INTRODUCTION

I

THIS PAPER, one of a series originating in the study of capital formation and financing in the United States, presents new estimates of residential construction, a quantitatively important and analytically distinctive component of capital formation in this country.

The classification of the paper as "technical" should not mislead the reader. Neither esoteric language nor complex tools are required to read and understand it; nor is it technical in the sense that it can interest only readers preoccupied with the measurement and analysis of residential construction. The classification "technical" merely means that the discussion is devoted largely to an explicit and detailed presentation of how the new estimates have been derived, not to substantive findings — except for some fascinating glimpses in Section 2. Analysis of the substantive findings suggested by these estimates is deferred to a prospective monograph, in which use can be made of a variety of other information bearing on residential construction and its financing.

The paper should in fact be of interest to everyone who has ever used economic records, and particularly time series. Economic time series are used widely for formulating hypotheses, testing generalizations, observing current changes in conditions, and providing bases for policy choices; and their proper use requires an understanding of how they are derived, what questions arise in the process of construction, and how these questions are answered.

II

There has been a great need for a comprehensive measure of the volume of residential construction, by type and location, over a period long enough to permit analysis of trends underlying the 14- to 20-year swings that characterize residential construction in this country. The series that have been available extend back only to 1915 and are subject to some important qualifications in the earlier years. The Bureau of Labor Statistics, however, has a large collection of basic data on construction permits reaching back to the later nineteenth century and covering a fair number of cities, and it has made these data available to our group. We have used them to construct a set of estimates covering the period from the late 1880's through 1929. This paper is largely an account of the conversion of this mass of potentially valuable raw data into a continuous, usable record.

1 The study was begun by the National Bureau of Economic Research in mid-1950. It is financed by a generous grant from the Life Insurance Association of America.
Four types of difficulty had to be faced and resolved, some satisfactorily, others with regrettable but unavoidable arbitrariness. The first type originates in the gap between a permit and a true record of construction activity. A permit is a rough and ready declaration of intention; it may lapse, and construction may never take place. Fortunately, this is a minor difficulty in most times and areas. More important is the bias in permit valuations toward omission of certain costs that should be included, e.g. architects' fees and land development costs, and toward understatement of even the costs that are included. Furthermore, if a true time series on work accomplished is to be obtained, allowance must be made for the time that elapses between the date at which a permit is issued and the date at which construction is begun, and for the distribution of work over time once construction is started.

The second type of difficulty is connected with the coverage of permit data within each city. Although only cities with a compulsory permit system are included in the sample, the possibility of under-reporting because of non-compliance had to be recognized. A related difficulty originates in the changes in the city boundaries, i.e. of the area within which the permit system is operating. For in trying to pass from the data for a varying or even a constant number of cities to some comprehensive total with a consistent definition over time, the scope of coverage of the raw data would have to be ascertained with maximum possible accuracy. Under-reporting of unknown magnitude or shifts in areas of coverage were, therefore, difficulties that had to be overcome.

The third type of problem arises in passing from the data representing a varying number of cities to estimates that relate to all urban population, in the country as a whole and preferably by regions. This, in a way, is the most crucial phase of the procedure; and the reader will find in Dr. Blank's paper a lively account of the questions faced and choices made. The interesting aspect of the procedure is the reliance on the assumed and, so far as possible, tested relation between the volume of construction and population growth in the permit cities -- a factor adequately covered by available data. It was the absence of any such known relationship and of comprehensive data upon which to apply even a conjecturally plausible one that made it impossible to do much with the raw data on nonresidential construction.

The fourth type of difficulty originates in the possibility of extending the estimates for total urban residential construction to cover construction in other areas. Dr. Blank has made this extension by assuming that (1) population growth in nonfarm rural areas can be considered a dominant factor in determining the volume of new housekeeping residential construction in those areas, and (2) permit data for the smaller cities can be used to infer the changing relation over time between new nonhousekeeping residential construction and population growth in nonfarm rural areas. The decision to extend the estimates to total nonfarm residential construction (rather than urban alone) necessarily involved further questions -- e.g. the allowance to be made for differences in
value per dwelling unit between rural and urban areas, and the adjustments to be made, in calculating changes in rural nonfarm population, for the effects of the shift of some nonurban areas into urban limits from one population census date to another.

III

Dr. Blank's account of how the four types of difficulty were resolved is revealing, both as to the many respects in which the situation described is typical of a vast number of other economic time series, and as to the several aspects that seem characteristic of residential construction alone. Like many other widely used economic time series, the primary data used here originate as by-products of administration. This means that there may be gaps in accumulation and errors in reported magnitudes, and that the coverage, either in space or over time, may be affected by differences in legislation and compliance. Typical also is the accumulation of the primary records in scattered and inaccessible depositories — to be mobilized at some focal place only under pressure of some nationwide problem or by some control agency aware of their value. Quite typical also in the story is the dependence of scholars upon the Government for the supply and treatment of the primary data. We all recognize that, despite the ingenuity of Dr. Blank and his colleagues, little could have been accomplished if the government had not used large resources in the late 1930's to assemble the data, and if the Bureau of Labor Statistics had not given great help in classifying the data prior to their tabulation. Popular prejudice to the contrary, the bricks of statistical estimates cannot be made out of straw alone: they require the clay of primary data and the molding effort of intelligence and hard work.

The specific characteristics of residential construction emerge clearly in Dr. Blank's account: the localized character of the industry and the corresponding multitude and variety of jurisdictions covering the enforcement of regulations and the collection of data; the large number of small producing units in the field; the lack of standardization of the product and the substantial quality changes over time; and the fact that not until recently did the field become a matter of concern for federal agencies, either as to policy or as to data collection. Similar obstacles to the securing of adequate, countrywide, consistent series may exist in some other fields. For residential construction they loom all the larger because of the need for a long statistical record to permit the analysis of movements that underlie the long cyclical swings characteristic of the industry.

The relevance of these difficulties in securing adequate data for an analysis of real capital accumulation in this country need not be stressed. That much effort and ingenuity had to be applied is obvious enough. It is also clear that several arbitrary decisions had to be made; that the estimates should be used only with proper attention to the assumptions that underlie their derivation;
and that the results are subject to improvement as the recent, more detailed data accumulate.

While Dr. Blank's work results in the revision of some older estimates, it would not have been possible without the contribution of past work. In a sense, this indebtedness to the work of earlier investigators is the most typical theme of the story, and is eloquently illustrated by the numerous references in the paper to Wickens, Chawner, Long, Newman, the Bureau of Labor Statistics, the Department of Commerce, and others. Earlier work was indispensable for the attainment of whatever advance to knowledge the new estimates represent; and they, in turn, should prove indispensable for further advance. While clearly of importance for our inquiry into long-term trends in real capital accumulation, they are even more useful as an additional stone in the structure of growing knowledge in the field, and to that extent, of the economy at large.

SIMON KUZNETS