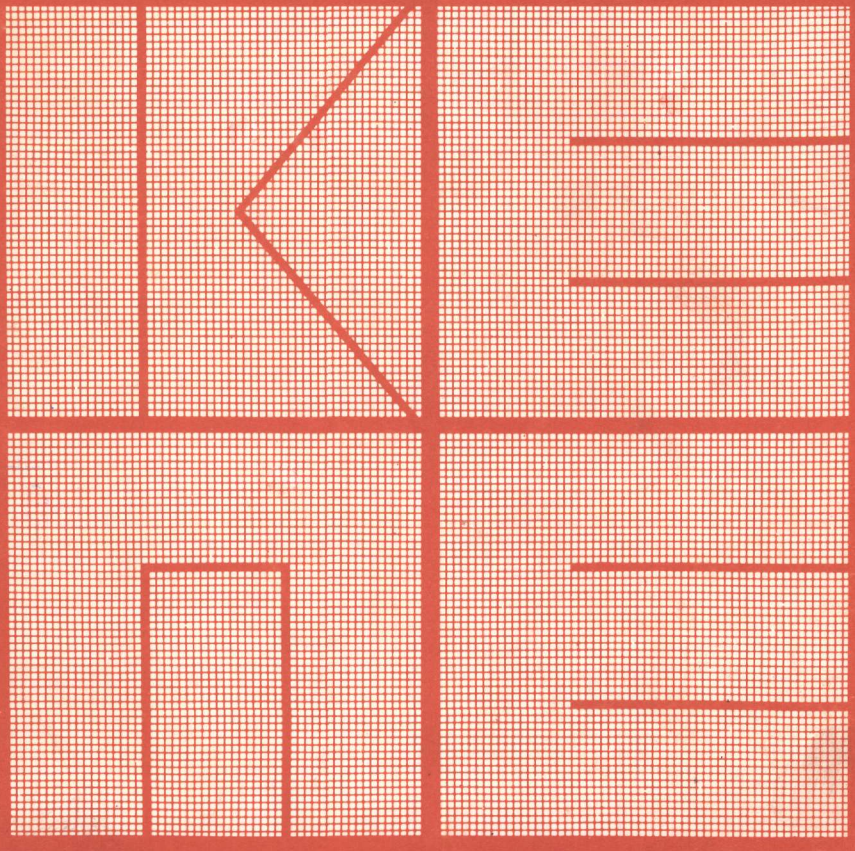
CENTRE OF PLANNING AND ECONOMIC RESEARCH

Reports on Planning 32

J. B. Grossman

A Study of the Greek Labor Force statistics



Athens 1986

A Study of the Greek Labor Force statistics

CENTRE OF PLANNING AND ECONOMIC RESEARCH

Reports on Planning 32

A Study of the Greek Labor Force statistics

Jean B. Grossman

Athens, December 1986

Opinions or value judgements expressed in this book are those of the author and do not necessarily represent those of the Centre of Planning and Economic Research.

CENTRE OF PLANNING AND ECONOMIC RESEARCH

The Centre of Planning and Economic Research (KEPE) was established as a research unit, under the title "Centre of Economic Research", in 1959. Its primary aims were the scientific study of the problems of the Greek economy, encouragement of economic research and cooperation with other scientific institutions.

In 1964, the Centre acquired its present name and organizational structure, with the following additional objectives: (a) the preparation of short, medium and long-term development plans, including plans for regional and territorial development and also public investment plans, in accordance with guidelines laid down by the Government; (b) the analysis of current developments in the Greek economy, along with appropriate short-term and medium-term forecasts; also, the formulation of proposals for appropriate stabilization and development measures; (c) the further education of young economists, particularly in the fields of planning and economic development.

The Centre has been and is very active in all of the above fields, and carries out systematic basic research in the problems of the Greek economy, formulates draft development plans, analyses and forecasts short-term and medium-term developments, grants scholarships for post-graduate studies in economics and planning and organizes lectures and seminars.

Within the framework of these activities, the Centre also publishes studies from research carried out at the Centre and lectures given by specially invited distinguished scientists.

The Centre is in continuous contact with similar scientific institutions abroad and exchanges publications, views and information on current economic topics and methods of economic research, thus further contributing to the advancement of the science of economics in the country.

PREFACE

The preparation of national development programmes presupposes a thorough analysis of the relevant data and the clearest possible understanding of the problems and the development potential of the greek economy.

In the context of the preparation of the Five Year Programme 1983-1987 an effort has been made for the systematic presentation of statistics and data as well as for the preparation reports and studies for section not sufficiently studied in the past. These papers, the work of KEPE staff, and greek and foreign analysts, more generally contribute to the furthering of economic research, which in Greece happens to be still limited.

The purpose of this series of issues is the publication of these studies which are considered to be of a more general interest and which assist in a better understanding of the process and goals of planning.

As usual in similar cases, the views and opinions expressed are these of the authors and they are not binding for KEPE.

LOUKIS A. ATHANASSIOU
Active Director

Centre of Planning
and Economic Research

ACKNOWLEDGEMENTS

This work was done for the Centre of Planning and Economic Research (KEPE) in Athens, Creece.

The author would like to thank the many people in the Greek Government with whom she talked in collecting the information needed for this work. In particular, she would like to thank Mr. Kontoulis, Mr. Kalambokidis and Mr. Hatzhipanayotou from the National Statistical Service of Greece (NSSG), Mr. Hatsiglou from the Ministry of Labor, and Mr. Efstratoglou and Mr. Zissimopoulos from the Manpower Employment Organisation (OAED). The staff at the Centre of Planning and Economic Research (KEPE) including Mrs. Economopoulou, Mr. Glytsos, Mr. Karmas, Mr. Kensakis and Mrs. Papaconstantinou were also particularly helpful. I especially thank the Director of KEPE, Mrs. Katseli, for setting up all my appointments.

The author would also like to thank the Foerder Institute for Financial Assistance with regard to the typing of this document.

JEAN B. GROSSMAN

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Labor force statistics are the most important source of data from which we can learn about the welfare of a nation's people. They not only make policy-makers aware of changes in the state of the economy, but they also shed light on the type of problems that economy is facing and the type of policy response that is best. However, to be useful the data need to be timely, appropriately defined, and of high quality. In this report we examine what data should be collected and how the data should be collected is discussed. These ideals are compared with what is actually collected and how it is actually collected. Suggestions are made as to what possible changes could be made to the existing system.

A. Labor Force Statuses

The first step in developing a suitable system of labor force statistics is to decide whose labor force participation should be followed. Until 1981, data were collected on four different categories of people: a) individuals over 10 years old, b) urban and semi-urban inhabitants over 14, c) many persons in the private sector and d) others employed in large manufacturing firms. In 1981, individuals in rural areas over 14 were added to the second category. The plurality of categories is not necessarily bad. We find that comparing the labor force situations of several categories can lead to deeper understanding of who is bearing the brunt in a recession or

gaining from a boom. However, the category definitions should be made with these possible comparisons in mind, rather than merely letting them develop haphazardly.

The next important step that must be taken is to decide how to operationally categorize individuals into labor force statuses and which statuses should be used. There is agreement among economists and policy-makers that, at a minimum, individuals should be categorized as employed, unemployed, or not economically active. The major issue to resolve is whether or not to collect the information needed to distinguish between those who are economically inactive for completely voluntary reasons and those who are inactive because they are discouraged about the prospects of finding a job. The advantages in separating the two groups are: b) being able to more accurately describe the deterioration in welfare during recessions, and b) having a measure of economically-determined unemployment which is more comparable across the urban, semi-urban and rural division. The advantages, however, are just about outweighed by: a) the cost of the additional data requirements, needed to make the distinction and b) the problem that the implementation of the discouraged worker concept still produced an imperfect measure.

The operational definitions of employment and unemployment are also considered in the paper. For example, should one require someone to work a minimum number of hours per week to be considered employed? Likewise, should one require the unemployed to search a minimum number of hours? How does one

determine if a non-working person is really trying to find a job? The employment hours criterion is rejected, since employment-unemployment statistics have never tried to distinguish between those who are fully employed and those who are not. This is not the role of labor force status figures. Other variables, such as hours worked are better suited for measuring employment intensity. Restrictions on the amount and type of search are also rejected because it is felt that the sincerety of search cannot be determined very well with hours or type-of search restrictions and what constitutes serious search behavior differs radically between occupations.

B. The Data that Exist

There are four sources of Greek labor market statistics: the labor Force Survey (LFS), employers' surveys, data from the Manpower Employment Organization (OAED) and the Census of Population. These are briefly reviewed below. The National Statistical Service of Greece (NSSG) collects all but the OAED data. The Census is the largest data set, but it is so infrequent (once every ten years) that it is only used to inform long run policy. The Labor Force Survey is the next largest survey. It is an excellent sourse of data about the general public in urban and semi-urban areas. The major complaint voiced against it is that the quarterly observations "jump around" too much, making intertemporal comparisons difficult. Comparability among quarters could be improved by reinterviewing some fraction of the individuals who were sampled in the last period, instead of drawing a new sample each

quarter. Survey costs would be likely to fall if this rotating sample technique was used. Another important problem has recently developed with the Labor Force Survey. As in 1981 individuals of the rural areas were added to the Labor Force Survey's scope. Whereas this was a laudable change, it was bought at the price of making the previously quarterly survey an annual one. This, seriously reduced the usefulness of the Labor Force Survey to national policy-makers. It is suggested that the quarterly feature of the survey is reinstated as quickly as possible, even if it means reducing the number of interviews collected.

The Government's employment service, OAED, gathers data on the number of people who are employed and unemployed. However, they gather this information on only a small part of labor force, primarily the experienced employees in the private sector, who have a permanent attachment to the labor force. As one might imagine, the variation in these OAED series differs from the series which cover a broader segment of the labor force. There are no particular problems with the OAED data, except that they cover a very special and small part of the labor force.

Three surveys of employers are conducted in Greece. The largest and oldest is in the manufacturing sector. The other two are in retail and in mining. The major problem confronting these surveys is that they are based on the firms that existed during the last Census of Manufacturing. Thus, they do not represent any of the new, emerging firms, and may

over represent other types of firms that have declined in number since the last Census. These employer surveys could be improved by incorporating the new Censuses more rapidly than in the past, and to study the characteristics of new firms and extinct firms. A study of emerging and dying firm can indicate how the "true" picture differs from the one given by the employer's surveys.

The existence of more than one labor force data source does not mean that effort is being duplicated unnecessarily. Each source provides the policy-maker with a view of the economy from a different perspective. One can learn a great deal by comparing the different data sets. In fact, one should construct the data sets with these comparisons in mind.

The last topic to be discussed in the report is the administration, presentation, and dissemination of the data. Several goals should be kept in mind when considering changes in the surveys themselves, or changes both in the administration and/or representation of data:

1. Every piece of data that is worth collecting is also worth making available in a timely and useful manner. Data should be readily available to analysts and policy-makers in a usable form, if they wish to do further analysis with it.

2. Future data needs, as well as current data needs, should be kept in mind when considering changes.

3. Survey changes should be done only if the benefit of change outweighs the monetary costs and the costs in terms of short-term deterioration in inter-temporal comparisons.

A summary of specific recommendations is included at the end of the paper.

I. GOALS OF LABOR FORCE STATISTICS

Labor force statistics should serve two important policy purposes: first, to indicate the state of a nation's economic performance; and second to indicate the economic welfare of the society. Industrial production series gives one indication of economic performance; but much more can be learned about the source of problems if production data are supplemented with labor market data. On the other hand, labor market data are the primary source of data from which one can discern the economic welfare of a nation. Thus, timely and accurate labor market information must be provided to policy-makers if wise economic programs are to be developed.

In this report, we discuss how labor force data can be used to inform policy-makers about the state of the economy and the welfare of the people. Then, we discuss how one gathers such data conceptually and operationally; from whom data should be collected; and how individuals should be categorized. The third section describes and evaluates the quality of the four major data sets: the Labor Force Survey, the data from the OAED, the employers' surveys, and the Census of Population. What can be learned by comparing these various sources is discussed in the fourth section. Administrative issues, such as inter- and intra-agency coordination, data presentation, and dissemination, are briefly discussed in the fifth section. The report ends with a brief summary of the recommendations that are made throughout the paper. The appen-

dices present statistical issues, such as how large a change in the unemployment rate must be observed before one can conclude that the economy has really deteriorated and what is a cost-effective way of designing any sample.

In this chapter we discuss the uses of labor market information. There are primarily three uses: indication of economic performance, indication of social welfare; and evaluation of policy intervention. These three uses are discussed below.

A. Economic Performance

Labor is one of a country's most important factors of production. The utilization rate of this factor, and espesially the differential utilization rates among industries and occupations, can provide much information about the health of an economy. For example, the drop in industrial production tells the policy-maker that something is wrong, but it does not tell him what type of problem has occurred or what type of policy is best to combat it. However, if one observed a high demand for technical workers, while unskilled workers were becoming unemployed, one might conclude that the problem is structural rather than demand-related. Retraining programs or other adjustment programs may be justified. Therefore, information about employment and unemployment is needed by occupation and productive ability of the economy.

B. Welfare

Similarly, in order to initiate appropriate income policies or to evaluate the impact of other government policies on the society's welfare, one needs information on the labor market. Welfare statements, though, are more difficult to make than production statements, because a worker's well-being does not increase continuously as the amount of hours worked increases. A retiree, for example, who works no hours may be happier than someone who holds two jobs. A worker's welfare depends on both income and hours worked. However, even if we had income data on individuals, it would be difficult to make comparisons among individuals in different labor force statuses. Is an unemployed individual receiving 500 drachmas per month better off than an employed individual earning 700 drachmas per month and working 48 hours per week? The unemployed individual enjoys 48 hours more leisure than the employed individual. The best that could be done is to make intra-group comparisons: an unemployed individual receiving 500 drachmas per month is probably better off than an unemployed individual receiving only 300 drachmas. Without income data, the welfare comparisons that can be made are even more restricted and require more assumptions. We assume that transitions from employment to unemployment (seeking a job) are involuntary and therefore represent a deterioration in the society's welfare, given that all other things are constant. Transitions between employment and not economically active are a little more difficult to interpret because some people make the transition completely voluntarily, while others may be workers who would

like to find a job but have given up seeking employment because the prospects are so dismal. An increase in the number of economically inactive discouraged workers represents a deterioration in welfare, but an increase in the number of economically inactive women due to childbirth may not.

Thus, to make welfare statements from observed changes in labor force status, one must assume that the structure of the labor market and the composition of the labor force remains constant. Therefore, welfare statements should be restricted to short-run relative statements, such as "the Greek people were less well off in 1980 and 1981 than they were in 1978-1979". One cannot say, for example, that the society is twice as bad off since all things have not remained constant. During this particular time period, the participation of women increased dramatically. One cannot evaluate the relative loss in welfare due to the unemployment of a second wage earner as opposed to the primary wage earner. Thus, we can only state that the society is less well off than before. This illustrates that, in addition to labor force status data, information on the structure and composition of the labor force and labor market is needed in order to interpret the changes in employment or unemployment data.

C. Policy Evaluation

In addition to providing policy-makers with information about the underlying production and welfare trends, labor force statistics are also called upon to determine short-

run effects of policy or exogenous shocks. These short-run indicators are needed to allow the Government to respond promptly to developing problems.

The need for frequent and accurate data on the labor market is clear. However, budgetary constraints restrict any country from gathering and disseminating data on everything that is desired. Thus, it is up to a Government to set priorities. Whatever data are collected, should be accurate and disseminated quickly enough to be useful. Therefore, policy-makers must make difficult tradeoffs between the number of general areas in which data will be collected and the level of detail in each area. Recently, a commission has been formed, headed by Mr. Kalambokidis, which considerer just these decisions for the data collected by the National Statistical Service of Greece (NSSG). In recommending changes, the Commission should keep in mind that the benefits from a change should outweigh not only the monetary cost but all the cost of reducing intertemporal comparisons with previously collected data.

This paper deals with not only the NSSG labor force data, but also labor force data from other sources. The quality.of these data is examined and suggestions for possible improvements are stated when appropriate.

II. LABOR FORCE STATUSES

A. The Appropriate Population to Consider

The first question that must be addressed in developing labor force statistic is from which individuals in the society should data be collected. From the previous discussion on uses, it would seem that data should be collected on the total potential productive labor supply. In this way, one can learn about: a) the economy's aggregate labor supply for production purposes and b) how many people, who could potentially earn money in the economy, are for welfare purposes.

Should anyone be excluded? It is generally assumed that children have a low probability of being in the labor force and are therefore excluded. However, at what age should they start being counted? At what age could they add substantially to production or family income? In agricultural societies, children as young as ten may provide a significant amount of labor towards farm production. However, in more developed societies or where there are widely enforced compulsary education laws, the minimum age is usually increased to the age at which education is no longer mandatory. Greece has a large, but shrinking agricultural sector. It therefore collects employment data under both regimes. In the Census, the minimum cutoff is ten, but in the Labor Force Survey (LFS), the cutoff is fourteen. Overall, the inclusion or exclusion of individuals' ages, ten through fourteen, make little dif-

ference. In 1971, for example, they comprised only 2.1 percent of the economically active. Even in the rural areas they made up only 2.8 percent of the labor force.

Should there be a corresponding cutoff of older individuals? The problem in implementing a cutoff is that there is no legal age at which employment must end, or even a clear cut-age at which the majority of individuals totally stop their productive activities. By making 65 a cutoff point, one would miss the large number of individuals who continue to work in some form and to some extent. Thus, a maximum age cutoff is not recommended. Greek data do not use one.

B. The Employment Category

Traditionally, the Greek labor force has been categorized into one of three groups: employed, unemployed, or not economically active. Let us examine each category to determine who should be considered in each.

Employment seems straightforward. However, there are issues that should be considered. First, what constitutes work? Second, does one have to work a minimum number of hours? Individuals that donate their time and labor to either the family or to economic institutions are currently not counted as working. Clearly, they are adding to the nation's production, directly or indirectly. However, their inclusion in the employed labor force would make the interpretation of employment as a point on the aggregate labor supply curve tenuous, since their current non-market labor supply is not directly

related to wages, and in fact, may be inversely related to wages. It would also confuse the interpretation of changes in the number employed as a change in welfare. Exclusion of household activities and voluntary work, as is done now, is compatible with current national income accounting.

A more difficult decision, in defining employment, is whether to require that an individual work a minimum number of hours in order to be counted as employed. It seems wrong to count someone who works only five hours a week with equal weight as one who works forty hours per week. However, this feeling stems from expecting more out of a categorical variable than one can get. Labor force statuses (employed, unemployed, not active) are mutually exclusive threshold indicators. An individual must fall into one, and only one of these categories. Thus, a part-time employee is given the same weight as someone who puts in overtime. The percentage of individuals in each group is a general indicator of labor force utilization. If a more precise measure of utilization is desired, one should use a more continuous variable, such as hours, rather than changing the logical thresholds. However, as we will discuss in a later section, even hours of work are subject to conceptual difficulties because of underemployment. It is important to keep in mind that employment and unemployment numbers only tell part of the story. An increase in unemployment is likely to understate the idleness of resources, since many "employed" individuals are not supplying as many productive hours of work.

C. The Unemployment Category

The principle behind defining unemployment is that the sum of unemployed and employed individual should be equal to the total labor force. In other words, unemployment should consist of those people who want to work, given full information on the state of the economy. The operational problem that arises in measuring unemployment is how can we objectively determine whether or not someone is seriously interested in finding a job. The solution that labor force surveys have employed is to inquire into a person's job search acivities. Specifically, in the Greek Labor Force Survey, a person must have actively searched for a job in the last week. Although this criterion is concrete and should indicate labor force attachment, one weakness in the criterion is in the definition of job search. Different people use different search techniques. The job search behavior for professionals can be very different than that of unskilled workers. One wants to include only serious search acivity, excluding actions that are merely pro forma or half-hearted. However, one runs real risks by categorizing labor force attachment by the type of search activity, because the same activity could be highly productive for one person but not productive at all for another person. One pollibility that could make the search criterion more rigid would be to require individuals to search for a minimum number of hours. However, for the same reasons a minimum hours criterion was dismissed for the employment category, it is dismissed for the unemployment category. If someone is not employ-

ed and engaged in even one hour of honest job searching, he or she has satisfied the threshold requirement of not being employed and wanting a job. The sincerity of the search cannot be determined by either the number of hours or the type of search.

D. Discouraged Workers

Not everyone who is not employed and wants a job continues to search. Workers, who have searched for a while with no luck or those who live in areas where are very few employment possibilities, are likely to be discouraged and stop searching. These individuals would not be categorized as unemployed but, under the Greek system, as not economically active. It is appropriate to distinguish between discouraged workers and those who are actively seeking work. However, for some purposes, it is useful to know how many workers have some marginal attachment to the labor force.

One advantage to collecting data on discouraged workers is that although the unemployment rate indicates the direction in which economy is moving, it is likely to understate the magnitude of the swings, especially in a country like Greece where the rural sector is so large. When times are bad, a greater proportion of workers who become unemployed will drop out of the labor force. In general, the diversity of employment opportunities are much less in rural and semi-rural areas than in urban areas. A particular region may depend heavily on one or two industries such as forestry or mining. If these indu-

stries are adversely affected, workers have few other job possibilities to exploit. They can learn more rapidly that there are no jobs. Then, they are faced with the choice of dropping out of the labor force or incurring migration costs in order to continue their search. There is, thus, differential search cost for individuals in rural, semi-rural, and urban areas. One would expect that the rate of discouraged workers would be higher in rural and semi-rural areas than in urban areas. Thus, to obtain a measure of labor force utilization that is comparable across areas, one should include discouraged workers.

There are several problems with collecting data on discouraged workers. What one would like to learn is whether someone who is unemployed accepts an appropriate job with an appropriate salary, if one was offered. Because the answer is in response to a hypothetical situation, the answers one would get from this question would be of unknown reliability. As in the operational definition of unemployment, the desire to work could be measured by the level of one's search activity, and how long ago did job search stop. But again, the cutoff on the number of weeks that have transpired since one's last search is arbitrary. Experimentation would have to be done to determine the appropriate cutoff in Greece. In the Canadian system four levels of search acivity are distinguished embracing people who: have searched in the last week, have searched within the last four weeks, have searched in the last 6 months or have not searched in the last six months.

A discouraged worker must also be one who is not searching because he believes there are no job opportunities. In other words he is discouraged when he considers that the probability of finding a job and the wage he is likely to receive are not worth the time and monetary cost to continue his job search. Thus, a survey question would also have to inquire into why he was not currently searching. All in all, four or five questions would have to be asked to determine if someone is a discouraged worker[1]. Although one is closer to the concept of the potential labor supply, it is still not perfect. For example, students and retirees are excluded from discouraged worker categories. However, unknown proportion of school attendees and retirees are individuals who chose an alternative activity because they could not find a job.

E. Not Economically Active

Not economically active is the last category into which someone in the population could fall. Examples of not economically active individuals are discouraged workers, wo-

---

1. The U.S. identifies discouraged workers - without a six month search cutoff point - as any individual who states that he would like to work, and is not searching because: 1) he believes there is no work; 2) he cannot find work; 3) he lacks the right training; 4) employers think he is too young or too old; or 5) of the discrimination by employers or other personal handicaps.

men who decide to have children since they cannot get a job, individuals who are forced to retire early or retire after losing their job and not finding another, and youth who cannot find a job and decide to go back to school. Currently in Greece discouraged workers are not identified and, therefore, fall into this category. If the population were in steady state, one would expect the proportion of individuals who would not join the labor market for a reasonable change in wage rates, to be fairly constant. Thus, deviation of the observed proportion of not economically active around its mean could be interpreted as the change in the number of economically motivated people. However, before the derivation can be interpreted in this way, some evidence of a steady state demographic situation must be exhibited. If such a state cannot be proven, one would not know how much of change was due to the state of the economy and how much was due to an increase in, for example, women of child-bearing age, or in the number of children. Other changes, such as those in pension policies could induce individuals to retire earlier; and therefore increase the proportion of economically inactive. Thus, in order to interpret changes in the proportion of economically inactive as being caused by economically induced changes, all other factors affecting an individual's labor supply decisions must be held constant.

III. THE CURRENT SYSTEM OF COLLECTION

Many types of labor market information are best obtained from individuals so that their economic decisions can be linked to their demographic characteristics (i.e. age, education, work experience, etc.). Other data should be obtained from employers so that information about the firm (i.e. size, product, price, etc.) can be known. Greece conducts both types of survey. In the following pages we will describe and evaluate the three major sources of labor market data: the Labor Force Survey conducted by the National Statistical Service of Greece (NSSG), the employers' surveys also conducted by the NSSG and data obtained from the OAED.

A. The Labor Force Survey

The Labor Force Survey (LFS) is a household survey that is conducted to collect data on many aspects of a household's labor force participation. Between 1974 and 1980 only individuals in urban and semi-urban areas were surveyed. Information was collected over 52 weeks but aggregated into quarterly numbers. Before 1981, data were collected on: household composition, education, marital status, location twelve months ago, labor force status of each family member over 14, unemployment duration, reasons for unemployment, the branch of the principal current or prior job, and the individual's occupation. In 1981, many questions were added about hours

worked, and questions were asked about all the jobs one held, not just the principal one[1].

1. Sample Design and Coverage

The LFS is a multi-stage stratified sample as are all surveys of this kind. It is an 1.5 percent sample of the population universe distributed geographically, such that it is proportional to the number of households in an area, at the time of the 1971 Census. It is constructed in such a way that it is a self-weighting sample representative of the population over 14 years old, living in non-institutional settings (i.e. not in the military, hospital, etc.).

Before 1981, this universe included only urban and semi-urban areas. However in 1981, the population was expanded to encompass the rural areas, and the questionnaire was revised to conform to EEC recommendations, which entailed mostly adding questions. However, the increased scope of the survey was obtained at the cost of the frequency of the data. Instead of collecting the data through out the year as had been done previously, the survey was conducted entirely in

---

1. The additional data collected were: hours and days worked, whether one worked less than, equal to or greater than one's usual hours and why, overtime hours, one's occupation one year ago, what type of job an enemployed worker was looking for (full time, part-time, self-employed, temporary), the type of job search activities conducted on, whether one conducted an on-the-job search.

the second quarter (again conforming to the EEC request that the data provided to them cover the month of April). Although this may allow for better intercountry comparisons, it seriously damages the usefulness of the survey for domestic policy-making. The Greek labor market is a very cyclical one, as can be seen from Figure 1, which depicts the private non-agricultural sector's unemployment. By observing only the performance in the second quarter, one would see an overoptimistic picture. In addition, any economy in which agriculture is a major branch of the economy, as it is in Greece, must be mindful of the great seasonality of agriculture. A survey, conducted in the rural areas during planting or harvest time, will give a more incorrect estimate of annual rural labor utilization than will a survey that is spread over the year.

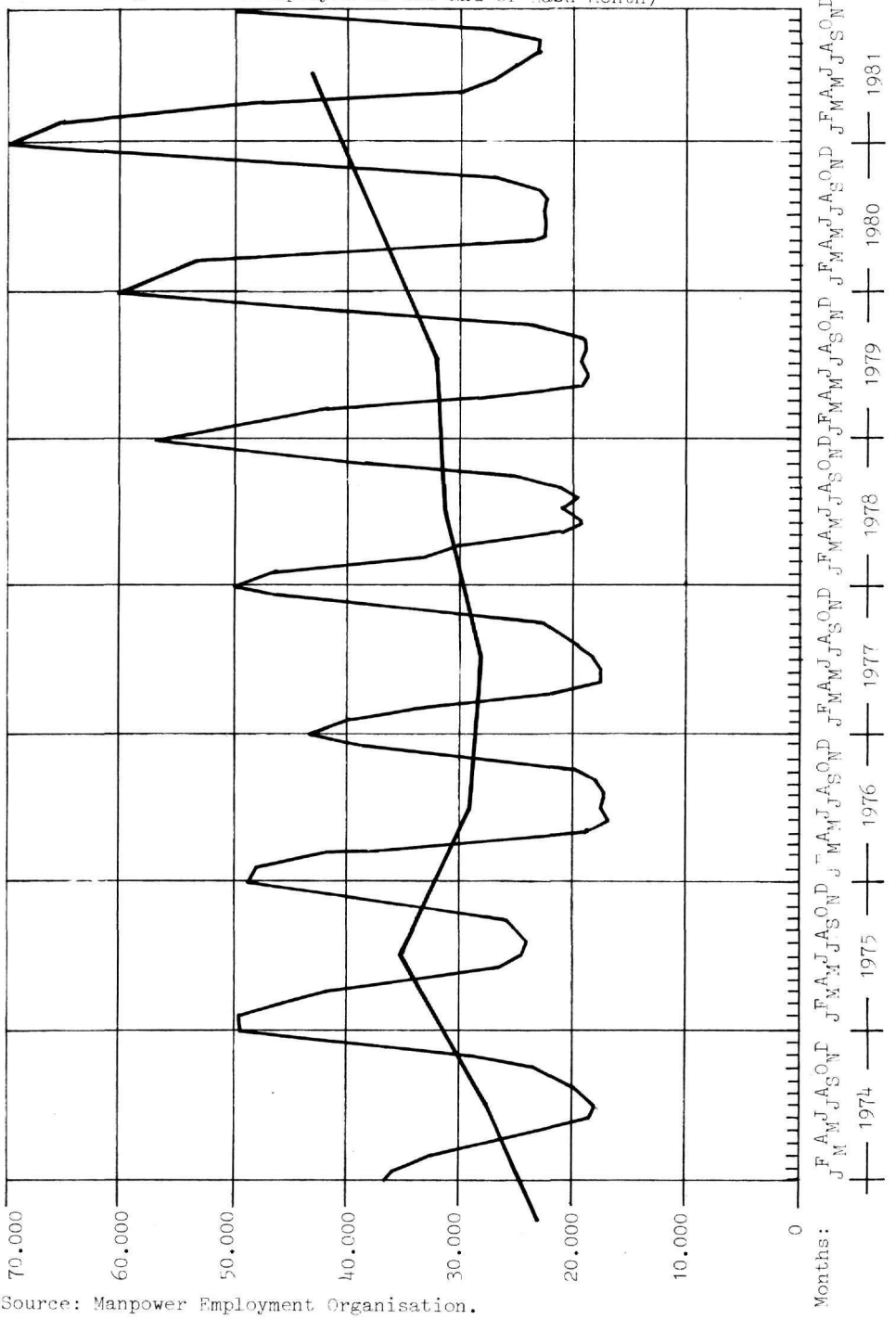
2. Accuracy of the Data

The only way to obtain an estimate of the true unemployment rate, with no possible error, is to take a census of the entire population. This, of course, would be very costly to do quarterly, or even annually. Therefore, the LFS is a sample of the population, as discussed in the previous section. The use of samples, however, introduces the possibility of error. This particular type of error is called sampling error. It arises from the fact that not everyone in the population is identical, some people are unemployed and some are employed. Thus, the estimate one obtains from one sample of people about their labor force status may not be exactly the same number that one would obtain if a different set of people were

Figure 1

The Seasonality of the Greek Labor Market

(Registered Unemployed at the End of Each Month)



Source: Manpower Employment Organisation.

surveyed. However, the more the seze of the samples grows, the more similar the answers will be. The LFS is a fairly large sample. It includes approximately 45,000 individuals. This ample size guarantees that we can be 90 percent confident that the true unemployment rate lies within a range of plus or minus two or three tens of a percentage point around the estimated rate[1]. In other words, nine out of ten times

---

1. The variance used to calculate the confidence interval must take into account that the LFS is a stratified sample rather than a simple random sample. In "Study on the Sampling Errors Associated with the 1979 Greek LFS" (by the NSSG), it was determined that the variance of the unemployment rate, calculated from the LFS, is 2.4 times larger than it would be if it was determined from a simple random sample. This number is called the design effect. The confidence interval, stated in the text, was calculated using the following formula:

$$t\sqrt{\text{Deff } (x)\ (1-x)/N},$$

where: x(1-x) is the variance of a proportion x; Deff is the design effect, here 2.4; N is the number of individuals, here 45,000; and t is the t-value associated with a 90 percent level of confidence. The confidence interval in the text was calculated under two assumptions: first, that the true unemployment rate is .02, and second that it is .06.
The calculations are:

$$1.645\sqrt{2.4(.02)\ (.98)/45{,}000=.0017}$$ (.17 percentage points)

and $$1.645\sqrt{2.4(.06)\ (.94)/45{,}000=.0029}$$ (.29 percentage points)

the rate in the population will lie in between the stated unemployment rate plus 0.3 and minus 0.3. Thus, the estimate of the national unemployment rate is fairly accurate. The unemployment rate of subgroups will be less precise, because of the size of these subgroup samples being smaller. For example, the unemployment rate among women can only be known up to a range of plus or minus fivetenths of a percentage point[1]. Regional and youth unemployment rates will also be less precise.

This discussion of sampling error is not meant to deny to usefulness of the estimates obtained from the LFS. Rather it is to remind policy-makers that small changes in the unemployment rate must be interpreted cautiously. These changes may indicate a true change in the state of the economy, or it may be that the true unemployment has not been altered and only the estimate has changed, due to sampling error. On the other hand, large changes in the national unemployment rate are also likely to indicate true changes in the state of the economy.

Appendix I discusses these issues more fully, defining what is meant by "large" and "small". Readers who are interest-

---

1. Using the same formula as in the previous foot-note and assuming the true unemployment rate among women is 9.0 percent (it was estimated to be 8.8 percent in 1981 for the urban and semi-urban areas):

$$1.645\sqrt{2.4(.09)(.92)/22,500} = .0048$$

ed in how to design an optimal survey, given cost considerations, are referred to Appendix 2.

Another factor that affects the quality of the collected data is the operational administration of the survey. The Greek Labor Force Survey is primarily administered by sixty full-time interviewers, who are employed by the NSSG and based in Athens. This administrative set up has two advantages: first, the data from different areas are more comparable than they would be if they were collected by different interviewers in different areas; second, the interviewers gain experience. Not only do more experienced interviewers reduce the number of skip logic and coding errors made, but refusal rates are lowered. Ideally, it would be best to have the whole survey done by these interviewers. However, for cost reasons, some local part-time interviewers are called in to help in the areas that are far away from Athens. To try to minimize the pollibility of differential data collection methods (and the possible comparability problem that introduces), all interviewers (full and part-time) receive training in Athens, which includes four days of field work. The NSSG might consider studying, whether or not the data seem comparable across regions. For example, one could examine the refusal rates in different areas and see if they were related to the number of part-time interviewers used.

Another operational practice, which could have an impact on the quality of the data, is the manner in which respondent substitutions are made. If an interviewer cannot find

her designated respondent at home during her first effort, she will attempt one callback. If the respondent still is not available, the household is dropped from the survey, and another household is substituted in its place. There are obvious cost reasons why this practice is followed. However, it should be remembered that by dropping difficult-to-contact households one may be introducing a bias. If difficult-to-contact households systematically differ from those that are contacted on labor market characteristics, the estimates we obtain from the LFS may be biased. For example, if the difficult-to-contact households are these in which members are more likely to be holding more than one job, then the LFS estimate of the average number of jobs held by someone, in the population or the proportion of the population who has more than one job, would be too low. The substitution practice would be to introduce a negative bias to these estimates. Thus, there is a tradeoff between cost and potential bias. The NSSG is currently trying to conduct an investigation into the size of the bias. If the bias turns out to be large, the NSSG may wish to rethink its substitution policy. It may wish to collect fewer total observations in favor of attempting more callbacks. Whether or not the bias is large or small, it is important not to allow the interviewers very much discretion concerning who they interview. There should be a rule, such as substitute the next residence to the right of the original one. This ensures that the interviewer cannot introduce a selection bias.

3. Problems and Suggestions

The most widely cited problem of the LFS is the difficulty that is inherent in making intertermporal comparisons. It is alleged that the sampling error is too large to be very useful. In Appendix I, we determine exactly how large a change must be, before we can conclude that it was due to a true change in the economy, rather than just a sampling error. This minimum difference turns out to be moderately large. The intertemporal comparisons could be made more precise by increasing the sample size, but this is quite costly. Alternately, the LFS could change its sampling technique. Instead of using an entirely new sample each quarter, some subset of individuals, who was interviewed previously, could be included again. Thus, the LFS would be a rotating sample of individuals. Because a significant portion of the sample would have been in last quarter's sample, the precision of intertemporal comparisons, which is one of the most important uses of the LFS, would be improved.

This technique of sample rotation is used in the United State's Current Population Survey, a survey that is very similar to the LFS. It has been shown that there are several cost advantages. First, a new sample does not have to be selected each time. Second, the demographic information does not have to be collected each time, so the interview is shorter on average. Third, the interviewer can find out when it is convenient for the respondent to be contacted again. Thus, there are fewer callbacks needed. Changing to a rotating

sample not only increases intertemporal precision but it lowers costs.

The rotation scheme that is suggested for a quarterly survey, such as the LFS, is one in which respondents are interviewed at least four times. They enter the sample one quarter and are interviewed that quarter and the next. They are not interviewed the next two quarters. Then, they are re-interviewed the following two. Thus, the rotation scheme is two qaurters in, two quarters out, two quarters in. Fifty percent of the sample is the same between quarters and fifty percent of the sample is the same between years. If it was decided that quarterly comparisons were not as important as annual ones, a scheme of one quarter in, three quarters out, one quarter in, three quarters out ... would improve annual comparisons but worsen quarterly comparisons. The 2-2-2 scheme treats quarterly and annual comparisons as equally important.

A serious problem, that has recently arisen, is the abandonment of a quarterly survey in favor of an annual one. As discussed in the earlier design section, this practice is likely to give a misleading view of the state of the economy. Both the extreme cyclical nature of the Greek economy and the presence of a significant agricultural sector suggest that policy-makers must be able to observe the state of the labor market more than just once a year in the second quarter. The expansion of the survey to include rural areas was an extremely worthwhile change, but it should not been brought for a less frequently administered survey. It would be better to

decrease the sample size rather to conduct the survey only once a year.

Currently the only welfare statements that can be made are about the transactions between employment and unemployment. Nothing can be said on how badly off are the unemployed and economically inactive, because no information on income is collected in the LFS. Obtaining data on the amount and the sources of household income would be very useful from a welfare policy standpoint. However, developing income questions, which obtain the desired information without offending the respondents, requires much thought and experimentation. It is a worthwhile expansion of the LFS if funds become available.

B. The Employers' Surveys

Three surveys of establishments are conducted in Greece: the Survey of Manufacturing, the Survey of Retail Establishments and the Survey of Mining Establishments. The Survey of Manufacturing began in 1962; the Survey of Retail Establishments began in 1974; and the Survey of Mining Establishments began in 1977. We will confine ourselves to discussing the manufacturing survey since it is the oldest and largest. The Survey of Manufacturing collects data on employment monthly, and quarterly data on hours, wages, overtime, and vacancies - all by sex.

1. Sample Design and Coverage

The Survey of Manufacturing is a sample of all the firms that reported having ten or more employees, during the last Census of Manufacturing, which is conducted every five years. From Table 1 it can be seen that only seven percent of all the manufacturing establishments have more than ten employees. However, this 6.5 percent employs 61 percent of the manufacturing workers. Surveying large firms is a cost-efficient way of learning about jobs, but the cost-savings is gained at the expense of representativeness. It must be kept in mind that, the data depict the situation of individuals who are employed in relatively large establishments.

Like the LFS, the manufacturing survey is a stratified randon sample. The firms found in 1973 were divided into six strata (10-19 employees, 20-29, 30-49, 50-99, 100-199 and 200+). All firms in the last stratum (200+ employees) are included in the survey on a census basis. The remainder of the sample is drawn from the other strata; the stratum sample size being inversely proportional to the stratum's standard derivation of employment. The resulting sample is not a simple self-weighting sample. In order to obtain unbiased estimates, sample observations must be weighted by the inverse of the probability of their selection. To the author's knowledge the

Table 1

Manufacturing Employment, by Size of the Firm
(September 30th, 1978)

Size of Firm by Number of Employees	Number of Firms	Percentage by Number of Firms	Total Employment	Percentage by Employment	Average Employee Size
Total	128,988	100.0	689,419	100.0	5.3
0-4	109,291	84.7	196,792	28.5	1.8
5-9	11,030	8.6	72,886	10.6	6.6
10-19	4,579	3.6	63,750	9.2	13.9
20-49	2,560	2.0	80,306	11.6	31.4
50-99	777	.6	55,216	8.0	71.1
100+	751	.6	220,469	32.0	293.6

Source: NSSG, Census of Manufacturing, 1978 Vol. II, Athens, 1981.

observations are not weighted. Thus, some bias is introduced into the estimates[1].

1. Selecting the sample in this manner is the cost minimizing way of minimizing variance. However, if it is known that the observations will not be reweighted, it is more appropriate to minimize mean squared error rather than variance to take explicit account of the bias-cost tradeoff.

2. Accuracy of the Data

The response rates to mail surveys, such as the employers' surveys, are normally much lower than in-person surveys.

Thus, one must be concerned about the possibility that those, who have the time and inclination to answer the interview, differ systematically from those who do not. The NSSG tries to minimize the problems caused by self-selection by making telephone or in-person callbacks if the firm does not respond. This greatly increases the representativeness of the sample.

The major aspect of the employers' surveys that threatens the integrity of the estimates is the out-dated nature of the sample. The samples of firms, presently included in the survey, were selected in 1973, ten years ago. At that time it was a random sample of firms that existed then, but things have changed. New firms have come into being. None of these are represented.

Firms that were in the sample may have gone out of business or changed size, so that they are no longer in their original stratum. In general, the distribution of firms has changed. Thus, it is difficult to say that the employers' surveys represent the today's employees in firms with ten or more employees today.

The NSSG takes two measures to solve this problem. First, if a firm that is in the sample ceases to exist, a re-

placement is sought by: a) examining the 1973 lists of firms in the stratum to which the dead firm belonged in 1973 and b) seeing if there is a firm in the same geographical area. If there is such a firm, it replaces the old one; if not, the old firm is dropped without a replacement, so that the sample shrinks. The sample was originally 1080 firms in 1973. It is currently 990 due to the death of firms. In addition to trying to replace firms that have gone out of existence, the NSSG adjusts the employment figures collected by it, using a factor that was developed in 1973 to account for the net effects of births and deaths. The adjustment is made to the employment index, not to the raw employment numbers.

Even with the measures that are taken, it is felt by most users of the data and the NSSG itself, that the sample under-represents small firms, because the firms that have survived since 1973 are likely to be bigger than they were in 1973 and those firms that are most likely to be in the lower strata are the new firms which are completely excluded from the sample. In addition, the replacement strategy is likely to result in the under-representation of firms in rural and semi-urban areas, since the probability of finding a replacement in these areas is much smaller.

3. Problems and Suggestions

The problems caused by changes in the composition of firms have already been discussed. There are several things that could be done to improve the situation. First, incorporating the information from the new Censuses more rapidly

would be tremendously useful; but the sample used by the employers' surveys will not incorporate the information of the 1978 Census of Manufacturing until 1983. If the sample frame and stratification schemes could be promptly changed every five years, the problems caused by using an old sample would be lessened. Perhaps tax information could be used to update the sample frame on a continual basis.

The adjustment factor should also be updated. The 1981 LFS indicated that there were 681 thousand workers employed in manufacturing and the adjusted employers' surveys number was 685 thousand. One would expect the employers' surveys to be larger, because the LFS counts people and the employers' surveys count jobs. If some people hold more than two jobs, the number of jobs should be greater than the number of employees. If the two numbers were correct, it would imply that only seven-tenths of one percent of the manufacturing workers hold more than one job. This seems too low. Besides recalculating the adjustment factor, it would be useful to use the 1978 Census to study how the new firms that came into being between 1973 and 1978 differed from the firms that were and continued to survive during that period. One could also study the characteristics of those firms that ceased to exist.

The characteristics of the firms with fewer than ten employees might also be examined. Unfortunately, it would be difficult to determine how sensitive their employment was to the economy, since one observes them only once every five years. Through such studies one could gain a better understanding of the possible biases of the employers' surveys.

The survey could also be improved by replacing the questions about quarterly vacancies' with question about the occupational composition of employment. Vacancy data is notoriously unreliable. It does not represent the desired demand for labor. Many "vacancies" are filled before they are empty without going through formal search procedures, such as placing advertisements. Although it would be nice to know employers' unfilled demand for labor, the information on vacancies that can be collected cannot be interpreted as the unfilled demand for labor, due to conceptual problems[1]. To try to get some measure of the demand for labor, more frequent occupational data should be collected. Quarterly data on occupational employed enables one to determine shifting demand requirements in the economy.

C. Data from the OAED

The OAED is one of the major dispensers of social insurance benefits. It administers the unemployment benefits for anyone who is covered by the largest social insurance system, IKA. It also operates as an employment service for anyone who

---

1. Between 1969-1973, the U.S. tried to collect vacancies data but abandoned it in 1973. It was found that firms: a) do not keep data on vacancy and b) hired many more people than they reported vacancies. In addition, frequent occurrence of internal promotion muddled the issue. A further discussion on the problems associated with vacancies is "job Vacancy Statistics" by Harry Frumerman; and the comment by Robert E. Hall in "Counting the Labor Force", Appendix 1.

wishes to use its services. Besides administering employment service programs, it collects data on those who are listed with the service. Therefore, from the OAED one can find out how many workers are employed and covered by the Social Insurance Institute (IKA) and how many people are unemployed and registered.

1. Description and Accuracy of the Data

In order to understand what the data represent, it is necessary to know a little about the Greek system of social insurance coverage. There are many agencies that provide workers with social insurance. The agency that covers an individual is determined by the branch of the economy which employs the individual. For example, the communication workers have their own system; the government workers have their own system; and banks' employees have their own system. The IKA is the largest agency covering all of the private sector. Employers must notify the IKA within eight days of hiring a new employee. It is the count of these employees that constitutes the OAED's employment numbers. To be counted as unemployed, a person must register at an OAED office. Anyone can do so; but few who do not qualify for benefits do so. In addition, not everyone who qualifies for benefits may register as unemployed with OAED because the benefits may not be worth the psychological and monetary cost of registering. To qualify for benefits, one must have worked at least three years (a minimum of 80 days per year) and have worked 125 days within the last 14 months. Hours worked in the immediately preceding two months

do not count towards the 125-day requirement. Days that were worked in non-IKA covered jobs do not count fully. Thus, new labor market entrants (youth and women) and occasional workers are not likely to show up as unemployed in the OAED data. The duration of benefits varies between two and five months, depending upon how long one has been covered. One receives in those months 60 percent of basic wage or a maximum of 600 drachmans per month. Benefits may be extended, in extraordinary conditions, for up to 40 additional days. There is an additional requirement that a benefit recipient must accept any job that the OAED locates for him or her as long as it: 1) pays the minimum salary he stipulated during the registration process, 2) is within the maximum commuting distance he stipulated and 3) is in the occupation he stipulated. Failure to accept the job means forfeit of benefits. Thus, the registered unemployed that is reported by the OAED are mostly experienced workers who felt it was worth registering. It excludes new entrants and occasional workers, as well as those who refuse OAED-found jobs or who are unemployed longer than their maximum benefit-receiving period. Again exhaustees could keep their names on the employment service list but few do so.

In 1981, the OAED data represented 42 percent of the labor force, as identified by the LFS. It constituted 63 percent of the urban and semi-urban labor force, as identified by the LFS. Unlike the LFS, the OAED data are a census of its special population, not a sample. Therefore, there is no sampling error. It accurately reports on a very special subset of the population.

Dropping exhaustees from the list of unemployed implicity assumes that someone who has not found a job in the required time (regardless of the state of the economy) is better categorized as economically inactive, rather than unemployed. A similar assumption is made for job-refusers. However, because there is no search requirement and one's willingness to take a job is only determined if the OAED finds one a job, some individuals may be incorrectly categorized as unemployed, rather than economically inactive. Offsetting this potential bias, the people who are likely to register are those who have worked for three years previously, thus, the likelihood that someone is voluntarily dropping out of the labor force is considerably reduced.

2. Suggestions

The OAED data, which are published monthly, serve a useful role as a timely indicator of the economy's and labor market's health. However, it would be useful if the OAED also published the data it collects of the occupations that unemployed seek and the minimum wages they are willing to receive. This data is currently collected, but not published. The desired occupation data could be used to discover structural imbalance or adjustment problems. The minimum wage would be useful in determining whether the unemployed expectations are just too high. It would also be useful if the data were broken down into the same nine regions used by the NSSG, rather than the currently used seven. In this way, the two data sources could be used in conjection with one another at a

regional level, as well as a national level, to get a fuller picture of the problems of the labor market.

D. The Census of Population

The Census of Population, conducted once every ten years, is useful to labor market analysts, primarily as the LFS's frame, Although the Census collects a variety on information on demographic and economic characteristics of the population in urban, semi-urban and rural areas, it is collected so infrequently that is cannot serve the many short-run needs of policy-makers. It could be useful, however, in more indepth cross-sectional studies of various types of economic behavior, such as labor force participation.

IV. COMPARING DATA SOURCES

It is bothersome and confusing when the results from different data do not agree. Usually, the discrepancy lies in the amplitude of month-to-month or quarter-to-quarter changes. These discrepancies can usually be explained by both the size of the various samples and the nature of the data sets. The different sources allow us to observe the economy from several different perspectives and to gain a deeper insight into the nature of any problem.

Three of the four data sources discussed: the Census, the LFS and the OAED data present information on the labor market in terms of individuals. The fourth, the employers' surveys, described jobs, not individuals. The Census covers the largest population. It includes any individual over the age of ten, living in any area in Greece: the urban, semi-urban, or rural areas. The LFS has the second largest universe, namely any individual over fourteen years of age. Before 1981, individuals in rural areas were excluded, but now the only difference in coverage between the Census and the LFS is the age cut-off. The OAED gathers data from a subgroup of individuals covered in the LFS, those who are employed by private firms. There is no age limit as to who can be covered by the IKA Social Insurance System; thus, there could be individuals aged then to fourteen in the OAED data that are not included in the LFS, but these are very rate.

The OAED data reflects the employment situation of the experienced, primary labor force individuals, who live in the urban and semi-urban areas. The pre-1981 LFS covered the same areas, but a much broader set of people. One would expect differences in the economic situations of these two groups of people. Through these differences, we can learn about the distribution of economic hardship between the primary (experienced) and secondary (occasional or temporary) labor force.

The key difference between the NSSG and the OAED data is that the NSSG unemployment data represent everyone who does not have a job and is looking for one in the urban and semi-urban areas, including those who are not eligible for unemployment benefits under the IKA unemployment system administered by the OAED. Figure 2 shows diagrammatically that the OAED collects data on only a small subset of the people for which the NSSG collects data. OAED data missed the large increase in the labor force in 1980 and 1981. The inability of the OAED numbers to detect the increase in the labor force provides insight into what types of individuals must comprise the expansion. They must be tsose not eligible for the social insurance, handled by the OAED and IKA. In other words, they would be the self-employed, temporary employees, and those who had not previously worked in IKA-covered jobs.

Turning to Figure 3, we see that before 1981, the majority of those, who were determined to be unemployed in the Labor Force Survey, did qualify for the OAED benefits; so they

Figure 2

Total Labor Force, According to the NSSG and the OAED Data

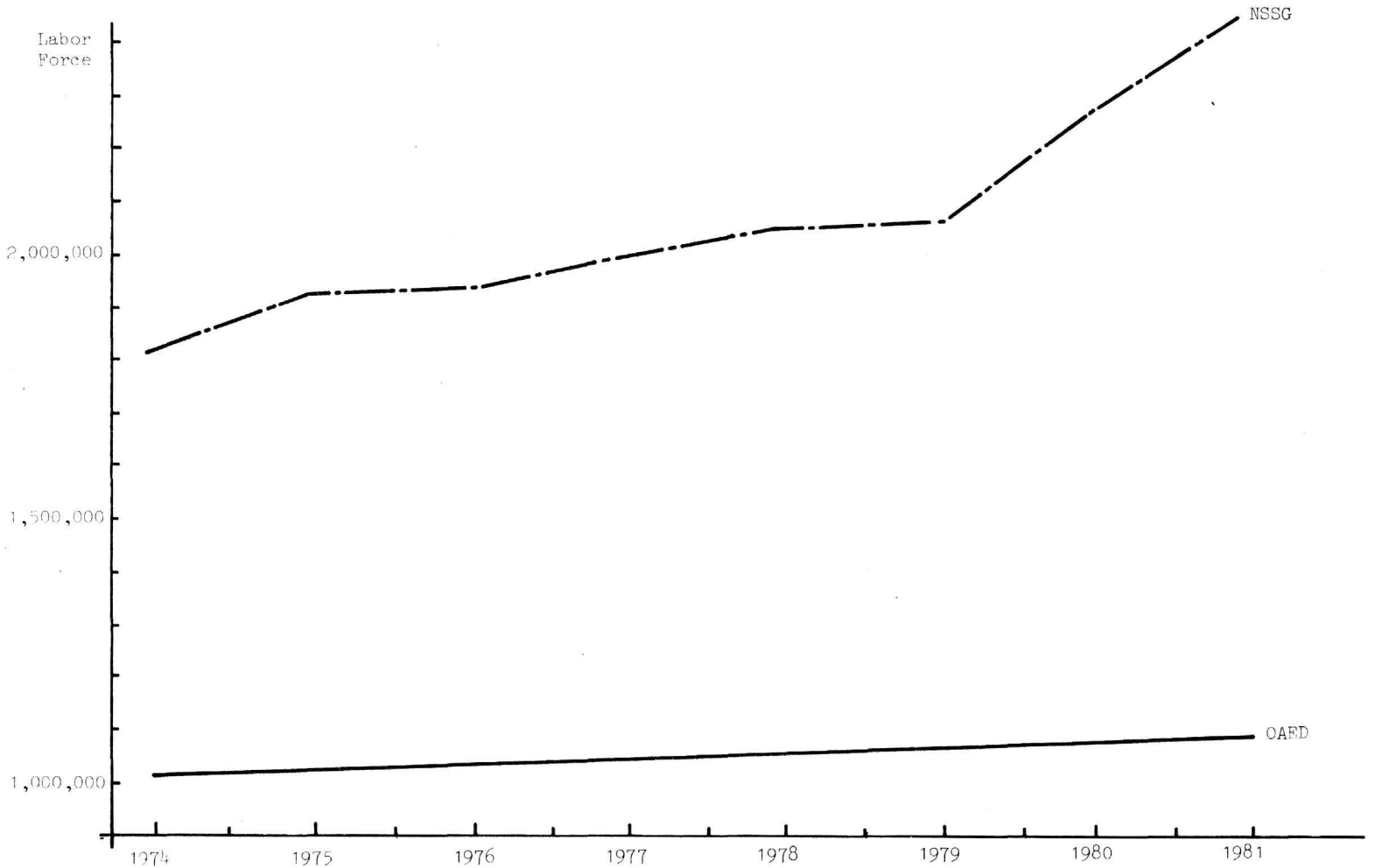
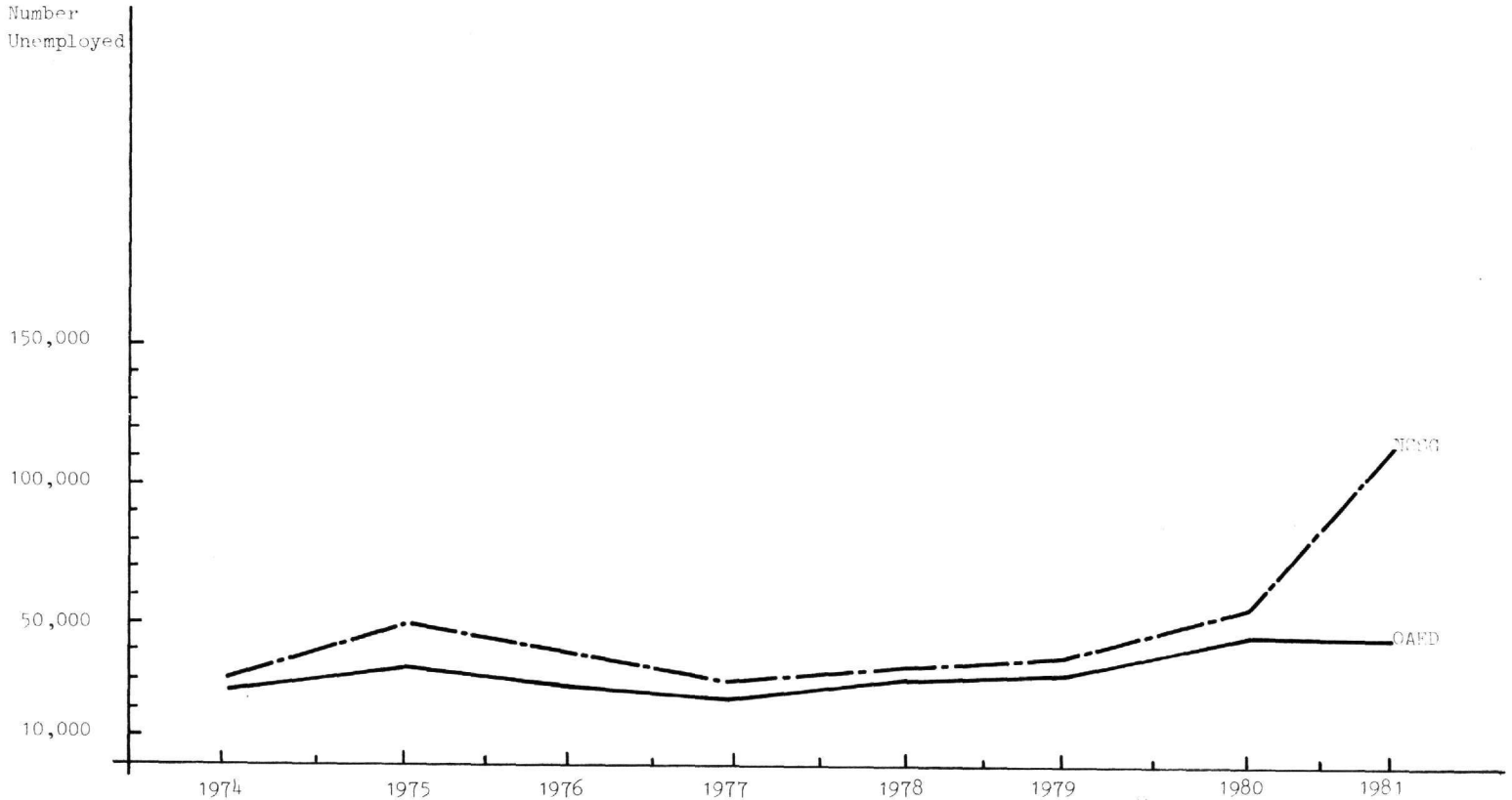


Figure 3

Total Unemployment, According to the NSSG and the OAED Data[1]
(For Urban and Semi-Urban Areas)



1. 1981 is purged of the rural component.

must have been private sector employees, not primarily youth and new entrants, long-run unemployed, or occasional workers. However, in 1981, the percentage of those who were indicated as unemployed in the Labor Force Survey and who also qualified for the OAED benefits, declined drastically, as can be seen in Table 2. In 1978, the ratio of the OAED unemployment to the NSSG unemployment was .87, and in 1981 it was .32. As one would expect, the number of unemployed persons found by the NSSG is always greater than the OAED number, and it is more sensitive to economic conditions. For example, in 1975, the ratio of the two numbers decreased. However, the serious stagnation of the 80's has caused unemployment to touch groups in the economy, previously not seriously affected, in particular youth and women.

Table 2

The Ratio of the OAED to the NSSG Data

Year	Number Unemployed	Number Employed
1974	.855	.600
1975	.700	.611
1976	.748	.642
1977	.804	.665
1978	.871	.701
1979	.835	.701
1980	.586	.685
1981	.318	.671

The unemployment rates, presented in Figure 4, reenforce this latter point. Before 1980, the unemployment rates, calculated by the NSSG and the OAED, moved together (although the NSSG rate was smaller, because they counted a larger number of people in the labor force). The OAED unemployment rate and the NSSG rates for men were particularly similar. However, in 1980, and especially in 1981, unemployment was no longer affecting primarily heads of households (what is sometimes called the primary labor force). The queues of women and youths looking for work increased greatly. By comparing the movements of these two series we are able to discover that the characteristics of unemployment have changed in the 80's. It is no longer dominated by members of the primary labor force. New entrants are comprising an increasing share.

Unlike the LFS, the employers' surveys describe the labor market as a group of jobs, rather than people. A person who holds more than one job would be twice as likely in included in the employers' surveys, than someone who only holds one job. Thus, changes in the employment figures reflect changes in moonlighting, as well as changes in basic employment. The major sample restriction is that one works in a large manufacturing firm that has survived since the last Census of Manufacturing. This restriction makes the employers' surveys fairly specialized. However, since the sectors which receive the employers' surveys (manufacturing, retail and mining) are the only sectors in which wage data are collected, the surveys allow analysts to observe the relationship between

Figure 4

Rate of Unemployment, According to the NSSG and the OAED Data
(For Urban and Semi-Urban areas)

wages and employment. Manufacturing is an important sector, but it must be remembered that it is fairly small in the Greek economy, being only nineteen percent.

V. ADMINISTRATION AND PRESENTATION

A. Coordination

There is a recently formed committee, headed by Mr. Kalambokidis which examines the degree of coordination among the various parts of the NSSG coordination. Timely exchange of information is important at all levels. For example, the employers' surveys could be made much more useful to policy-makers if the Census of Manufacturing was released more rapidly and the survey's sample frame could be updated more rapidly. Likewise, rapid processing and dissemination of some sample of the Census of Population would help policy-makers in making economic projections and in benchmarking the LFS of the corresponding year[1]. However to be useful, data must be exchanged rapidly.

Not only should there be timely exchange of the actual data, but the various parts of the NSSG and other agencies should strive to coordinate the definitions they use. For example, before 1981, the LFS's definition of employment and the Census's definition differed. The Census required someone to work in a job for ten or more hours per week. The LFS had no minimum hours standard. This difference meant that the Census and the LFS employment numbers could not be compared,

---

1. The entire Census does not have to be processed in order to get a pretty accurate picture of the population.

even if the rural individuals and those between the ages of ten to thirteen were excluded from the Census. Fortunately, the 1981 Census's employment definition is now compatible with that of the LFS.

B. Presentation and Dissemination

A cardinal rule of surveying is that every item of data that is collected should be important enough to be published. Thus, the demographic data, as well as the economic data, gathered in the LFS should be published. The data should be presented in a form that is the most useful to policy-makers. Historical continuity is very important, since most policy uses of the data involve intertemporal comparisons. Therefore, if a change in a survey is made both the information that is consistent with the old definition and with the new definition should be published, if possible. Thus, for example, the NSSG should publish all the tables it will produce from the 1981 LFS, both including and excluding the rural areas. In this way, policy-makers can use the rural-excluded sample to compare with pre-1981 trends, while knowing the rural-included figure too.

Another dissemination goal is to give all the information to all the relevant agencies. Thus, although the public may only get a smaller subset of the data (i.e. that are published in the Annual Statistical Report), policy-makers will have the additional detail that is needed many times. Currently, agencies can request to be sent various special printouts. However, many agencies do not request the full set of print-

outs and/or discard them. It is important to realize that although a current piece of information is not needed now, future analysts may want to examine a time-series of that data. It costed time and money to develop such time-series when such data existed, but were never stored or were discarded. The NSSG itself may consider publishing a time-series volume.

The LFS contains a wealth of microeconomic data. Unfortunately, few analysts can use it because of the way it is stored. Currently, the computer system used by the NSSG is incompatible with most of the other computer systems in the Government. Therefore, in order to extract data or obtain a special table, one must either wait for the NSSG Computer Center to produce the needed data table, or translate and copy the tape and do it oneself. The problem with the latter choice is that converting data between two incompatible systems is a very slow, tedious, and error-prone process. Data is bound to be lost, and once the conversion is made, time will have to be devoted to "cleaning" the new tape. The Government should seriously consider making its computers compattible with one another.

VI. SUMMARY OF RECOMMENDATIONS

Many suggestions have been made throughout this paper. All of them connot be done, given the limitation of a budget. Thus, when deciding on what should be done, one must weigh the benefit of a change to current and future poli-makers against the possible short-run problem of intertermporal non-comparability and the cost of the change.

Three changes that could be made immediately with existing data are:

1. Publish all tables from the 1981 LFS, inclusive and exclusive of the rural sector, to minimize the non-compatibility problem.

2. Publish and distribute all tables from all labor market data sources to all users of labor market data.

3. The agencies that analyze labor market data should exchange their work. Specifically, there should be more exchange of information among: the Ministries of Labor, Finance, National Economy and the KEPE.

Substantive changes that should receive serious consideration as soon as possible, are:

1. The Labor Force Survey should be reinstated as a quarterly, rather than annual survey; and the survey should use a rotating sample to improve intertemporal comparisons.

2. Income data should be collected on the Labro Force Survey.

3. The employers' surveys should be extended to more industries.

Other changes or items to be considered are:

1. Instead of collecting vacancy data, which is inaccurate, occupational data should be collected in the employers' surveys.

2. Effort should be made to make the employers' surveys more like probability samples.

3. The Government should require non-IKA insurance agencies to turn over their data on employment and unemployment to the OAED; so that a labor market on the insured population might be obtained readily.

4. The OAED should publish the data they collect from unemployed individuals, concerning the occupation and wage they seek.

5. The Government should periodically fund supplements to the LFS to study special areas of interest. One of the first studies should be an investigation of underemployment-- what questions could be incorporated into the regular LFS to get an underemployment, what are good measures of it, etc.

6. The employers' surveys should incorporate the new sample frame from the Censuses, as quickly as possible.

APPENDIX 1

Minimum Detectable Intertemporal Differences for the LFS

As discussed in the paper, all estimates derived from a sample have a possibility of a sampling error associated with them. Because of the existence of this sampling error, we can only know the true unemployment rate up to a range. The larger the sample, the smaller the range.

One can use the estimate of the unemployment rate for two purposes: 1) to describe the state of the labor market, and 2) to test whether or not the true unemployment rate has increased or decreased. For example, in 1979, the urban and semi-urban unemployment rate was calculated to be 1.8 percent. From the discussion in the paper and footnote we know that, the estimate tells us that we can be 90 percent certain that the true unemployment rate was between 1.5 and 2.1 (=1.8±.3). Likewise, the 1980 estimate of 2.8 percent tells us that we can be 90 percent confident that the true unemployment rate lies between 2.5 and 3.1 percent. It looks as if the labor market was worse off in 1980 than in 1979. However, in order to make this statement statistically sound, one must determine what is the minimum change needed to be observed, before one can say that the unemployment rate has increased or decreased. The variability of the sampling error for intertermporal changes is higher than the variablility of one year's value. In fact, if the samples used in calculating the unemployment rate

in each year are independent, the variance of an intertemporal comparison is twice that of the underlying variable, unemployment. Thus, it is hard to tell whether a non-zero difference between two unemployment rates is due to a sampling error or to a true change. The smallest change that one has to observe, before one can conclude that the unemployment rate has increased or decreased is calculated as:

$$D = (t_1+t_2)\sqrt{\frac{S^2}{N} + \frac{S^2}{N}}$$

where: D is the smallest significant change or what is called the minimum detectable difference; S^2 is the sampling variance; N is the number of people in the sample each year, t_1 is the critical t-value, associated with how confident you want to be, that you would observe a true change (i.e. the power); and t_2 is the critical t-value, associated with how confident you want to be, that if there was no change your test would tell you that. Thus, to be 90 percent confident that you would observe a true change, and 90 percent confident that your test will not falsely indicate a change in the national unemployment rate, we would have you observe a minimum change of D, before you would conclude the unemployment rate increased or decreased, where:

$$D = (1.282+1.645)\sqrt{\frac{2(2.4)(.06)(.94)}{45,000}} = .0072 \text{ or } .72$$

percentage points.

In the example we examined that the change between 1979 and 1980 is from 1.8 percent to 2.8 percent; we see that 2.8 - 1.8 = 1.0 is greater than .72. Thus, we can be 90 percent confident that indeed the unemployment rate did increase. The change between 1980 and 1981 from 2.8 to 5.5 percent is also far greater than .72 percent. We can therefore be confident that this also reflects a real worsening in the labor market.

To summarize, it is statistically easier to describe the labor market than to make intertemporal comparisons. The change in the unemployment rate has to be such, that the two estimates are farther apart than the needed, so that the two confidence intervals do not overlap. This is partly because the variation of a difference of variables is bigger than the variation of the variable, and partly because, besides worrying about how confident one can be in the description of each year's unemployment rate; one wants to have a certain level of confidence, that one will conclude there is a change if it occurs. In the case of the Greek LFS, one can be at least 90 percent confident, that the national unemployment rate is plus or minus three tenths of one percentage of the estimate. In order to make the conclusion that the unemployment rate has

changed, that it has 90 percent confidence, one must observe a change of .72 percentage points. For example, one would have to observe the unemployment rate go from 4.5 percent to 5.22 percent.

APPENDIX 2

Designing a Cost-Effective Sample

As was stated in the text and in Appendix 1, all estimates that are derived from a sample have a sampling error associated with them. The larger the sample, though, the more likely it is that the sample estimate will be close to the true value. In more technical words, the larger the sample, the smaller is the standard derivation of the sampling error. But how large a sample is necessary to best answer the major policy questions, given a budget? The answer depends on five factors:

1. How much variation is there in the variables one is trying to measure.
2. How precisely one wants to measure the key variables.
3. How confident one wants to be, that the sample would detect a policy-significant change, if one was to occur.
4. How small a policy-significant change is.
5. What is the cost of collecting the data.

If everyone in Greece worked the same number of hours, we would only need to ask one person how many hours he or she worked, in order to obtain a good estimate of the average population. If, on the other hand, some people worked a lot while others worked only very little, we would need to ask more people in order to get a good sense of the average. In order to obtain

a given level of precision in a sample estimate, one needs a larger sample if the variable differs greatly among people than if does not vary very much.

How precisely one wants to measure something will also affect the optimal sample size. If it is important to be nimety percent confident that the true population value lies within a range of plus or minus 1 percent of the sample estimate, one would need to collect many more observations than would be needed, to be just as confident that the true value lies within a range of plus or minus ten percent of the estimate[1].

Every sample is conducted for a purpose. Government surveys are usually conducted so that policy-markers can be informed about important aspects of the society. Underlying this purpose, there is an assumption that if some factor changes significantly, policy-markers will take some action. For example, the Labor Force Survey is mainly conducted to measure unemployment. There is a certain change in the unemployment rate that will trigger a policy response or, at least, policy concern. This is called a policy-significant change.

---

1. The sampling error does not decline linearly with the sample size. In other words, a sample that is twice as big does not cut the sampling error in half. Rather it varies with the square root of the inverse sample size. A sample twice as large reduces the sample error to 70.7 percent (1/2) of the old error. A sample three times as big reduces the sampling error to 57.7 percent of the original size.

An important role of a survey is to indicate when a policy significant change has occurred. Because of sampling error, one cannot be 100 percent certain that the survey will detect this policy-significant without conducting a census. However, the larger the sample the more confident one can be, that if a policy-significant change took place the survey would pick it up. In addition, giving up particular sample size, the smaller the policy-significant change is, the harder it is to detect. Thus, policy-makers must realize that in order to detect very small changes, large samples need to be collected.

Another important factor affecting the final sample size is cost. Cost influence a sample not only by imposing a maximum budget limit but by possible affecting a certain number of the various types of observations to be collected. For example, if collecting information in a rural setting is much more expensive than collecting information in an urban setting, one would be led, for cost reasons, to have a lower sampling rate in rural areas than in urban areas (assuming cost is the only way in which rural and urban settings differ). Although differing collection costs do not enter into the usual statistical determination of sample size, other than imposing a maximum budget, cost considerations should be taken into account, when designing a cost-effective sample.

These five factor all affect what is a cost-effective sample. We now have a technical discussion of exactly how one determines the sample size.

The definition of a cost-effective sample is a sample which most precisely measures the variables that one is interested in, given a particular budget. This translates statistically into minimizing the sample variance of the key policy-significant changes, subject ot a maximum budget. Technically, this is minimized with respect to the objective function:

$$\left(\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n} \right)$$

subject to the constraint:

$$c_a n_a + c_b n_b + \ldots + c_m n_m \leq B.$$

where: $n = n_a + n_b + \ldots + n_m$.

The term in parenthesis is the variance of the change in a key variable from one period of time to the next; σ_1^2 is the variance in the first period; σ_2^2 is the variance in the second period. If one thinks that the underlying variance will not change between years, then the term in parenthesis simplifies to $2\sigma^2/n$; n is the total number of observations collected in each period-- the sample size. The budget constraint

is the second equation; c_i is the average cost of collecting an observation of type i. The different "i's" could be different regions or different stages of development (urban, semi-urban, rural) or any other important budgetary disaggregation; n_i is the number of observations of type i that one collects; B is the maximum budget[1].

The solution to this constrained maximization is:

$$\frac{n_i}{n_j} = \left(\frac{\sigma_2}{\sigma_1}\right)\left(\frac{c_1}{c_2}\right)^{\frac{1}{2}} \qquad i, j = a, b, \ldots, m$$

This implies that fewer observations are collected in the costly areas and more observations are collected in areas where the varinace is greater.

---

1. If there are more than one major policy question, the objective function becomes a weighted average of variance of the policy tests, such as:

$$Z_1\left(\frac{\sigma_1^2}{n} + \frac{\sigma_1^2}{n}\right) + Z_2\left(\frac{\sigma_1'^2}{n} + \frac{\sigma_2'^2}{n}\right) + \ldots + Z_r\left(\frac{\sigma_1''^2}{n} + \frac{\sigma_2''^2}{n}\right)$$

where: Z_1 is the policy-weight, associated with the relative importance of that policy question.

One may also wish to put on an additional constraint that, nine times out of ten one wants to detect a policy-significant change of D. One can impose another degree of likelihood of observing a change; nine times out of ten, ninety-five times out of one hundred, etc. This additional constraint takes the following implicit mathematical form:

$$D = (t_1 + t_2) \left(\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}\right)$$

where: D is the minimum policy-significant difference; σ_1^2 and σ_2^2 are the variances of the key variable in the first and second period; n is the number of observations that were collected in each period; t_1 is the t-value, associated with how confident one wants to be, that one will statistically detect a change of D if it occurs, which means that if one wants to be 90 percent confident one takes on the value of 1.282; and t_2 is the t-value, associated with how confident one wants to be, that the test will not incorrectly indicate that there was a change when, in truth, there was none. With this additional constraint the problem becomes:

$$\text{minimize} \quad \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}$$

"w.r.t. n"

s.t.

(1) $$c_i n_i - B$$

(2) $$n - \left(\frac{t_1 + t_2}{D}\right)^2 (\sigma_1^2 + \sigma_2^2)$$

where $n = \Sigma n_i$. The solution of this constrained maximization operationally defines the cost-effective sample.

IN THE SAME SERIES

No 1 G. Katiforis. General Planning Guidelines. Athens 1984 (in Greek)

No 2 T. Giannitsis. International Specialization and Division of Labour between Greece and Newly Industrializing Countries. Athens 1984 (in Greek)

No 3 P. Koutsogiannis. Fundamental of Industrial Policy. Athens 1984 (in Greek)

No 4 Ch. Kamenidis, C. Oustapassidis, H. Papadopoulou, A. Semos, I. Tzimitra-Kalogianni. The Cost of Grading and Packing Peaches: Determining Factors and Means of Reducing it. Athens 1984 (in Greek)

No 5 L. Ananikas, G. Daoutopoulos, O. Iakovidou, A.Papadaki-Klavdianou, G. Siardos. Farm-Succession in Greek Agriculture. Athens 1984 (in Greek)

No 6 Th. Vakalios. Life Quality: Orientation-Directions. Athens 1984 (in Greek)

No 7 I. Kastritsi-Kathariou, V. Kiortsis. An Economic and Technical Study for the Exploitation of the Artemia – Crustacean in Greece. Athens 1985 (in Greek)

No 8 N. Polidoridis, Ch. Vossos, N. Karanikolaou, V. Pappas, Ch. Triantafillou, A. Tsamadias, G. Kontogiannopoulou-Polidoridis. University Location and Regional Development: Investigation of the Possibilities to Locate Universities Based on Regional Development Criteria. Athens 1985 (in Greek)

No 9 N. Polidoridis, A. Vernardis, Ch. Vossos, N. Karanikolaou, V. Pappas, Ch. Triantafillou, A. Tsiamadas. Urban Land Values: Investigation of Factors Forming and Affecting Land Values of the Urban Read Estate. Athens 1985 (in Greek)

No 10 J. Samaras. Problematic and Dynamic Aspects of the Public and Private Sector. Athens 1985 (in Greek)

No 11 V. Stavrinos in collaboration with M. Maragoudaki. Social Cost of Road Accidents in Greece for the Year 1980. Athens 1985 (in Greek)

No 12 G. Kokkinakis. Study on the Development of Telecommunications in Greece during the Five-Year Plan. Athens 1985 (in Greek)

No 13 Th. Georgakopoulos, J. Lofzidis. The Taxation of Property in Greece. Athens 1985 (in Greek)

No 14 N. Konsolas, in collaboration with C. Athanassopoulos, El. Sidiropoulos, C. Rokos, A. Papadaskalopoulos. A Study of Evaluation and Readaptation of the Location and Development Program of the Industrial Areas. Athens 1985 (in Greek)

No 15 V. Proïmos, G. Nikiforidis, N. Pallikarakis, A.Bezerianos, G. Panagiotakis. Proposal for a National Network of Medical Technology in Greece. Athens 1985 (in Greek)

No 16 M. Negreponti-Delivani, in collaboration with V.Portaritou-Kresteniti, J. Papadimitriou, G. Makris. D. Makridou-Boussiou, A. Karakos. Measures to Encourage Industrial Investments. Athens 1985 (in Greek)

No 17 M. Nikolinakos, G. Gekos, P. Paraskevaïdis. Development Strategy of the Cooperative Sector in Greek Agriculture. Athens 1985 (in Greek)

No 18 A. Tsaftaris, D. Roupakias. Reorganization of the Arable Crops and the Production of Protein Feed in our Country. Athens 1985 (in Greek)

No 19 G. Alogoskoufis. Macroeconomic Policy and Aggregate Fluctuations in Greece:1951-1980. Athens 1985 (in Greek)

No 20 G. Zotos. An Investigation of the Phenomenon of Quantity Surcharge in Greece. Athens 1985 (in Greek)

No 21 J. Vlahos. The Shipbuilding and Ship-Repair Industry in the Piraeus-Perama-Skaramaga Area. Athens 1986 (in Greek)

No 22 M. Moraïtou-Apostolopoulou, V. Kiortsis, G. Verriopoulos, A. Nikolaïdou. Preliminary Study of Biological Conditions of the Amvrakikos Golf: Possibilities of Aquaculture. Athens 1986 (in Greek)

No 23 G. Zacharatos. Program for the Touristic Development of the N. Plastira (Megdova) Lake Area. Athens 1986 (in Greek)

No 24 E. Theodori-Markogiannaki, P. Kavadias, D. Katochianou. Basic Data by Nomos and Region. Athens 1986 (in Greek)

No 25 G. Stefanopoulos, A. Sofos, A. Tsoukas, in collaboration with D. Paraskevopoulos, N. Sarantakos. Analysis and Planning of the Greek Petrochemical Industry. Athens 1986 (in Greek)

No 26 Ch. Georgopoulos. Study of Fiber-Optic Technology and Applications in Greece. Athens 1986 (in Greek)

No 27 G. Papayoannou. Direct Financing in Greece. Athens 1986 (in Greek)

No 28 Yannis M. Ioannides. Public Investment Financing and Social Risk Allocation. Athens 1986

No 29 G. A. Giannopoulos. Effects of Alternative Strategies for the Development of Transport Infrastructure in Greece. Athens 1986 (in Greek)

No 30 H. E. Charalabidis. Shipping and Economic Development. Athens 1986 (in Greek)

No 31 M. Konstantopoulou. Emigrant Remittances. Econometric Analysis and Forecasts for the Period 1983-87. Athens 1986 (in Greek)