Testing the diversity of consumer preferences in Asia, Europe and North America

Oleksandr Movshuk
Visiting Researcher, ICSEAD

Working Paper Series Vol.99-18
September 1999

The International Centre for the Study of East Asian Development, Kitakyushu

Testing the diversity of consumer preferences in Asia, Europe and North America[*].

by Oleksandr Movshuk

International Centre for the Study of East-Asian Development,

Kitakyushu, Japan

September 22, 1999

We introduce a new approach to testing the diversity of unobservable consumer preferences in cross-section data. The approach is applied to three international cross-sections with Asian, European and North American countries for a large number of goods and services at three levels of aggregation. At the most aggregated level of consumption, the study found mostly insignificant differences across countries in consumer preferences. On the other hand, at the least disaggregated level we identified a number of commodities with highly unusual national consumption patterns. However, the latter commodities represented a small portion of analyzed commodities, indicating that consumer preferences are broadly comparable across countries.

JEL classification numbers: C15, C21, D12.

Keywords: consumer preferences, residual diagnostics, Monte Carlo testing.

[*] I would like to thank Yonas Biru from the World Bank for kindly providing data from the International Comparison Project. Several crucial suggestions from Kanemi Ban, Charles Yuji Horioka, Shinichi Ichimura, William James and Eric Ramstetter are also greatly appreciated.

Introduction.

In this paper we explore the international diversity of consumer preferences for a large number of goods and services, using data from three benchmark studies of the International Comparison Project (ICP). Our approach focuses on the predictive ability of the Almost Ideal Demand System (Deaton, Muellbauer, 1980a) after the ICP data are divided into the estimation and prediction subsets. Then we verify whether the consumption pattern in the prediction subset corresponds to the consumption pattern in the estimation subset.

The null hypothesis is that international differences in consumption for a specific good and service can be explained by conventional economic factors (such as income and price effects in the demand analysis), with insignificant residual deviations from the theoretical (e.g., predicted) consumption pattern. To increase the test power, the test algorithm endogenously allocates to the estimation subset only those countries that constitute the most regular consumption pattern. Then we evaluate the diversity of consumer preferences by how far the remaining countries in the prediction subset deviate from the regular part of data, using studentized prediction residuals.

The test statistic does not follow any of known standard distributions, but its significance can be evaluated by p-value, obtained by the Barnard (1963) Monte Carlo procedure, and previously applied by Bewley and Theil (1987)[1]. If the test's p-value is sufficiently close to one, then the observed gap between the regular and outlying subsets of data can be attributed to random factors, indicating no room for systematic differences in unobservable preferences. Conversely, if the calculated p-value is close to zero, then the null hypothesis that there are no residual national peculiarities in consumption can be rejected with high confidence.

The suggested approach differs from the vast majority of past studies of international consumer preferences. The latter repeatedly focused on the *estimation* aspect of the linear regression model by analyzing the international diversity of estimated regression *parameters*, as, for instance, in Lluch, Powell and Williams (1977), Kravis, Heston and Summers (1982). In contrast, our approach is based on the *predictive* ability of the regular part of data and the magnitude of studentized prediction *residuals*. Intuitively, if the whole sample of countries follows the same data-generating process with no substantial outliers, then it should be possible to predict adequately the consumption pattern for any composition of countries in the prediction subset.

Econometrically, our emphasis on the prediction aspect is closely related to the Chow (1960) test of predictive failure. However, there is one crucial difference. While the Chow test assumes that the partition pattern into the estimation and prediction subsets is known *a priori*, in this study the partition pattern is data-dependent.

The plan of the paper is as follows. Section 1 reviews related studies that examined the hypothesis of similar consumer preferences across countries. Section 2 describes our testing procedure to identify subsets of unusual countries with respect to predicted consumption patterns. Data sources and regression specifications are discussed in section 3, while section 4 contains major results. Section 5 concludes the paper.

Section 1. Related approaches to evaluate the diversity of consumer preferences across countries.

Despite the large number of studies that dealt with international differences in tastes, many of them have had rather tenuous links with the theory of consumer behavior. Besides, even if theoretically plausible regression specifications were used, quite often there was a little concern about misspecifications in the estimated regression equations. We will consider these two groups of studies in turn.

---

[1] In econometrics, the Monte Carlo test is better known as 'parametric bootstrap' (e.g., Horowitz, 1997).

The seminal paper by Houthakker (1965) is perhaps the most widely known study with no firm basis in the theory of consumer behavior. The study used the double-logarithmic specification

$$\log q_i = \beta_0 + \beta_1 \log p_i + \beta_2 \log \mu + \beta_3 t \qquad (1)$$

where q_i is real per capita expenditures on good *i*, p_i is the price index of *i*, normalized by a total price deflator, μ is the total expenditures in real terms, and *t* is a time trend. Though (1) does not have the additivity property of conventional utility functions, Houthakker still claimed that the specification 'remains without serious rivals in respect of goodness of fit, ease of estimation and immediacy of interpretation" (p. 278), and estimated the specification for 5 major consumption categories in 13 OECD countries.

Most subsequent studies of international demand functions avoided such an *ad hoc* approach. The most popular functional form has been the linear expenditure system[2], since the system satisfies all general restrictions of the demand theory. Besides, it does not require a large number of parameters in derived regression specifications[3]. Up to present, Lluch, Powell and Williams (1977) appears to be the most comprehensive application of the linear expenditure system to international data. The study covered 8 major expenditure categories in 17 developed and developed countries. Unfortunately, most data in the study followed the same upward trend, and no attempt was made to avoid spurious regressions that are typical in such trending data. For instance, out of 134 reported regressions, 104 ones had R^2 statistic larger than 0.95, while the Durbin-Watson statistic was less than unity in 60 regressions. According to the rule of thumb due to Granger and Newbold (1974), such combinations of test statistics indicate a large number of spurious regressions. Although the presence of spurious regressions greatly increases critical values from the *t* and *F* distributions, the

[2]See, for example, in Goldberger and Gamaletsos (1970), Parks and Barten (1973), Lluch and Powell (1975), Lluch, Powell and Williams (1977).

[3] Nevertheless, the double-log specification continues to surface occasionally in more recent studies of international demand function, including studies that used data from various benchmark studies of the International Comparison Project (Kravis *et al*, 1975, p. 279; Kravis *et al*, 1982, p. 357).

reviewed study still applied conventional critical values. This must have greatly inflated the nominal level of the hypothesis testing well above the declared 5% level, producing, in particular, too many rejections of the null hypothesis of common tastes.

The pitfalls of spurious regressions with time series data can be avoided by estimating demand functions with cross-section data. The most extensive study has been done by Kravis, Heston and Summers (1982). The study used ICP data for 1975 to address the question whether consumers in 10 Asian and 15 European countries had the same demand functions. The study applied the Chow ANOVA test of structural stability to 21 basic categories of food consumption, 25 summary consumption categories, 7 more aggregated 'grand' categories as well as 4 major consumption expenditures. The study concluded that approximately one-forth of the most disaggregated categories failed the structural test at 5% significance level (p.365). On the other hand, at more aggregated level the hull hypothesis of the same demand functions in Asia and Europe was less frequently rejected, with, finally, no rejections at all for the highest level of aggregation.

Unfortunately, the study did not report which consumption categories violated the hull hypothesis of common tastes. Besides, the application of the Chow test was valid under a number of crucial assumptions, which are unlikely to hold with the small samples of 10 and 15 countries. First, the application of the Chow test implies that the disturbance term in the Asian and European subsets had the same variance, whereas the authors themselves observed that the variance in Asian equations was "definitely larger" (p. 365). Second, the distribution of the Chow test statistic is derived on the assumption that regression disturbances are normally distributed in both subsets. Clearly, the standard reference to the central-limit theorem is quite tenuous with samples of 10 and 15 countries.

Given these pitfalls of parametric demand studies, it is not surprising that there was a growing interest in using non-parametric approaches, such as the revealed preference approach. The approach draws budget lines for two consumers (such as 'representative consumers' from different countries) and compares their actual choices at alternative sets of prices. Apart from several weak axioms of consumer behavior, no specific demand function has to be specified. One useful corollary of the revealed-preference analysis is its ability to check whether the two considered consumers share the same unspecified utility function, thus making possible testing the null hypothesis of common international tastes.

Unfortunately, the non-parametric approach also has a few limitations. For example, the approach can be applied only to consumption *bundles* instead of the more informative cases of specific commodities. Besides, if budget lines for consumers do not intersect (such as in comparisons between rich and poor countries), the revealed preference approach has no power to detect differences in tastes, no matter how large the differences may be.

The revealed-preference approach was first applied to international consumption data by Kravis, Heston and Summers (1982, p. 354-357). Using ICP benchmark data for 1975, the study compared consumer preferences in 34 developed and developing countries, analyzing total consumption bundles with 108 categories of goods and services. Despite the conspicuous diversity of analyzed countries from various continents, the study failed to detect even a single country-pair that violated the hypothesis of common preferences.

Dowrick and Quiggin (1994) applied the revealed-preference approach to 1980 data from the ICP that covered 60 countries. When consumption bundles contained 10 broad categories of expenditures, the study detected only two country-pairs (Finland-Austria and Nigeria-Zimbabwe) when the hypothesis of common preferences did not hold, though as many as 1700 country-pairs were considered. However, both these violations were insignificant at 5% level. Moreover, when Dowrick and Quiggin constructed consumption bundles with 38 categories of expenditures, the hypothesis of common tastes was supported by every country-pair.

To some extend, the overwhelming rejection of common tastes in the study can be attributed to its rather unusual composition of national consumption bundles. Dowrick and

Quiggin considered not only consumption goods, but also other categories of expenditures, including even investments in non-residential buildings, producer durables and inventory change as well as government expenditures. Such expenditures have little to do with private consumption despite the authors' assertion that investments are 'claims on future rather than present consumption' (p. 336). Besides, both applications of the revealed-preference approach considered consumption bundles with too many commodities, so that even if some international differences in tastes on *specific* commodities existed, they must have 'melted away' in the aggregated bundles.

In sum, despite the considerable effort to compare unobservable consumer preferences across countries, there has been no study that identified *specific* countries where consumer preferences on *specific* goods and services were unusual. The goal was most closely approached by Kravis, Heston and Summers (1982) when they applied the Chow test to Asian and European countries. However, conclusions of the study were limited to groups of countries, separated *a priori*. In the following section, we will introduce an approach that endogenously separates countries with typical and unusual consumption patterns for specific goods and services.

Section 2. Description of testing procedure.

Consider the conventional linear regression model $Y = X\beta + u$, where Y is ($n \times 1$) vector of observations on a dependent variable, X is ($n \times k$) matrix of n observations on k independent variables (assumed to be fixed and of full rank), β is ($k \times 1$) vector of unknown regression coefficients, and u is ($n \times 1$) vector of unobservable regression disturbances, with $u \sim N(0, \sigma^2)$.

In the context of international demand analysis, the vector Y typically refers to expenditures per capita on some good or service, while X contains data on k economic factors of consumption (such as income and price effects). All k variables are observable. In contrast, individual tastes – the focus of this study – are not observable, and therefore they must be relegated to the disturbance term *u*. Specifically, if consumer tastes in some countries are distinctive compared with the rest of analyzed countries, then the elements of vector *u* for the former countries become large in magnitude, unless some other omitted factors of consumer demand (such as climate or religion) offset the deviations in tastes.

Since the vector *u* is not observable, it is only possible to use some estimates of *u*. A natural substitute for *u* is ($n\times 1$) vector of OLS residuals $\hat{u}=Y-X\hat{\beta}$, where $\hat{\beta}=\left(X'X\right)^{-1}X'Y$. Since $\hat{u}=\left(I-V\right)u$, where $V=X(X'X)X'$, the vector of OLS residuals is a linear transformation of regression disturbances. In scalar form, the relationship simplifies to $\hat{u}_i=u_i-\sum_{j=1}^{n}v_{ij}u_j$.

Though $\hat{u}$ converges to *u* asymptotically (Theil, 1971, p. 378-379), in finite samples their correspondence may be quite poor. First of all, if some v_{ij} are large, then the second term of $\hat{u}_i=u_i-\sum_{j=1}^{n}v_{ij}u_j$ may dominate the first term, thus making $\hat{u}$ an inferior image of *u*. Second, if $\text{var}(u_i)=\sigma^2$, then $\text{var}(\hat{u}_i)=\sigma^2(1-v_i)$, with v_i denoting the i^{th} diagonal element of matrix *V*. Thus, observations with unusually large v_i (also called 'high leverage points' by Belsley, Kuh, and Welsch (1980)) tend to have estimated OLS residuals with small variance. In consequence, observations with unusually large *u*, but 'high leverage', may not have distinct OLS residuals, thus greatly complicating the identification outlying observations with high leverage by OLS residuals.

To avoid unequal variance of OLS residuals, the residuals can be standardized by $\sqrt{1-v_i}$. If regression disturbances u_i $(i=1,...,n)$ have zero mean and constant variance σ^2, so do the standardized OLS residuals $\hat{u}_i\big/\sqrt{1-v_i}$. Even more informative are studentized residuals $t_i=\hat{u}_i\big/\hat{s}_{(i)}\sqrt{1-v_i}$, where $s_{(i)}$ is the OLS estimate of σ^2 with the i^{th} observation omitted. It can be shown that under the null hypothesis $u \sim N(0,\sigma^2)$, studentized residuals t_i follow the Student's t-distribution with $n-k-1$ degrees of freedom (Cook, Weisberg, 1982), thus making possible statistical inference about the magnitude of unobservable regression disturbances.

Unfortunately, the t_i statistic can be informative about unobservable u_i only if there is only one deviant observation with large u_i. Conversely, t_i may have low power if there are several outliers. This is because $t_i = \hat{u}_i \big/ \hat{s}_{(i)}\sqrt{1-v_i}$ is algebraically equivalent to the studentized prediction residual

$$\widetilde{t}_i = (y_i - x_i\beta_{(i)}) \Big/ \hat{s}_{(i)}\sqrt{1 + x_i^{'}(X_{(i)}^{'}X_{(i)})^{-1}x_i} \tag{2}$$

where $\hat{\beta}_{(i)}$ as the OLS estimator of β with the i^{th} observation omitted, so that $\hat{\beta}_{(i)} = (X_{(i)}^{'}X_{(i)})^{-1}X_{(i)}^{'}y_{(i)}$, where both $X_{(i)}$ and $y_{(i)}$ have i^{th} row omitted (Hadi, Son, 1990). If there are several regression outliers, part of them may continue to affect the estimated $\beta_{(i)}$, since only a single i^{th} observation has been omitted. Moreover, if these multiple outliers resemble each other, then the prediction error $y_i - x_i\beta_{(i)}$ in the nominator of (2) may not become large, producing so-called 'masking effect' when multiple outliers make each other difficult to detect. In the next section, we will demonstrate the masking effect of $\widetilde{t}_i$ with a numerical example.

The masking effect can be avoided if, instead of a single observation, several outlying observations are excluded in the estimation of $\hat{\beta}$ (in other words, when the full sample is divided into the estimation and prediction subsets, with the latter containing multiple observations). To achieve high power, the estimation subset should also consist of regular observations, while potential outliers must be confined only to the prediction subset. In the suggested testing procedure, the separation into the estimation and prediction subsets is achieved as follows:

Step 1. Apply a robust regression to all data, using, for example, the least trimmed squared (LTS) estimator of Rousseeuw (1984).

Step 2. Sort observations by absolute values of their residuals from the LTS fit. Denote the sequence of absolute LTS residuals from the smallest to the largest by $e_1 \leq e_2 \leq e_3 \leq ... \leq e_{n-2} \leq e_{n-1} \leq e_n$.

Step 3. Partition data into the original estimation subset B_1 with the smallest $k+1$ absolute LTS residuals and the prediction subset P_1 with the rest of $n-(k+1)$ observations.

Step 4. Estimate OLS regression with observations in B_1, and then calculate studentized prediction residuals $\tilde{t}_i$ for all observations belonging to P_1.

Step 5. Find an observation with the smallest absolute $\tilde{t}_i, i \in P_1$, and record it as $\tilde{t}_{k+2}^{\min}$.

Step 6. Expand the estimation subset by the least outlying observation in P_1 (the one with $\tilde{t}_{k+2}^{\min}$), and repeat from step 1 to step 6.

Step 7. Continue until the prediction subset contains the last remaining observation. Record the last test statistic as $\tilde{t}_n^{\min}$ and stop.

After repeating the test algorithm $n-(k+1)$ times, one obtains the sequence of $\tilde{t}_{k+2}^{\min}$, $\tilde{t}_{k+3}^{\min}$,…, $\tilde{t}_n^{\min}$ test statistics for progressively decreasing subsets of potential outliers. Since the test assumes no knowledge about the number and sign of outlying observations, one possible criteria to identify the most likely subset of regression outliers by the largest $\left|\tilde{t}_j^{\min}\right|, j = k+2,...,n$.

As noted above, under the null hypothesis $u \sim N(0,\sigma^2)$, the sequence of $\left|\tilde{t}_j^{\min}\right|, j = k+2,...,n$ follows the Student's t-distribution with $j-k-1$ degrees of freedom. However, it is important to note that the null distribution of successive $\tilde{t}_{k+2}^{\min}$, $\tilde{t}_{k+3}^{\min}$,…, $\tilde{t}_n^{\min}$ test statistics depends on the varying degrees of freedom (ranging from 1 d.f. for $\tilde{t}_{k+2}^{\min}$ to n-k-1 d.f for $\tilde{t}_n^{\min}$), so that the sequence of $\tilde{t}_{k+2}^{\min}$, $\tilde{t}_{k+3}^{\min}$,…, $\tilde{t}_n^{\min}$ test statistics is not equivalent. To make the sequence comparable, one may transform the sequence of $\tilde{t}_{k+2}^{\min}$, $\tilde{t}_{k+3}^{\min}$,…, $\tilde{t}_n^{\min}$ into standard normal deviates *N(0,1)* by available normalizing transformations, as mentioned by Hawkins (1991, p. 223). Hawkins also suggested the following normalizing formula due to Wallace (1959, p. 1125):

$$z_j^* = \frac{8v+1}{8v+3}\sqrt{v\log_e\left(1+\frac{\tilde{t}_j^2}{v}\right)} \tag{3}$$

where $v = j-k-1$ denotes the corresponding degrees of freedom, with v ranging from 1 to $n-k-1$.

After transforming $\tilde{t}_{k+2}^{\min}$, $\tilde{t}_{k+3}^{\min}$,…, $\tilde{t}_n^{\min}$ into the sequence of directly comparable test statistics z_{k+2}^*, z_{k+3}^*,…, z_n^*, the most outlying subset is identified by the final test statistic $Z = \max\left|z_j^*\right|$, $j = k+2,...,n$.

In addition to (3), several other normalizing transformations can be used as well. Wallace (ibid.) suggested an alternative transformation with better accuracy even for small v, but at the cost of more complicated formula:

$$z_j^{**} = \left[1-\frac{2}{8v+3}\sqrt{1-e^{-s^2}}\right]\left\{v\log_e\left(1+\frac{\tilde{t}_j^2}{v}\right)\right\}\sqrt{v\log_e\left(1+\frac{\tilde{t}_j^2}{v}\right)} \tag{4}$$

where $s = \dfrac{0.184(8v+3)}{v}\left\{\log_e\left(1+\dfrac{\tilde{t}_j^2}{v}\right)\right\}^{-1/2}$

More recently, Bailey (1980, p. 305) suggested another transformation with superior accuracy for large v:

$$z_j^{***} = \frac{8v+1}{8v+9}\sqrt{\frac{v+19}{12}\log_e\left\{1+\frac{\tilde{t}_j^2}{v+1/12}\right\}} \tag{5}$$

Once the test statistic $Z = \max\left|z_j^*\right|$ is obtained by using either (3), (4), or (5), the statistical significance of Z may be evaluated. Due to the preliminary ordering of data by LTS residuals the sequence of $\tilde{t}_{k+2}^{\min}$, $\tilde{t}_{k+3}^{\min}$,…, $\tilde{t}_n^{\min}$ is not independent, thus making the analytical distribution of Z intractable. However, the distribution of Z can be approximated by the Barnard (1963) Monte Carlo procedure. The procedure (also known as parametric bootstrap) contains the following three steps.

First, one needs to generate many subsets of artificial data according to the null hypothesis of no regression outliers, or $H_0 : u_i \sim \mathrm{N}(0,\sigma^2)$, with σ^2 not generally known. However, since the sequence of $\tilde{t}_{k+2}^{\min}$, $\tilde{t}_{k+3}^{\min}$,…, $\tilde{t}_{n}^{\min}$ is the ratio of estimated mean and standard deviation, the distribution of *Z* does not depend on unknown parameters β and σ^2 (in other words, the statistic *Z* belongs to pivotal statistics). Therefore, without loss of generality, one can obtain the distribution of *Z*, using any arbitrary values for β and σ^2. For example, fixing all these parameters at unity, the artificial data can be generated by $y_b = 1 + x_1 + ... + x_k + N(0,1)$.

Second, the test statistic Z is calculated from ($n \times 1$) vector y_b and ($n \times k$) actual matrix X, and the calculated bootstrap test statistic Z_b is stored. The procedure is repeated *B* times. Upon completing, all test statistics Z_b are sorted in absolute values.

Third, we get p-value for the test statistic by counting how many times the actual test statistic *Z* exceeds Z_b from the simulated data under the null conditions. Formally, the approximate p-value $\hat{p}(Z) = \dfrac{1}{B+1}\sum_{s=1}^{B} I\left(Z_B > Z\right)$, where $I(\cdot)$ is the indicator function. Under the mild regularity conditions, as $B \to \infty$, the estimated p-value will tend to the true p-value (Horowitz, 1997). Moreover, it can be shown that for pivotal and two-sided test statistics like *Z*, the bootstrap approximation makes error of order $O(n^{-2})$. In contrast, the traditional asymptotic approximations make errors of size $O(n^{-1})$ (*ibid.*), thus supporting the advantage of using the Bernard procedure to approximate the p-value of non-standard (but pivotal) test statistic *Z*.

In the next section, we will illustrate the sequence of test algorithm, as well as the accuracy of previously mentioned normalizing transformations for $\tilde{t}_{k+2}^{\min}$, $\tilde{t}_{k+3}^{\min}$,…, $\tilde{t}_{n}^{\min}$.

Section 3. Data and regression specification.

a) Data.

We used data from three ICP benchmark estimates for 1975, 1980 and 1985. The ICP data were obtained by request from the World Bank. These data included per capita expenditures on 110 categories of consumption goods and services, and were expressed in domestic currencies and at purchasing power parities.

The ICP data are collected in international cross-sections with no linkages in time, precluding the modeling of consumer preferences for durables[4]. We also omitted several consumption categories that contained a large number of missing data. Finally, we often followed the comment by Kravis *et al* (1975, p. 49) that so called 'residual categories' of ICP data (such as 'milk products, not else specified') may contain a high share of internationally incomparable data[5], and omitted these residual categories from the final database of analyzed goods and services.

Eventually, our international database covered Asian, European and North American countries and included 24, 26 and 25 countries for 1975, 1980 and 1985 cross-sections (table 1), and these cross-sectional data contained 40, 53, and 54 categories of goods and services.

Since the original categories of goods and services were highly disaggregated, we also grouped the data according to the concept of multistage budgeting (or utility tree). Assuming the weak separability in preferences, the concept allows the determination of consumer preferences only by a subset of a few related commodities, ignoring the impact from less relevant commodities (Deaton, Muellbauer, 1980b). We assumed the following structure of the utility tree:

Aggregation level 1. Consumers allocate expenditures to the most aggregated categories of data, such as food, clothing, fuel and power, medical services, purchased transport and the like (see table 2-1).

[4]Unlike non-durables and services, the consumption of durables is extended over a long time, and is affected by the past and future conditions in the economy.
[5]Since it is highly unlikely that national statistical offices interpret such categories in similar ways.

Aggregation level 2. Subsequently, consumers subdivide consumption expenditures at the most aggregate level into lower levels of aggregation. For example, food is further subdivided into 10 less aggregated expenditures, such as bread/cereals, meat, fish, milk, etc. (see table 2-2).

Aggregation level 3. Finally, expenditures on bread and cereals are further subdivided into expenditures at the lowest desegregation level, consisting of rice, flour, bread, bakery products, etc. (see table 2-3). Meat, fish, milk and other commodities from the aggregation level 2 are subdivided similarly.

In most cases, the pattern of the multistage budgeting in tables 2-1, 2-2, and 2-3 follows the standard ICP classification scheme, given in Kravis *et al* (1982, pp. 60-66), which, in turn, closely corresponds to the standard classification of national accounts.

b) Regression specification.

Our regression specification was based on the Almost Ideal Demand System of Deaton and Muellbauer (1980a), since it provides a first-order approximation to any set of demand systems. The demand system included the income term and price terms (at the relevant level of multistage budgeting):

$$w_i = \beta_{i,0} + \log_e\left(\mu/\overline{P}\right) + \sum_j \beta_{i,j} \log_e\left(P_{i,j}/\overline{P}\right) \tag{6}$$

where w_i denotes i^{th} commodity's share in total nominal expenditures (i.e., expenditures in national currency) with respect to a relevant sub-system, $P_{i,j}$ is purchasing power parity for i^{th} commodity, $\overline{P}$ is the aggregated purchasing power parity for the whole sub-system, and M is total consumption expenditures (in national currencies) on all goods, included in the relevant sub-system.

To illustrate the suggested testing procedure, we will use actual ICP data from 1980 cross-section and consider demand for coffee. In addition to coffee (indexed 1.9.1 in table 2-3), other closely related commodities include tea (index 1.9.2) and cocoa (index 1.9.3). All required ICP data are given in table 3. As shown in the second column of table 3, there are quite substantial international discrepancies in the expenditure shares, spend on coffee. For example, in seven countries (Belgium, Denmark, Germany, Hungary, Italy, Norway, and Spain) coffee accounts for more than 90% of all expenditures in the summary category 'coffee, tea, cocoa', with the median share for the whole sample as high as 86%.

In contrast, in Hong Kong, Ireland, Japan, Korea, Sri Lanka, and the United Kingdom the share of coffee is less than 60%. Given the wide variation in the dependent variable, it is interesting if the variation can be accounted by the conventional factors such as income and price effects.

Table 4 describes the sequence of steps of the proposed testing procedure. Before proceeding to the test, it is instructive to look at studentized residual $t_i = \hat{u}_i \big/ \hat{s}_{(i)}\sqrt{1-v_i}$, which is a conventional 'leave-one-observation-out' outlier diagnostic (t_i is given in second column of table 4). With $n = 26$ and $k = 5$, the 5% critical value for the largest studentized residuals equals $t_{n-k-1}^{\alpha/n} = t_{21}^{0.05/26} = 3.54$. As table 4 shows, it is Hong Kong that has the largest studentized residual 2.87, but the statistic falls short of its 5% critical level. In other words, the conventional outlier diagnostic detects no substantial regression outliers.

Now we turn to the description of the suggested testing algorithm to identify observations with large *u*. First, observations are sorted by the absolute residuals from the least trimmed squared (LTS) fit. These residuals are given in the third column of table 3. The subset of '*k+1* most-regular observations (with smallest magnitude of LTS residuals) includes Germany, Greece, Finland, France, Portugal and Israel.

Also note that the United States, Sri Lanka, Indonesia, Japan, the United Britain, Ireland and Hong Kong has quite large LTS residuals compared with the rest of countries. However, the distribution of LTS residuals under the null hypothesis of normally distributed *u* is not analytically tractable in finite samples, so that it is not possible to conclude 'how large are large' LTS residuals. In contrast, the distribution of OLS residuals is understood much better, and this advantage is used during the second step of the proposed test.

After applying the OLS to the subset of '*k+1* most regular observations', we calculate studentized prediction residuals $\tilde{t}_i, i = k+2, k+3, ..., n-1, n$ for the rest of sample, as given by (2). Then we find an observation with the minimum studentized prediction residual. As shown in forth column of table 4, it is Canada with $\left|\tilde{t}_{k+2}^{\min}\right| = 8.097$. We record the first test statistic. At the second step, we augment the original estimation subset B_1 with Canada to obtain subset B_2. Then we re-estimate the demand system with observations in B_2 to find the next least outlying observation in the prediction subset P_2 (Norway, with studentized prediction residual 2.198).

After repeating steps 1 to 6, the estimation subset is augmented by Austria, Belgium, Poland and so on, until we reach the last loop, when the final member of the prediction subset is Hong Kong.

The largest absolute test statistic in the sequence of $\tilde{t}_{k+2}^{\min}$, $\tilde{t}_{k+3}^{\min}$,..., $\tilde{t}_{n}^{\min}$ is 8.097 from the first loop. Under the null, the test statistic is t-distributed with 1 d.f., so that 8.097 has 0.078 two-tail probability, as shown in column 5 of table 4. The corresponding exact inverse CDF from the standard normal distribution $N(0,1)$ is 1.761 (column 6).

The next test statistic (2.198) corresponds to the t-distribution with 2 d.f., from which the exact two-tail probability equals 0.159 and the inverse CDF of $N(0,1)$ is 1.408. After normalizing the next test statistics in the same way, we find that the largest normalized test statistic equals to 4.470. The statistic corresponds to the United States and six other countries that were more even more outlying: Sri Lanka, Indonesia, Japan, UK, Ireland, and Hong Kong. Consequently, these seven countries constitute the most outlying subset, with the test statistic Z equal 4.470.

The following columns in table 4 demonstrate the precision of the above-mentioned normalizing transformations (3), (4), and (5). With 1 d.f., z_i^{**} is the most accurate, with the absolute error of only 0.006, while z_i^{*} and z_i^{***} deviate from the exact z_i by 0.085 and 0.034.

The table also demonstrates that if the number of degrees of freedom gets larger, the accuracy of all considered normalizing formulas greatly improves, with most absolute errors about one unit in three decimal places, which should be sufficient for most practical purposes.

After finding the subset of most-outlying countries with $Z = \max|t_i| = 4.47$ (where $i = k+2,...,n$), we evaluate the statistical significance of the calculated test statistic. By applying the Barnard's Monte Carlo test $B = 499$ times with normally distributed artificial *u*, we ordered Z_B by absolute values, and the 5% significance critical value was given by the average of 475th and 476th quintiles of Z_B. The critical value was 3.29, so that the null hypothesis can be rejected at 5% confidence level. In fact, *no* Z_B exceeded the actual $Z = 4.47$, so that the approximate p-value for this Z is essentially zero, indicating that the identified seven outlying countries have very unusual preferences for coffee compared with the pattern in the other countries.

This result is also recorded in table 6 that contains categories with the most outlying countries for 1980 cross-section. To complete the example, consider the column labeled 'coffee'. There are 7 normalized test statistic (as well as their sign), so that the least outlying case of the US is filled by -4.47, followed by Sri Lanka with –4.84 and finally the most unusual case of Hong Kong with –7.40. At the bottom, the table contains the actual test statistic (4.47), its 5% critical value from the Barnard procedure with $B = 499$ (3.29), and, finally, the p-value for the statistic (0.000).

Section 4. Major results.

Here we consider consumption categories that had the largest international diversity in 1975, 1980 and 1985 cross-sections (tables 5, 6, and 7, correspondingly). To save the space, we provide results only for consumption categories where the test statistic *Z* had p-values less or equal to 0.200. Results for all consumption categories are available on request.

a) 1975 cross-section.

Consider the first consumption category in table 5 – ‘recreation’. There are 3 positive outliers (Netherlands, Romania, and Thailand) and 3 negative ones (Iran, Korea, and Luxembourg). The test statistic equals 3.48, and its p-value under the null hypothesis is 0.156. Thus, the identified outlying countries are nearly significant at conventional significance levels.

A number of other consumption categories demonstrated more significant results with smaller p-values. Examples include ‘spices’ (indexed 1.8 in table 2-2) with positively outlying Japan, and also ‘rice’ (indexed 1.1.1), where *all* Asian countries in the sample are positively outlying. These categories have test statistics 4.69 and 4.46, with corresponding 0.000 and 0.002 p-values.

Other highly significant cases include ‘fresh fruits’ (indexed 1.6.1), with negatively outlying Denmark, France, Ireland, Netherlands, UK and USA, and ‘condiments’ (indexed 1.10.4), with positively outlying Korea, Philippines, Spain, Sri Lanka, Thailand, and Yugoslavia.

b) 1980 cross-section.

As shown in table 6, in the case of ‘rice’ we again find a familiar pattern when all Asian countries (plus Portugal) have unusual pattern consumption, with p-value for the category at zero. ‘Bread’ has also a large number of significant outliers, and most of them are negative. There are 3 positive outliers in the consumption of ‘noodles’ (Canada, Italy and Japan).

There is an interesting pattern in the consumption of ‘potatoes’, with 4 negatively outlying Asian countries (Hong Kong, Indonesia, Korea, and Sri Lanka). Apparently, this is because ‘potatoes’ are coupled in the ICP classification with ‘manioc, arrowroot and other tubers’ (indexed 1.8.2 in table 2-3) that have relatively large consumption in Asia.

There is also an intriguing substitution relationship between ‘coffee’ (already discussed in section 3) and ‘tea’. In contrast to seven negative outliers in the case of coffee, there are two positive outliers for ‘tea’ (Hong Kong and Ireland) that essentially offset the negative preferences for ‘coffee’ in these countries.

Finally, Hong Kong, Japan and Sri Lanka showed a quite large positive preference for ‘condiments’ in the ‘sugar, sweets, and condiments’ branch of the utility tree. This is apparently due to relatively small demand for ‘sugar’ in Asian countries, revealed among results for the following 1985 cross-section.

Though the majority of identified outliers in consumption appears to be feasible, the negative residual deviation of Japan for fresh fish is obviously odd. We attribute the result to substantial changes in classifying fish during the ICP studies. While Japanese national account data in 1975-1985 constantly allocated roughly 75 per cent to ‘fresh fish’ (out of ‘total fish’), the corresponding shares in the ICP data are 0.777, 0.274, 0.432. In particular, the sharp drop in 1980 appears to be due to the ICP extension of ‘canned and preserved fish’ (indexed 1.4.2 in table 2-3), which is the only complement category to ‘fresh fish’.

The latter case is an important reminder that the disturbance term u has a comprehensive meaning, and may refer not only the unobservable differences in preferences, but also to all other factors of consumption (climate, religion, etc.) which are unaccounted for in the estimated demand system.

Here we find an interesting analogy with a conceptually similar measurement of unobservable ‘productivity’ by unexplained *residuals* after estimating a production function. Arguably, both these ‘residual measures’ are essentially indicators of ‘our ignorance’ rather than the alleged productivity or preferences. Nevertheless, since both preferences and productivity are not directly observable, there may be little hope to find a more precise and direct approach to evaluate the magnitude of these interesting economic phenomena.

c) 1985 cross-section.

As shown in table 7, there are a number of diversity patterns in consumption that are familiar from previous cross-sections. First, Asian countries and Portugal again show a large positive deviation in ‘rice’, with p-value essentially zero. Italy again is a large positive outlier for ‘noodles’. There are also similarities for ‘coffee’ and ‘tea’. Specifically, Australia, Ireland, New Zealand, Portugal, Sri Lanka, and UK (essentially – former members of the British Empire) have negative deviation for ‘coffee’ (with 0.096 p-value), while Hong Kong, Japan, Korea and Sri Lanka are positively outlying in the consumption of ‘tea’ (with 0.010 p-value). Finally, all negative deviations in the consumption of ‘sugar’ are Asian countries. It is noteworthy that among these countries, Hong Kong, Japan, and Sri Lanka already had positive deviation in the related ‘condiments’ category in 1980 cross-section, so that there may be a substitution pattern between ‘sugar’ and ‘condiments’ in these three countries.

Section 5. Conclusions.

This paper introduced a testing procedure to identify subsets of observations with unusually large regression disturbances. The procedure was applied to international data on consumption from the International Comparison Project in 1975, 1980 and 1985.

Despite the lack of exact correspondence between these cross-sections (due to changes in country coverage and available consumption categories), there were a few consumption categories that showed consistent patterns of unexplained consumption that had significant test statistics Z. Most notably, such cases included rice, bread, noodles, coffee, tea, sugar and condiments. We also often found distinct disparities between Asian and other countries in the consumption of rice, potatoes, sugar and condiments.

However, the clear-cut ‘inter-continental’ pattern was often absent, as in the case of coffee, when substantial negative deviation was detected not only in Asian countries, but also in countries sharing the colonial heritage of the British Empire. Using the example of coffee, we also demonstrated the insufficiency of conventional measures of regression outliers (e.g., studentized residuals). The latter failed to identify even a single outlying country at 5% significance level, though the suggested test detected as many as 7 outlying countries with a highly significant test statistic.

In sum, though we identified a number of commodities with highly unusual national consumption patterns (especially at the most disaggregated level), these cases represented a small portion of analyzed commodities. Therefore, with the exceptions of cereals, tubers, non-alcoholic drinks and few other commodities, consumer preferences appear to be broadly comparable across countries.

References.

Bailey B. J. R. (1980) “Accurate normalizing transformation of Student’s *t* variate”. *Applied Statistics*, 29, 304-305.

Barnard G. A. (1963) “Comment.” *Journal of Royal Statistical Society, Series B,* 25, pp. 294.

Belsley, D. A., Kuh, E. and Welsch, R. E. (1980) *Regression diagnostics: identifying influential data and sources of collinearity.* New York: Wiley.

Bewley R. and H. Theil (1987). “Monte Carlo testing for heteroscedasticity”. In T. B. Fomby and G. F. Rhodes Jr. (eds.), *advances in econometrics* (c*omputation and simulation)*. Vol. 6, JAI Press: Greenwich.

Chow, G. (1960). ‘Tests of equality between sets of coefficients in two linear regressions’, *Econometric*, vol. 28, pp. 591-605.

Cook R. D. and S. Weisberg (1982). *Residuals and influence in regression.* Chapman and Hall: New York.

Deaton A. S. and J. Muellbauer (1980a). “An almost ideal demand system”, *American Economic Review*, 70, pp. 312-326.

Deaton, Angus and Muellbauer, John. (1980b). *Economics and consumer behavior*. Cambridge: Cambridge University Press.

Dowrick Steve, and John Quiggin (1994). “International comparison of living standards and tastes: a revealed-preference analysis”, *American Economic Review*, vol. 84, no.1, pp. 332-341.

Goldberger A. S. and T. Gamaletsos (1979). “A cross-country comparison of consumer expenditure patterns”, *European Economic Review*, 1, 357-400.

Granger C. W. J. and P. Newbold (1974). “Spurious regressions in econometrics,” *Journal of Econometrics,* 2, 111-120.

Hadi A. S., and M. S. Son (1990). “Some properties and relationships among several uncorrelated and homoscedastic residual vectors”. *Communications in Statistics: Theory* *and Methods*, 19, pp. 2625-2642.

Hawkins D.M (1991). “Diagnostics for use with regression recursive residuals“. *Technometrics*, vol. 33, pp. 221-234.

Horowitz, J.L. (1997). “Bootstrap methods in econometrics: theory and numerical performance“. In D.M. Kreps and K.F. Wallis, eds. *Advances in economics and econometrics: theory and application*s, Seventh World Congress, vol. 3, Ch. 7. Cambridge, U.K.: Cambridge University Press.

Houthakker H.S. (1965). “New evidence on demand elesticities“. *Econometrica*, 33, pp. 277-288.

Kravis Irving, Alan Heston, and Robert Summers (1982), *World product and income: international comparison of real gross product*. The Johns Hopkins University Press, Baltimore and London (for the World Bank).

Kravis Irving, Zoltan Kenessey, Alan Heston, and Robert Summers (1975), *A system of international comparison of gross product and purchasing power*. The Johns Hopkins University Press, Baltimore and London (for the World Bank).

Lluch C. and A. Powell (1975). “International comparisons of expenditure patterns”. *European Economic Review*, 5, 275-303.

Lluch C., A. A. Powell and R.A. Williams (1977). *Patterns in household demand and savings.* Cambridge, U.K.: Cambridge University Press.

Parks R. W. and A. Barten (1973). “A cross-country comparison of the effects of prices, income and population composition on consumption patterns”. *Economic Journal*, 83, 834-852.

Rousseeuw, P. J. (1984). “Least median of squares regression”. *Journal of American Statistical Association,* vol. 82, pp. 851-857.

Theil H. (1971). *Principles of econometrics*. New York: John Wiley.

Wallace D. L. (1959). “Bounds on normal approximations to Student’s and the Chi-Distributions”. *Annals of Mathematical Statistics*, vol. 30, pp. 1121-1130.

Table 1. Country coverage.

	1975	1980	1985
1	Austria	Austria	Australia
2	Belgium	Belgium	Austria
3	Denmark	Canada	Belgium
4	France	Denmark	Canada
5	Germany	Finland	Denmark
6	Hungary	France	Finland
7	Iran	Germany	France
8	Ireland	Greece	Germany
9	Italy	Hong Kong	Greece
10	Japan	Hungary	Hong Kong
11	Korea	Indonesia	Ireland
12	Luxembourg	Ireland	Italy
13	Malaysia	Israel	Japan
14	Netherlands	Italy	Korea
15	Philippines	Japan	Luxembourg
16	Poland	Korea	Netherlands
17	Romania	Luxembourg	New Zealand
18	Spain	Netherlands	Philippines
19	Sri Lanka	Norway	Portugal
20	Syria	Philippines	Spain
21	Thailand	Poland	Sri Lanka
22	Great Britain	Portugal	Sweden
23	United States	Spain	Thailand
24	Former Yugoslavia	Sri Lanka	Great Britain
25		Great Britain	United States
26		United States	

Table 2-1.
Major categories of consumption.

1. Food, beverages, and tobacco.
2. Clothing.
3. Footwear.
4. Gross rent, fuel and power.
5. Furniture, furnishings, household equipment.
6. Medical care and health expenses.
7. Transport and communication.
8. Recreation, entertainment.
9. Education, and cultural services.
10. Personal care

Table 2-2.
Food, beverages, and tobacco.

1.1 Bread and cereals
1.2 Meat
1.3 Fish
1.4 Milk
1.5 Oils and fats
1.6 Fruits, vegetables and tubers
1.7 Coffee, tea and cocoa
1.8 Sweets, spices
1.9 Beverages (alcoholic and non-alcoholic)
1.10 Tobacco

Table 2-3.

Bread and cereals

1.1.1 Rice, glazed or polished
1.1.2 Maize, meal and flour of wheat, barley, and other cereals.
1.1.3 Bread and rolls
1.1.4 Biscuits, cake, tarts, pies and other bakery products
1.1.5 Cereal preparations, preparations of flour, starch
1.1.6 Macaroni, spaghetti, noodles, vermicelli, and similar products.

Meat

1.2.1 Fresh beef and veal
1.2.2 Fresh lamb and mutton
1.2.3 Fresh pork
1.2.4 Fresh poultry
1.2.5 Dried, salted, smoked, canned meat, meat preparations (bacon, ham, sausages, etc).

Fish

1.3.1 Fresh or frozen fish and other seafood.
1.3.2 Canned and preserved fish and other seafood.

Milk, cheese and eggs

1.4.1 Fresh milk
1.4.2 Milk products (evaporated, condensed, dried milk, cream, yogurt).
1.4.3 Cheese
1.4.4 Eggs and egg products.

Oils and fats

1.5.1 Butter
1.5.2 Margarine, edible oils, peanut butter, mayonnaise, other edible oils, lard and other edible fat.

Fruits

1.6.1 Fresh fruits (orange, tangerine, lemon, lime, grapefruit, banana, mango, pineapple, apple, pear, cherry, grape, melon, plum, strawberry, and the like).
1.6.2 Dried, frozen, preserved fruits, juices, fruit peel, nuts, and parts of plants preserved by sugar.

Vegetables

1.7.1 Fresh vegetables (beans, cabbages, carrots, cucumbers, eggplants, garlic, ginger, onion, peas, spinach, lettuce, tomatoes, edible seeds, herbs, mushrooms and the like)
1.7.2 Dried, frozen, preserved vegetables, vegetable juices vegetable soups.

Tubers

1.8.1 Potatoes.
1.8.2 Other tubers (manioc, arrowroot, sweet potatoes, and other starchy roots).

Coffee, tea, cocoa

1.9.1 Coffee
1.9.2 Tea
1.9.3 Cocoa

Sugar, sweets and condiments

1.10.1 Sugar
1.10.2 Jam, preserves, marmalades, jellies, syrup, honey
1.10.3 Chocolate, sugar confectionery, ice cream
1.10.4 Salt, spices, vinegar, prepared baking powders, sauces.

Alcoholic beverages

1.11.1 Spirits
1.11.2 Wine and cider
1.11.3 Beer

Table 3. Data for illustrative example

	w_{coffee}	$\log_e\left(\mu/\overline{P}\right)$	$\log_e\left(P_{coffee}/\overline{P}\right)$	$\log_e\left(P_{tea}/\overline{P}\right)$	$\log_e\left(P_{cocoa}/\overline{P}\right)$
Austria	0.897	3.69	0.374	0.248	0.571
Belgium	0.952	3.78	0.403	-0.046	0.499
Canada	0.677	4.60	0.234	0.225	0.745
Denmark	0.906	4.57	0.381	0.122	0.572
Finland	0.934	3.82	0.370	1.019	1.048
France	0.738	3.81	0.296	0.210	0.589
Germany	0.917	3.83	0.398	-0.022	0.398
Greece	0.793	2.58	0.367	-0.043	0.366
Hong Kong	0.213	2.86	0.577	-0.277	-0.210
Hungary	0.946	3.30	0.420	-0.245	-0.011
Indonesia	0.616	2.67	0.808	-0.725	0.605
Ireland	0.162	3.50	0.976	-0.386	0.291
Israel	0.844	3.41	0.371	0.188	0.306
Italy	0.942	3.24	0.402	-0.161	0.401
Japan	0.387	3.36	0.364	-0.278	0.949
Korea	0.437	0.53	0.281	-0.204	0.768
Luxembourg	0.849	4.03	0.371	-0.080	0.480
Netherlands	0.873	4.43	0.404	-0.252	0.374
Norway	0.935	3.80	0.389	1.036	0.195
Philippines	0.668	2.04	0.212	1.886	0.974
Poland	0.767	2.41	0.346	-0.059	0.380
Portugal	0.853	2.15	0.421	-0.283	0.118
Spain	0.947	3.21	0.400	0.200	0.000
Sri Lanka	0.513	2.65	1.539	-0.850	1.200
Great Britain	0.477	3.84	0.709	-0.527	0.214
United States	0.861	3.58	0.764	0.045	0.508

Table 4. Test statistics for the illustrative example.

	t_i	e_i	$\widetilde{t}_i$	Tail prob. of $\widetilde{t}_i$	z_i (exact)	z_i^{*}	z_i^{**}	z_i^{***}	$\lvert z_i - z_i^{*}\rvert$	$\lvert z_i - z_i^{**}\rvert$	$\lvert z_i - z_i^{***}\rvert$
Germany	0.56	0.06									
Greece	0.41	0.08									
Finland	0.41	0.12									
France	0.45	0.12									
Portugal	1.04	0.12									
Israel	0.26	0.12									
Canada	1.22	0.47	8.097	0.078	1.761	1.676	1.755	1.727	0.085	0.006	0.034
Norway	0.30	3.60	2.198	0.159	1.408	1.402	1.409	1.410	0.006	0.001	0.002
Austria	0.43	0.43	2.835	0.066	1.839	1.830	1.840	1.837	0.009	0.001	0.002
Belgium	0.79	1.30	2.016	0.114	1.580	1.579	1.580	1.581	0.001	0.000	0.000
Poland	0.34	0.80	2.283	0.071	1.804	1.802	1.803	1.804	0.002	0.001	0.000
Hungary	0.99	0.33	2.187	0.071	1.803	1.802	1.802	1.803	0.001	0.001	0.000
Italy	0.97	2.11	2.177	0.066	1.839	1.838	1.838	1.839	0.001	0.001	0.000
Korea	0.72	0.77	-2.124	0.066	1.835	1.835	1.835	1.835	0.001	0.001	0.000
Spain	0.85	0.12	1.910	0.088	1.703	1.703	1.703	1.703	0.000	0.000	0.000
Luxembourg	0.17	0.40	-2.208	0.052	1.945	1.945	1.945	1.945	0.001	0.001	0.000
Denmark	0.20	0.66	-1.791	0.101	1.641	1.641	1.641	1.641	0.000	0.000	0.000
Philippines	1.16	0.12	2.613	0.023	2.279	2.278	2.278	2.279	0.001	0.001	0.000
Netherlands	0.22	1.81	-1.813	0.093	1.680	1.680	1.680	1.680	0.000	0.000	0.000
USA	0.86	21.18	-6.895	0.000	4.470	4.474	4.483	4.479	0.003	0.012	0.009
Sri Lanka	1.67	67.07	-1.718	0.106	1.614	1.614	1.614	1.614	0.000	0.000	0.000
Indonesia	0.36	26.18	-0.866	0.399	0.843	0.843	0.843	0.843	0.000	0.000	0.000
Japan	1.97	10.80	-2.661	0.016	2.398	2.398	2.398	2.398	0.001	0.001	0.000
Great Britain	1.05	28.88	-4.112	0.001	3.408	3.406	3.407	3.408	0.002	0.001	0.001
Ireland	2.58	52.78	-4.453	0.000	3.640	3.637	3.638	3.639	0.002	0.001	0.001
Hong Kong	2.87	30.20	-2.874	0.009	2.598	2.598	2.598	2.598	0.001	0.001	0.000

Note: test statistics were obtained after estimating specification (6) with data in table 3. t_i is studentized residual $\hat{u}_i \big/ \hat{s}_{(i)}\sqrt{1-v_i}$, calculated with all observations. The critical value for the maximum studentized residual is given by $t_{n-k-1}^{\alpha/n}$, and for $\alpha = 0.05$ it equals $t_{21}^{0.05/26} = 3.54$. e_i denotes absolute residuals from the robust LTS fit. Column with heading $\widetilde{t}_i$ contains the sequence of $\widetilde{t}_{k+2}^{\min}$, $\widetilde{t}_{k+3}^{\min}$ …, $\widetilde{t}_{n}^{\min}$ test statistics. z_i is the exact normalizing transformation of $\widetilde{t}_i$. z_i^{*}, z_i^{**} and z_i^{***} are normalizing transformations of $\widetilde{t}_i$ according to (3), (4), and (5).

Table 5. Selected results for 1975 cross-section.

	Recrea-tion	Eduaca-tion	Meat	Fish	Spices	Tobacco	Rice	Bread	Poultry	Fish fresh	Eggs	Fruits fresh	Sugar	Condi-ments
Austria										-4.03			-3.44	
Belgium										-3.47				
Denmark		5.65								-4.29		-4.86		
France														
Germany												-4.11		
Hungary														
Iran	-3.48			-3.57			4.46			4.41				
Ireland												-3.75	-4.40	
Italy														
Japan					4.69		7.33	-4.46					-6.00	
Korea	-4.45					3.72	7.77	-5.73					-6.66	5.30
Luxembourg	-5.21	-3.55								-3.63				
Malaysia							7.08	-4.39						
Netherlands	6.36		-5.29									-3.47		
Philippines		4.87					6.02	-3.58					-3.72	4.27
Poland		4.03	4.15					4.16		4.20				
Romania	4.71					-3.50				3.12			-3.86	
Spain		-3.34												5.07
Sri Lanka		4.34					6.80		-2.96		-3.37			5.61
Syria			-3.94					4.82		-3.84				
Thailand	5.58		-5.72				4.63							7.20
Great Britain								-3.04				-3.99	-4.21	
United States												-5.46		
Yugoslavia			4.33											4.41
Z-statistic														
Actual	3.48	3.34	3.94	3.57	4.69	3.50	4.46	3.04	2.96	3.12	3.37	3.47	3.44	4.27
5% critical	3.88	3.79	3.70	3.77	3.73	3.69	3.18	3.19	3.22	3.27	3.38	3.26	3.27	3.26
p-value	0.156	0.200	0.018	0.088	0.000	0.096	0.002	0.102	0.130	0.084	0.054	0.024	0.030	0.002

Table 6. Selected results for 1980 cross-section.

	Food	Health care	Fish	Spices	Rice	Bread	Noodles	Cereals	Pork	Poultry	Fish fresh	Milk pres.	Cheese	Eggs	Pota-toes	Other tubers	Coffee	Tea	Cocoa	Condi-ments
Austria				4.19																
Belgium																				
Canada				-4.31		-6.00	5.34	6.07	4.22					4.10						
Denmark																				
Finland									-3.87				-5.39							
France													6.70						2.95	
Germany						3.98														
Greece				3.59		-5.64			-3.44				3.84							
Hong Kong	3.92	4.21		-3.40	6.40			5.09						5.20	-4.03	4.03	-7.40	3.42		5.01
Hungary													-5.17	4.01						
Indonesia					5.02				-4.85						-5.19	5.19	-4.96			
Ireland				-4.95									-4.18				-6.41	3.46		
Israel				4.46		-6.11			-4.98	3.29			-3.12							
Italy	4.64						4.36						6.92							
Japan					7.39		3.48				-3.89						-5.07			3.75
Korea					7.77	-5.07									-4.42	4.42				
Luxembourg	-3.43																			
Netherlands						-3.39														
Norway	-3.81					-5.41			-3.58				4.77							
Philippines	5.40				6.05							4.96		-4.45						
Poland																				
Portugal			3.65		3.99	-4.60			-3.31				-5.82							
Spain													-4.64	3.93						
Sri Lanka					6.71	-4.27			-3.94					-5.41	-5.78	5.78	-4.84			3.81
Great Britain								4.51									-6.19			
United States																	-4.47			
Z-statistic																				
Actual	3.43	4.21	3.65	3.40	3.99	3.39	3.48	4.51	3.31	3.29	3.89	4.96	3.12	3.93	4.03	4.03	4.47	3.42	2.95	3.75
5% critical	3.64	3.64	3.69	3.73	3.33	3.33	3.35	3.33	3.32	3.32	3.23	3.37	3.37	3.35	3.20	3.14	3.29	3.35	3.32	3.34
p-value	0.114	0.004	0.058	0.134	0.000	0.036	0.034	0.000	0.058	0.068	0.010	0.000	0.116	0.006	0.000	0.002	0.000	0.038	0.192	0.006

Table 7. Selected results for 1985 cross-section.

	Food	Clothing	House	Health care	Fish	Bevera-ges	Rice	Bread	Noodles	Cereals	Pork	Meat proc.	Milk pres.	Eggs	Coffee	Tea	Cocoa	Sugar	Spirits
Australia						3.67					3.31				-3.45				
Austria											4.93								
Belgium																			
Canada								-3.64				-2.95							
Denmark	-3.95										6.00								
Finland	-4.41			-3.69							-3.53							3.00	
France																			
Germany					4.19														
Greece	5.10				-6.29	3.93						-4.22	3.96						
Hong Kong						4.18	4.28				7.03	-3.98	4.43	3.34		4.42		-7.21	
Ireland					3.88	5.12									-4.99				
Italy				6.12					3.33			-3.65						4.72	
Japan					7.11		6.59				5.05					3.92		-6.98	
Korea	-3.26			-5.10		4.57	7.21	-3.43				-4.32				3.83		-5.49	
Luxembourg																			
Netherlands											5.50								
New Zealand	-4.83														-4.16				
Philippines							5.07					-4.14	5.18						
Portugal		-3.32	3.87				4.39			4.92					-3.70		6.17	3.44	3.19
Spain	-4.20										-4.81	-3.43						5.25	
Sri Lanka				-5.79			5.97				-7.38	-3.14	5.32		-3.17	4.35			
Sweden					4.83														
Thailand			-4.31	-4.11		5.52												-6.76	
Great Britain															-3.64				
United States				5.32	-5.57														
Z-statistic																			
Actual	3.26	3.32	3.87	3.69	3.88	3.67	4.28	3.43	3.33	4.92	3.31	2.95	3.96	3.34	3.17	3.83	6.17	3.00	3.19
5% critical	3.74	3.75	3.75	3.74	3.70	3.70	3.35	3.35	3.35	3.35	3.22	3.22	3.35	3.36	3.35	3.36	3.34	3.32	3.24
p-value	0.182	0.148	0.032	0.058	0.020	0.194	0.002	0.038	0.058	0.000	0.030	0.170	0.004	0.056	0.096	0.010	0.000	0.136	0.064