Gathering and Imputing Data about the Family and Business

This paper discusses the rationale and challenges of collecting financial data from the household and the business and the imputation strategies employed for missing data in a national study entitled, "Family Businesses: Interactions of Work and Family Sphere."

George W. Haynes, Montana State University[1]
Sharon Danes, University of Minnesota[2]

Gathering Data

Small business are typically family-owned and operated businesses, where family and business resources are often intermingled. Research on small businesses generally disregards the intermingling of resources and attempts to assess the impact of changes on the business only. This data collection effort bridged this substantial gap by designing a survey instrument that carefully assesses the management, marketing and financing structure of both the family and business entities.

The business financial information was more challenging to gather than the household financial information. Household and business managers living on a farm/ranch had difficulty separating the financial resources between the farm/ranch business and household entities. Some business managers were reluctant to reveal the size (gross sales) or profitability of the firm. Other respondents were not familiar with accounting terminology. In several cases, the contributions to the household reported by the household and business managers differed substantially. While this data set presents a substantial imputation challenge for financial related questions, these types of data collection problems are not unique to this data set.

Missing Data Assessment and Imputation Strategy

Sample size, degree and characteristics of the missing data and randomness of that missing data are considerations when deciding questions about the handling of variables with missing information. The performance of any missing-data method depends on the mechanisms that lead to missing values (Little & Rubin, 1987; Paulin & Ferraro, 1994) Three types of mechanisms are possible: missing completely at random (MCAR), missing at random (MAR), or nonignorable nonresponse. Each of these response mechanisms has implications for the strategy to be used for addressing missing-data problems.

The first level of imputation will utilize the standard accounting identities. A second level of imputation will employ methods proposed by Little and Rubin (1987). Stochastic regression will be used in cases where most of the financial data is reported. In other instances, where most of the financial data is missing a cold deck method will be employed using the Survey of Consumer Finances for household financial data and National Survey of Small Business Finances for business financial data. Given the problems in separating business and household finances, separate populations weights are necessary for the farm/ranch businesses and households so the analyst has the choice of either including or excluding these observations. This carefully imputed survey will enable the co-mingling of household and business resources to be described and analyzed.

References

Little, R.J.A. & Rubin, D.B. (1987). Statistical analysis with missing data. New York: John Wiley & Sons.
Paulin, G.D. & Ferraro, D.L. (1994). Imputing income in the Consumer Expenditure Survey. Monthly Labor Review, 117(12), 23-31.

Endnotes
1.   Assistant Professor, Department of Health and Human Development
2.   Professor, Family Social Science Department