

night and no matter how it started organizationally, with what kinds of authorities, with what kinds of ultimate objectives, it is going to take a considerable period of years for a capability of this kind to come into being and we are going to have to make a beginning with certain segments of the file which we feel are most critical to public policy requirements and which will have the biggest payoff. We will need to generate experience with those records and expand the system as we go along. There won't be any option to proceeding this way. I would take exception to the pilot program notion only in this respect. It seems to me, to make an effective beginning on this thing, we've got to do something more than set off some kind of disassociated pilot program in the corner and say, "Let's run it for 5 years and come back and take a look at it." There is a certain minimum threshold of getting started and that minimum threshold probably represents setting up an organization which has some kind of interagency authority and funding and staffing.

We will need this to even make this kind of beginning, to get into the problem.

Senator MILLER. Thank you, Mr. Chairman.

Chairman TALMADGE. Thank you, Senator Miller.

Our next witness is Dr. Richard Ruggles, of Yale University.

You may proceed as you see fit.

STATEMENT OF DR. RICHARD RUGGLES, PROFESSOR OF ECONOMICS, YALE UNIVERSITY

Mr. RUGGLES. During the past 2 years a heated controversy has developed over the future of the Federal statistical system. A proposal for a national data center which would pool data from many government agencies has been labeled as an "invasion of privacy" and even a "threat to democratic institutions." To put this controversy into its proper perspective, there are certain directly related questions which should be considered. First, was the data base in the past adequate for the needs which it was intended to serve? Second, what has been the impact of the computer on the processing and use of data, and how has it affected the existing situation? Third, how is the present structure of the Federal statistical system related to its functioning and to the adequacy of the data base? On the basis of the answers to these questions, it can then be asked what steps, if any, should be taken now to assure that the data system will be capable of meeting future needs fully and efficiently, without invading personal privacy.

THE ADEQUACY OF THE DATA BASE

The Federal statistical system as it now exists evolved in response to the changing forces in our economy. The first requirement, specified in the Constitution, was for a head count, which had as its major purpose the determination of representation in the Congress. The demographic census has now become one of the basic sources of economic and social data, describing not only the geographic distribution of people, but summarizing their economic and social characteristics. Most nations today consider the demographic and industrial censuses as the backbone of their statistical systems. As the economy developed, other kinds of data were created as a byproduct. Customs data yielded