The 1982 Characteristics of Business Owners Dataset*

Richard Boden
Alfred Nucci

ABSTRACT. The 1982 Characteristics of Business Owners (CBO) dataset is a microdata file containing data on approximately 91,000 small businesses and 100,000 owners of those businesses. This dataset is unique in that it unites information on the personal attributes of business owners with information on various attributes of their businesses. The CBO dataset is also distinguished by the large volume of data that it contains on minority and female business owners and their businesses. This paper provides an overview of the coverage and contents of the dataset, and the design of the survey upon which it is based. In addition, the strengths and limitations of the CBO dataset are discussed and illustrated through synopses of studies that made use of these data.

I. Introduction

The 1982 Characteristics of Business Owners (CBO) dataset is one of the most comprehensive sources of data on the characteristics of both businesses and their owners.[1] Before the 1982 CBO cross-sectional microdata became available, researchers interested in such issues as the relationship between individual attributes and entrepreneurial performance traditionally drew upon household-based datasets (none of which was designed to support such inquiries), or upon datasets developed from surveys conducted by institutions representing small business interests.[2] Moreover, data on self-employed individuals — especially self-employed women and minorities — are scarce in household-based surveys because self-employment is a relatively rare phenomenon. In contrast, female and minority business owners (and their businesses) are, by design, well represented in the 1982 CBO.

This paper describes various aspects of the dataset, including coverage, contents, and survey design. The CBO's analytical strengths and limitations are documented and illustrated through discussion of two studies that made use of the data.

II. Coverage

The 1982 CBO dataset contains microdata on 91,000 businesses and approximately 100,000 owners.[3] This sample represents approximately 11 million businesses and 14 million owners in a proper subset of the U.S. business population as defined by the Internal Revenue Service (IRS) business tax universe — viz., sole proprietorships, and partnerships and 1120S corporations with fewer than 10 partners or shareholders, respectively. Excluded from the CBO universe are all regular corporations, as well as partnerships and 1120S corporations with 10 or more common owners.

To put the coverage of this dataset into perspective, Table I contrasts the segment of the U.S. business population that is represented in the 1982 CBO with the entire U.S. business population in 1982. Some simple arithmetic reveals that, excluding regular (1120C) corporations, the businesses represented in the 1982 CBO dataset comprise approximately 83% of the businesses in the Internal Revenue Service (IRS) tax universe, and account for about 70% of those businesses'

Final version accepted on June 1, 1990

Richard Boden
U.S. Small Business Administration
1441 L Street
NW Washington, DC
U.S.A.

and

Alfred Nucci
U.S. Bureau of the Census
Room 1587-FOB 3
Washington, D.C. 20233
U.S.A.

Small Business Economics 2: 301—305, 1990.

TABLE I
Number and receipts of businesses represented in the 1982 Characteristics of Business Owners dataset contrasted with the entire 1982 U.S. business population

Legal form of organization	Number of businesses	Sales ($millions)
Sole proprietorships		
IRS Universe	10,105,515	433,665
CBO Universe	9,338,523	368,544
Partnerships		
IRS Universe	1,514,212	251,609
CBO Universe	488,963	100,618
1120S corporations		
IRS Universe	564,219	235,011
CBO Universe	320,690	178,248
All corporations		
(IRS Universe)	2,925,933	6,156,994

Source: Special tabulations of 1982 Characteristics of Business Owners data by the Center for Economic Studies, U.S. Bureau of the Census, and Internal Revenue Service (IRS), *Statistics of Income, 1982* (various issues).

sales. Adding regular corporations, the businesses represented in the 1982 CBO account for 67% of all U.S. businesses in 1982, and 9% of all business receipts.

III. Sample design

The CBO dataset represents the product of a two-phase survey. Undertaken first were the Survey of Minority Owned Business Enterprise (SMOBE) and the Survey of Women Owned Business (SWOB). These quinquennial surveys collect information on the gender and ethnic status of business owners. The information is appended with administrative records data on various business characteristics (such as annual receipts, payroll, and geographic and industrial location). A well-defined subset of the 1982 SMOBE and SWOB responses served as the sampling frame for the 1982 CBO survey.

The CBO "sample" is comprised of five stratified, random samples — of approximately equal size — called "panels."[4] The panels, which correspond to classifications of businesses by their owners' demographic characteristics, are: Hispanic, other minorities (primarily Asian), Black, women and nonminority men.[5] Businesses' two-digit SIC code and geographic (state) location served as sampling strata. The survey was designed to ensure ample collection of data on female, minority, and nonminority male business owners.

Each CBO record contains two sampling weights, an owner weight and a business weight, which reflect the size of the respective populations of owners and businesses represented by the observation. The weights are adjusted for survey nonresponse, are panel-specific, and allow calculation of population estimates.

IV. Contents

The 1982 CBO dataset contains two basic classes of data: data obtained from the CBO and SMOBE/SWOB survey instruments, and data obtained from Internal Revenue Service business income tax records.[6] The CBO survey instrument contains 25 basic questions on owner characteristics (such as education, hours worked per week, and marital status), and business characteristics (including business start date, demographic composition of employees, and net income).[7]

All of the data obtained from the CBO survey instrument are categorical. Response data for questions concerning quantitative variables (such as net income or the amount of money individuals used to become business owners) are in interval form. Data on receipts, payroll, and number of paid employees are in the form of point values. The presence of data obtained from the survey instrument depends upon survey response (case response) and response to corresponding questions on the survey instrument (item response). By contrast, all data obtained from the SMOBE/SWOB surveys and IRS tax returns are present for every CBO observation.

V. Strengths and limitations

The principal strengths of the 1982 CBO dataset are its linking of business and owner information and its volume of data on female and minority business owners. This dataset also contains a variety of data elements not found in any other single data source, including:

(a) business entry capital — i.e., the amount of individuals used to become business owners;

(b) the percent of business entry capital that was borrowed;
(c) identification of sources of both borrowed and equity business entry capital;
(d) number of years that owners owned their businesses;
(e) number of years of owners' prior, paid managerial and nonmanagerial (wage) employment experience;
(f) indicators of business forfeiture and dissolution.[8]

These data elements provide a basis for empirical research on entrepreneurship that cannot be entertained by other datasets.

The CBO dataset has some aspects that limit its usefulness. Many of these aspects are shared by other cross-sectional, survey-based datasets. Others are more peculiar to the 1982 CBO dataset and worthy of brief discussion.

One of the more troubling aspects of the dataset concerns the possible presence of bias in the response data. The CBO survey instrument was mailed in 1986 to businesses that responded to the precursor SMOBE/SWOB surveys in 1985. The SMOBE/SWOB surveyed businesses that were known (by virtue of their filing of tax returns) to be in existence in 1982. Without question, some of these businesses ceased operation between 1982 and 1985. Unfortunately, there is no practical way of assessing the extent to which businesses that ceased operation between 1982 and 1985 are over-represented in the group of businesses excluded from the CBO survey sampling frame (by virtue of nonresponse to the SMOBE/SWOB surveys).[9]

Also, some of the response data in the CBO dataset may be unreliable. The CBO survey instrument — which was mailed in 1986 — asked owners to provide retrospective answers regarding the characteristics of themselves and their businesses as they prevailed in 1982. For example, owners were asked in 1986 questions concerning how many hours per week they spent managing or operating their businesses in 1982, and what percent of their employees in 1982 were women. Many owners may not have an accurate recollection of such information four years later.[10]

Finally, by definition, there is no way to use the CBO dataset (alone) to model the processes by which individuals select into business ownership and, hence, into the CBO sample, which pertains to business owners, exclusively. This limits the strength of inferences (using these data) that can be made about, for example, the causal relationship between individuals' human capital and their success as business owners.[11]

VI. Applications of the 1982 CBO dataset

The CBO dataset can entertain inquiries with either the business or the owner as the unit of analysis.[12] This section highlights two studies that are representative of the analyses for which this dataset is well suited.

Using the 1982 CBO dataset, Timothy Bates (1989b, 1989c) set out to identify the reasons why businesses owned by White (nonminority) men "out-perform" those owned by Blacks in terms of two basic criteria — annual receipts size and survival rate.[13] In the course of this research, Bates uncovered empirical regularities concerning the relationship between business performance and owners' human capital, initial financial investment in their businesses, and the debt/equity composition of that investment. Moroever, Bates found that Black/White business performance differentials can be partially traced to differences between Black and White owners' average level of education and financial scale of entry into business ownership.

Bates estimated a single-equation regression model — separately for Black- and White-owned businesses — with annual receipts as the dependent variable. The model specification is straightforward: output (receipts) is expressed as a function of inputs and a set of control variables. Regressors include measures of businesses' productive inputs (hours worked per week by owners and business entry capital), a set of owner attributes (including education, age, and previous paid managerial experience), industry variables, and a variable identifying businesses that were started or acquired by their owners between 1980 and 1981.

Bates found certain variables to be important determinants of annual receipts for both Black- and White-owned businesses. For both groups, the estimated influence of businesses' productive inputs were found to be positive and significant — businesses' annual receipts were found to be an increasing function of hours worked by owners and the size of owners' original financial investment in their businesses. Also, for both Black- and White-owned businesses, Bates found that the annual receipts of businesses owned by individuals with four or more years of college education were higher than the receipts of businesses owned by lesser educated individuals, all other things equal.

Bates notes that, at least in his sample data, Blacks used substantially less money (than White men) to enter business ownership, and were less likely to have four or more years of college education. Thus, these two variables account for some of the difference between the annual receipts size of Black- and White-owned businesses.[14]

Using discriminant analysis, Bates identified a key set of factors associated with exit from business ownership. Here, too, Bates' findings for Black-owned businesses parallel his findings for businesses owned by White men. In particular, for both Black and White owners, the tendency to remain in business (over the 1982—1986 period in question) was found to be positively associated with higher levels of education and higher levels of business entry capital.

Using cross-tabulations of CBO data pertaining to sole proprietorships, Faith Ando (1988) compared the attributes of White (nonminority) business owners and their businesses, with those of American Indian, Asian, Black, and Hispanic sole proprietors and their businesses. Within each of these five ethnic groups, Ando also compared owner and business attributes between male and female sole proprietors.

Ando's tabulations yielded some interesting findings and clues to fruitful directions for further research. One of her more notable findings is that gender differences in the characteristics of business owners and their businesses vary substantially with ethnicity. On the basis of this finding, Ando suggests that analyses of gender differences in business owners and businesses are best performed within ethnic groups — i.e., nonminority women should be contrasted with nonminority men, Black women should be compared to Black men, etc.

VII. Conclusion

The success of the 1982 Characteristics of Business Owners dataset has resulted in the funding of a 1987 CBO survey. While the dataset based upon the 1987 survey is under development, the 1982 CBO dataset will itself be the focus of data enhancement projects. One of these is a match of 1982 CBO microdata with 1987 SMOBE/SWOB data. The product of this project will be an augmented 1982 dataset that can be used to examine such issues as the relationship between business owner attributes and business survival and growth.

Notes

* Economist, U.S. Small Business Administration, Washington, D.C. The views expressed in this paper are the authors and do not necessarily represent those of the U.S. Small Business Administration.

1 The 1982 Characteristics of Business Owners data series was produced and is maintained by the U.S. Bureau of the Census, and was jointly funded by the Minority Business Development Agency (MBDA) of the U.S. Department of Commerce and the U.S. Small Business Administration.

2 Most notable of these household-based surveys are the decennial Census of the Population and the monthly Current Population Survey (CPS) and Survey of Income and Program Participation (SIPP). It should be added, however, that private-sector institutions representing small businesses, such as the National Federation of Independent Business (NFIB), have produced very useful datasets based upon surveys of their membership. These datasets have spawned some important empirical studies on entrepreneurship, such as that of Dunkelberg and Cooper (1990).

3 Overall, about 80% of the targeted business owners responded to the 1982 CBO survey. Survey response rates vary by panel, however, with the nonminority male panel exhibiting the highest response rate (84%) and the Black and Hispanic panels exhibiting the lowest response rates (75%). Among responding owners, response rates for individual survey questions are generally quite high — typically in excess of 90%. For further information on survey and item response rates, see Nucci (1989).

4 Although the business served as the sampling unit, all common owners of businesses were surveyed. About 25,000 owners in each of the five panels were surveyed.

5 The demographic classification of a business is based upon the demographic status of 50% or more of the business' owners. In cases of uniform divisions of owners across demographic groups, the demographic classification of the business proceeds according to a hierarchical scheme developed by the U.S. Bureau of the Census.

6 The SMOBE/SWOB survey data provided identification of owners' ethnicity and gender. IRS data provided identification of businesses' legal form of organization, industry and geographic classification, number of owners, annual receipts, payroll, and number of employees.

[7] A copy of the CBO survey instrument can be found in Nucci (1989).

[8] The 1982 CBO survey instrument was mailed in 1986 to owners of businesses known to be in existence in 1982. One question asked owners whether or not they still owned (in 1986) the businesses that they owned in 1982. Another asked whether or not the businesses that they owned in 1982 were still operating.

[9] The same problem arises if businesses ceased operation between the mailing of the SMOBE/SWOB survey instruments in 1985 and the mailing of the CBO survey instrument in 1986. However, if businesses responded to the SMOBE or SWOB surveys, then records corresponding to those businesses in the 1982 CBO will have data for data elements extracted from administrative records (such as annual receipts, number of employees, annual payroll, etc.), even if they did not respond to the CBO survey. These data can be used to compare certain attributes of businesses that responded to the CBO with those of nonresponding businesses, or to model the determinants of CBO response status.

[10] One instance of suspect response data — that is not necessarily attributable to the lag in the mailing of the CBO survey instrument — warrants explicit mention. The only data on business profits (net income) in the 1982 CBO dataset were obtained from responses to a survey instrument question. A large number of respondents interpreted this question as asking for liabilities and assets. These respondents provided positive and negative net income data. Item nonresponses were coded for these cases.

[11] For instance, individuals with higher-than-average levels of human capital most likely accrue higher-that-average wage earnings and may thus require relatively high (expected) self-employment earnings to induce them into business ownership. Thus, if analyses of the CBO data reveal that individuals with comparatively high levels of human capital (as measured by education or years of paid employment experience, for example), the extent to which this is merely a manifestation of the underlying entrepreneurial selection process (as opposed to evidence of a direct relationship between human capital and entrepreneurial performance) is indiscernible.

[12] Each record in the 1982 CBO corresponds to an individual business owner. Business-based analyses of multiple owner businesses, then, might involve some sort of aggregation of owner attributes across records. Common owners of firms can be linked by means of the Census File Number (CFN) which is part of each record.

[13] In the sample of businesses at the focus of Bates' research, the average annual receipts of Black-owned businesses are only about 57% of those businesses owned by White men, and Black business owners were about 5% more likely (than their White male counterparts) to have relinquished ownership of their businesses sometime between 1982 and 1986.

[14] Bates also notes that — insofar as it affects their ability to borrow start-up capital — Blacks' lower average level of education and lesser initial financial investmet in their businesses may not be independent phenomena. Much of Bates research — including portions of the research highlighted in this paper — has focused on the elucidating the reasons *why* the financial scale at which Blacks enter business ownership is lower than that of Whites. Bates' contributions in this area are beyond the scope of this paper, however.

References

Ando, Faith, 1988, 'Minorities, Women, Veterans and the 1982 Characteristics of Business Owners: A Preliminary Analysis', unpublished report prepared for the Office of Advocacy, U.S. Small Business Administration, under SBA contract SBA-3026-OA-88.

Bates, Timothy, 1989a, 'Financial Capital and Small Business Startups', unpublished report prepared for the Office of Advocacy, U.S. Small Business Administration, under SBA contract SBA-3026-AER-88, the Ford Foundation, and the National Science Foundation (NSF Grants SES84-01460 and SES87-13643).

Bates, Timothy, 1989b, 'Entrepreneur Factor Inputs and Small Business Longevity', Bureau of the Census, Center for Economic Studies Discussion Paper No. 89—4.

Bates, Timothy, 1989c, 'Entrepreneurial Human Capital Inputs and Small Business Longevity', *Review of Economics and Statistics* (forthcoming).

Bates, Timothy, 1989d, 'Small Business Viability in the Urban Ghetto Milieu', *Journal of Regional Science* (forthcoming).

Bates, Timothy, and Alfred Nucci, 'An Analysis of Small Business Size and Rate of Discontinuance', *Journal of Small Business Management* (forthcoming).

Dunkelberg, William, and Arnold Cooper, 1990, 'Investment and Capital Diversity in the Small Enterprise', in Zoltan Acs and David Audretsch (eds.), *The Economics of Small Firms: European Challenge*, Dordrecht: Kluwer Academic Publishers, pp. 119—134.

Nucci, Alfred, 1989, 'The Characteristics of Business Owners (CBO) Database', Bureau of the Census, Center for Economic Studies Discussion Paper No. 89—9.

U.S. Bureau of the Census, 1971, *Minority-Owned Businesses: 1969*, Washington, D.C.: Government Printing Office.

U.S. Bureau of the Census, 1980, *Selected Characteristics of Women-Owned Businesses 1977*, Washington, D.C.: Government Printing Office.

U.S. Bureau of the Census, 1986a, *1982 Survey of Minority-Owned Business Enterprises, Asian American, America Indians, and Other Minorities*, Washington, D.C.: Government Printing Office.

U.S. Bureau of the Census, 1986b, *1982 Survey of Minority-Owned Business Enterprises, Hispanic*, Washington, D.C.: Government Printing Office.

U.S. Bureau of the Census, 1986c, *1982 Economic Censuses, Women-Owned Businesses*, Washington, D.C.: Government Printing Office.

U.S. Bureau of the Census, 1985, *1982 Survey of Minority-Owned Business Enterprises, Black*, Washington, D.C.: Government Printing Office.

U.S. Bureau of the Census, 1987, *1982 Characteristics of Business Owners*, Washington, D.C.: Government Printing Office.