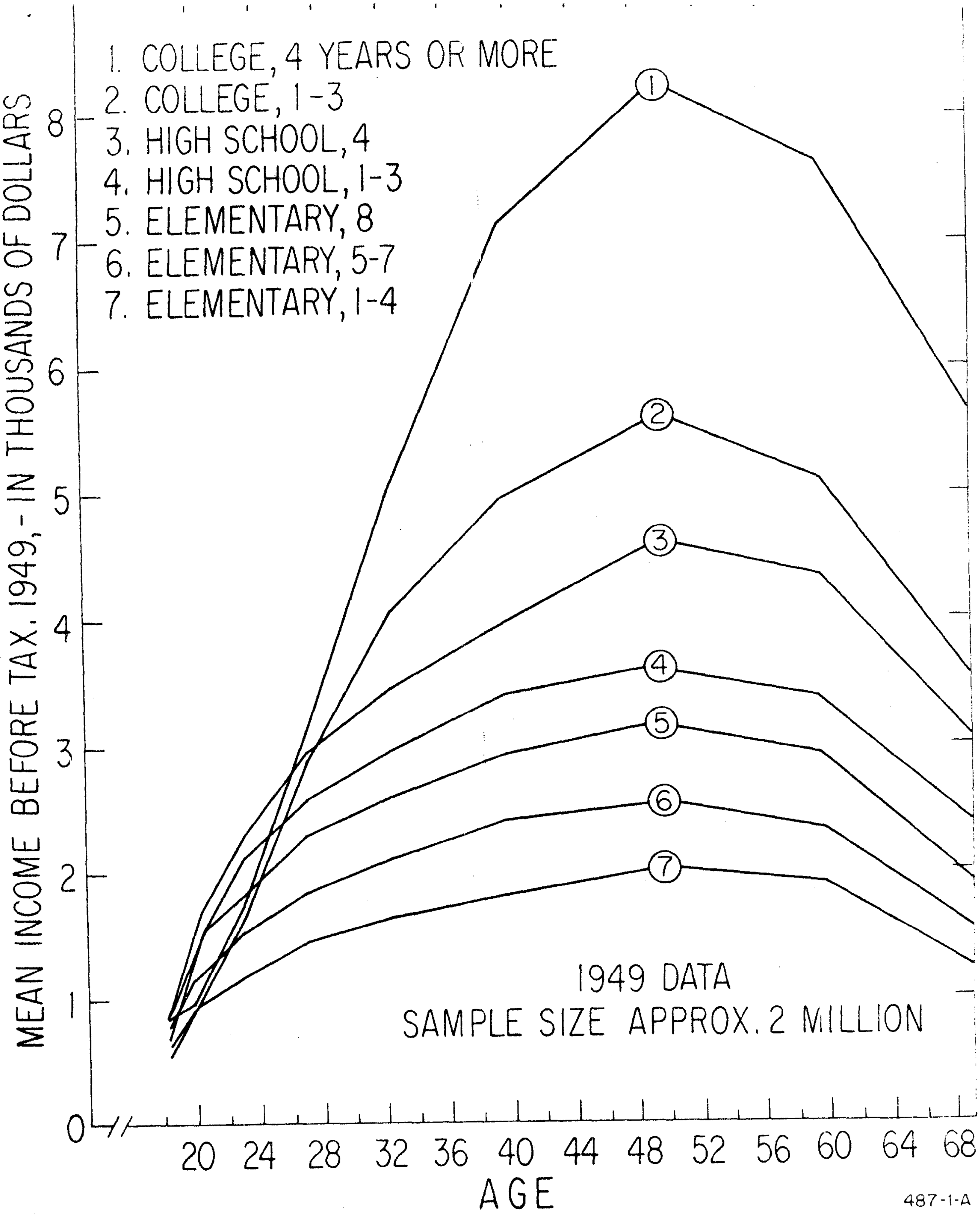

FIGURE 2 --DATA OF H.S.HOUTHAKKER, OP. CIT., FROM TABLE 1. MEAN INCOME BEFORE TAX BY AGE AND YEARS OF SCHOOL COMPLETED