ELSEVIER

Journal of Econometrics 107 (2002) 235–257

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Information theoretic measures of the income distribution in food demand

J.T. LaFrance[a,*], T.K.M. Beatty[a], R.D. Pope[b], G.K. Agnew[c]

[a] *Department of Agricultural and Resource Economics of Economics and Giannini Foundation of Agricultural Economics, University of California, Berkeley, CA 94720-3310, USA*

[b] *Department of Economics, Bringham Young University, Provo, UT 84602, USA*

[c] *Department of Agricultral and Resource Economics, University of Arizona, Tucson, AZ 85721, USA*

Abstract

A new method to nest, estimate and test the rank and functional form of the income terms in an incomplete system of demand equations is developed. Information theory is employed to infer the U.S. income distribution from data on quintile and top five percentile income ranges and intra-quintile and top five percentile mean incomes. Maximum entropy income distributions are combined with data on the U.S. demand for 21 food items to estimate U.S. food demand over the period 1919–1995, excluding 1942–1946.

JEL classification: D12; E21

Keywords: Functional form; Incomplete demand systems; Information theory; Maximum entropy; Rank of demand models

1. Introduction

This paper exploits the richness of incomplete demand systems to extend aggregation in nonlinear functions of income to incomplete demand systems (hereafter IDS). We develop methods to completely nest weakly integrable versions of the almost ideal demand system (AIDS), the linear approximate AIDS (LA-AIDS), the quadratic AIDS (QAIDS), the price independent generalized linear (PIGL)—including the price independent generalized logarithmic (PIGLOG) form, the quadratic PIGL (QPIGL), an extended version of the linear expenditure system (LES)[1] and an extended quadratic

* Corresponding author. Tel.: +1-510-643-5416.

E-mail address: lafrance@are.berkeley.edu (J.T. LaFrance).

[1] "Extended QES" indicates that *supernumerary income* is income minus a quadratic form in prices and that there is an $n \times n$ matrix of price effects in addition to the intercepts in the QES demands.

0304-4076/02/$ - see front matter
PII: S0304-4076(01)00122-1

expenditure system (QES). These methods permit us to test for and estimate the rank and functional form of the income terms in aggregable incomplete demand systems.

Rank three demand models require at least three summary statistics from the income distribution, e.g., for a QPIGL model in expenditure form we need the cross-sectional means of $m_h^{1-\kappa}$, m_h, and $m_h^{1+\kappa}$, where m_h is the income level of family h, $h=1,\ldots,H$, say, and κ is the PIGL coefficient on income, while for a QAIDS model we need the means of m_h, $m_h \ln(m_h)$, and $m_h[\ln(m_h)]^2$. To calculate these means, we need information on the distribution of income. The U.S. Bureau of the Census annually publishes the quintile ranges, intra-quintile means, top five percentile lower bound for income, and the mean income within the top five percentile range for all U.S. families. We use information theory to obtain annual maximum entropy income distributions that satisfy each of the intra-range percentile and conditional mean conditions for the period 1910–1999.

These estimates of the income distribution are combined with aggregate annual time series data on per family U.S. food expenditures for 21 individual food items over the period 1919–1995, excluding 1942–1946 to account for the structural impacts of World War II.[2] In addition to annual measures of food expenditures, prices, and the income distribution, we incorporate measures for the distribution of the U.S. population by age and the ethnicity of the U.S. population in the IDS specification. The results of the empirical application strongly suggest that a full rank three model is essential, and that all forms of the AIDS–IDS model are rejected in favor of a QPIGL–IDS that is nearly an extended QES.

The rest of the paper is organized as follows. The next section extends the aggregation results of Gorman and others to incomplete demand systems that can be written in a PIGL or PIGLOG form. The third section describes the implementation of the maximum entropy procedure to estimate the U.S. income distribution. Section 4 presents a summary and discussion of the empirical results, focusing primarily on the rank of the demand model and the functional form of the income terms. The final section summarizes and concludes. Proofs of our main results and additional discussions and derivations of the modeling approach are contained in an expanded paper (LaFrance, Beatty, Pope and Agnew (2001), hereafter LBPA), which is available from the authors upon request.

2. A flexible aggregate incomplete demand system

It is a generally accepted proposition that a modern demand system should be sufficiently flexible to represent a rich set of qualitative and quantitative behaviors. Clearly,

[2] See LaFrance (1999a, b) for empirical evidence for the exclusion of World War II and the stability of U.S. food demands over this sample period. The 21 food items included in the data set can be conveniently grouped into four categories: (1) *dairy products*, including fresh milk and cream, butter, cheese, ice cream and frozen yogurt, and canned and dried milk; (2) *meats*, *fish and poultry*, including beef and veal, pork, other red meat, fish, and poultry; (3) *fruits and vegetables*, including fresh citrus fruit, fresh noncitrus fruit, fresh vegetables, potatoes and sweet potatoes, processed fruit, and processed vegetables; and (4) *miscellaneous foods*, including fats and oils excluding butter, eggs, cereals and bakery goods, sugar and sweeteners, and coffee, tea and cocoa.