

LSMS Working Paper
Number 85

Demand Analysis and Tax Reform in Pakistan

Angus Deaton and Franque Grimard

The World Bank
Washington, D.C.

Copyright © 1992
The International Bank for Reconstruction
and Development / THE WORLD BANK
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

All rights reserved
Manufactured in the United States of America
First printing February 1992

To present the results of the Living Standards Measurement Study with the least possible delay, the typescript of this paper has not been prepared in accordance with the procedures appropriate to formal printed texts, and the World Bank accepts no responsibility for errors.

The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of the data included in this publication and accepts no responsibility whatsoever for any consequence of their use. Any maps that accompany the text have been prepared solely for the convenience of readers; the designations and presentation of material in them do not imply the expression of any opinion whatsoever on the part of the World Bank, its affiliates, or its Board or member countries concerning the legal status of any country, territory, city, or area or of the authorities thereof or concerning the delimitation of its boundaries or its national affiliation.

The material in this publication is copyrighted. Requests for permission to reproduce portions of it should be sent to the Office of the Publisher at the address shown in the copyright notice above. The World Bank encourages dissemination of its work and will normally give permission promptly and, when the reproduction is for noncommercial purposes, without asking a fee. Permission to photocopy portions for classroom use is not required, though notification of such use having been made will be appreciated.

The complete backlist of publications from the World Bank is shown in the annual *Index of Publications*, which contains an alphabetical title list (with full ordering information) and indexes of subjects, authors, and countries and regions. The latest edition is available free of charge from the Distribution Unit, Office of the Publisher, Department F, The World Bank, 1818 H Street, N.W., Washington, D.C. 20433, U.S.A., or from Publications, The World Bank, 66, avenue d'Iéna, 75116 Paris, France.

ISSN: 0253-4517

Angus Deaton is Professor of Economics and Public Affairs in the Department of Economics at Princeton University. Franque Grimard is a Ph.D. candidate at the Department of Economics at Princeton University.

Library of Congress Cataloging-in-Publication Data

Deaton, Angus.

Demand analysis and tax reform in Pakistan / Angus Deaton and Franque Grimard.

p. cm. — (LSMS working paper, ISSN 0253-4517 ; no. 85)

Includes bibliographical references.

ISBN 0-8213-2059-9

1. Taxation—Pakistan. 2. Pricing—Pakistan. 3. Food—Prices—Pakistan. 4. Household surveys—Pakistan. I. Grimard, Franque.

II. Title. III. Series.

HJ5168.5.D43 1992

336.2'05'095491—dc20

92-3418

CIP

ABSTRACT

Pakistan, like many LDCs, derives most of government revenue from indirect taxation. However, the system of taxes and subsidies has grown up piecemeal over the years, and it is unlikely that it cannot be improved. Any tax reform proposal must deal with the basic issues of equity and efficiency. Changes in taxes benefit or hurt people according to their demand and supply patterns, while the effects on government revenue depend on the way in which supply and demand respond to prices. Empirical demand analysis can inform the debate by delineating consumption patterns, and by estimating the responsiveness of demand to price. Previous exercises in price reform for Pakistan have been forced to make very restrictive assumptions about consumer preferences, and have typically used demand systems that prejudice what are the desirable directions of price reform.

In this paper, the methodology of Deaton (1988, 1991) is extended and applied to the 1984–85 Household Income and Expenditure Survey. A theory of quality variation based on separable preferences is developed, and the implications for welfare and empirical analysis laid out. The prices of oils and fats and of sugar do not vary very much in the survey data, and the symmetry and homogeneity restrictions from the theory play an important part in obtaining sharp estimates of own and cross-price elasticities. The parameter estimates suggest that there are significant cross-price elasticities between the high-calorie foods, wheat, rice, sugar, and oils, and the presence of these substitution patterns means that the effects of potential price reforms are quite different from those that would be estimated using the traditional and more restrictive assumptions. Based on demand patterns alone, it would be desirable to raise government revenue by raising the *consumer* price of rice. However, in Pakistan it is not generally possible to decouple the producer and consumer prices of rice, so that a full analysis of policy change would also depend on the supply responses, which are not considered in this study.

ACKNOWLEDGEMENTS

We are grateful to Harold Alderman, Valerie Kozel, David Newbery, Nicholas Stern, and especially Dave Coady, for help in the preparation of this paper. Deaton is grateful to the Lynde and Harry Bradley Foundation for research support, and to the Department of Applied Economics, Cambridge, and ST/ICERD at the London School of Economics for hospitality.

FOREWORD

This paper is about "getting prices right." It is an empirical examination of food prices in Pakistan, and of what changes in prices would mean for the allocation of resources and for the real incomes of consumers, many of whom are very poor. The results show how important it is to take into account substitution between different foods when considering a reform to the price of one of them. For the case of Pakistan, the low consumer price of rice is desirable on neither efficiency nor distributional grounds although remedying the situation in practice is likely to affect producer prices and producer incentives, the effects of which are not dealt with in this paper.

This work is part of a broader program of research in the Population and Human Resources (PHR) Department on the extent of poverty in developing countries and on policies to reduce poverty, especially in the context of general price and policy reform. Aside from the substantive findings for Pakistan, one of the objectives of this and similar work is to demonstrate the usefulness of household survey data for policy discussions in developing countries.



Ann O. Hamilton
Director

Population and Human Resources Department

TABLE OF CONTENTS

Section 1.	Introduction	1
Section 2.	Demand Analysis of the 1984-85 HIES	4
Section 3.	The Analysis of Tax and Price Reform	27
	References	37
	Appendix: Alternative Methods of Estimation	39

LIST OF TABLES

Table 1	Structure of the 1984-85 HIES Sample	5
Table 2	Budget Shares, Unit Values, and Fractions Consuming and Buying	6
Table 3	Price Variation by Province, Quarter, and Village: Rural Pakistan 1984-85	8
Table 4	Income and Household Size Coefficients in Share and Unit Value Regressions	11
Table 5	Estimates of Own and Cross-Price Elasticities: Pakistan 1984-85	19
Table 6	Estimates of Own and Cross-Price Elasticities: Pakistan 1984-85	25
Table 7	Efficiency Aspects of Price Reform: Rural Sector	33
Table 8	Equity and Cost-benefit Ratios for Tax Increase	35
Table A1	Alternative Estimates of Total Expenditure and Price Elasticities	42

Introduction

As is the case in many developing countries, the government of Pakistan derives most of its tax revenue from indirect taxation. In the 1970's and 80's, more than 80% of federal government revenue was collected from indirect taxes, some half of which is accounted for by customs duties, see Ahmad and Stern (1991, Table 2.5). The tax and subsidy system generates major distortions in relative prices. In the mid-80s, domestic wheat and rice prices were approximately 70% of world prices at official exchange rates, while the price of refined sugar was more than 60% above world prices. Since the tax and subsidy policies that account for these differentials have grown up piecemeal over the years, the structure can almost certainly be improved, even given the revenue requirements of the government. The theory of tax reform has been extensively developed in recent years, and Ahmad and Stern's work applies that theory to the case of Pakistan. The basic ingredients are the usual ones of equity and efficiency. Changes in taxes benefit or hurt people depending on how much they buy or sell of the goods in question, so that the effects on equity depend on demand and supply patterns, as well as on the extent to which the government is concerned to use taxes and subsidies as a tool of welfare policy. The efficiency effects, by contrast, depend on how demand responds to price changes, the estimation of which is the main topic of this paper.

The effects of a tax change on tax revenue depend on how much of the good is consumed (or supplied), and on how the demand for the good and for other goods changes in response to the price change. It is generally important to consider not only *own-price* responses, but also *cross-price* effects. In principle, it is clear that substitution effects between goods mean that changes in the price of one may have substantial effects on the tax yields from others. Theoretical work has also shown that both optimal tax structures and desirable directions of tax reform are sensitive to the way in which the price effects are specified. For example, if it is assumed that cross-price elasticities are negligible, efficiency considerations will make it desirable to tax more highly those goods whose price elasticities are low, a prescription that would typically lead to differential tax rates. But such a result is not at all robust. Equally plausible assumptions about preferences can easily generate recommendations for uniform tax rates, see Deaton (1981, 1987a), even in the absence of any concern for the distribution of income. One much used set of preferences, the linear expenditure system, implies that it will always be desirable to increase taxes on lowly taxed goods, and decrease

taxes on highly taxed goods. In this case, not only is uniformity desirable, but any move towards uniformity is itself welfare improving. Given these widely different results, it is important that proposals for tax reform should be firmly based on measurement, and not on arbitrary prior assumption.

While few would deny the desirability of measurement in principle, there remain formidable difficulties in the way of estimating price responses in Pakistan as in most other developing countries. Demand analysis in developed economies has typically relied on intertemporal price variation to identify demand responses. Accurate estimation of a full set of own and cross-price elasticities typically requires a large number of observations, and such time-series data are not available for most countries in the world. There are a number of solutions to this problem. The first is to make a greater use of theoretical assumptions, so that lesser requirements are placed on the data. This is the route followed by Ahmad and Stern, who use a variant of the linear expenditure system described in full in Ahmad, Ludlow and Stern (1988). The second is to make use of spatial variation in prices. Alderman (1988) has followed this route and has estimated a version of Deaton and Muellbauer's (1980a) almost ideal demand system using expenditure data from 1979 and 1982 household surveys matched to independent regional price information. These prices are for 12 selected urban areas only, so that for rural households, or for urban households far from collection centers, the match may be a poor one and the price inaccurate.

In this paper, we apply and extend the methodology of Deaton (1988, 1991) to data from the 1984-85 Household Income and Expenditure Survey (HIES). As in Alderman's work, we rely on spatial variation in price to identify the demand responses. However, rather than using the independent price data, which is all that was available to Alderman, we use the price data collected in the HIES itself. Each household who buys a good reports the amount bought as well as the amount paid, so that a price, or at least a unit value, can be derived. While these unit values are not prices and cannot be treated as such, so that the econometric analysis is not completely straightforward, there are as many unit values as there are households purchasing each good, so that there are a very large number of observations to work with. Section 2, which is the main section of the paper, shows how the methodology can be applied, and works through the calculations. We also explain how demand analysis with unit values is related to standard demand theory, and how the restrictions from the theory can be used to estimate a complete demand system using data for only a subset of goods.

Section 3 examines the implications of the empirical results for tax reform in Pakistan, tracing out the efficiency and distributional consequences of possible price changes. The results are quite different from what would be obtained without a flexible method for estimating the effects of price changes. Commodities such as wheat have low price elasticities and low income elasticities, so that they are good candidates for taxation on one count, and poor candidates on the other. But our estimates suggest that edible oils are income inelastic but price elastic, so that taxes are undesirable on both distributional and efficiency grounds. The estimates also suggest that there are important interactions among the goods whose prices are distorted, particularly wheat, rice, oils, and sugar, so that policies that change the price of one will have quantitatively important effects through the taxes and subsidies on the others. A brief Appendix considers some alternative procedures, including Alderman's, and compares the results.

Demand Analysis of the 1984-85 HIES

The household survey that is used here is the National Household Income and Expenditure Survey which collected data from 9,119 rural and 7,461 urban households in Pakistan between July 1984 and June 1985. Table 1 shows the distribution of both urban and rural sample households over the four provinces of Pakistan. For each province and sector, the households are distributed approximately equally over the four quarters of the sample year. The table also shows the numbers of "clusters" in the survey, and their distribution by sector and province. These clusters play an important part in the demand analysis below. The HIES of Pakistan, like other household surveys, minimizes travel costs by selecting, not individual households, but villages or, in urban areas, blocks, and then interviews groups of households within each cluster. Since these households are located near to one another, and are interviewed within a few days of one another, it is reasonable to assume that they face the same prices for the goods that they buy, particularly in rural areas, where there is often a single village market. Note that there are between 10 and 12 households in each cluster. Such figures are common in household surveys; for surveys of different sizes, it is typically the number of clusters that varies, with the size of each cluster much the same.

This paper is concerned with the demand for seven foods, listed in Table 2, which also shows the fractions of *total* consumption accounted for by each. Note that these budget shares are not the shares in the aggregate of consumption over all households, but the average of the budget shares for each of the 9,119 rural and 7,461 urban households. Households with zero budget shares are *included*; the analysis of tax reform requires the effects of price changes on demand averaged over all households, whether they purchase or not. Rural households spend 51% of their budget on food, urban households rather less at 43%. Apart from the catchall category other food, wheat (including bread) and dairy products are the two largest categories in the table.

Means of unit values are also shown in the Table. These are computed from those households who make market purchases of the commodity under consideration. The columns labelled "%b," show the fractions of households in each sector who make such purchases while the larger fractions, labelled "%c," are those who report consumption of each good, either from market purchases or from own-production. Although the survey contains both quantity and value information for consumption from own-production, we use only the market purchases in computing unit values, and treat

Table 1: Structure of the 1984-85 HIES Sample
(numbers of households and clusters)

Province	Households		Clusters	
	Rural	Urban	Rural	Urban
Punjab	5,474	3,793	440	316
Sindh	1,755	2,270	152	196
N.W. Frontier	1,402	994	110	92
Baluchistan	488	404	55	44
Total	9,119	7,461	757	648

as missing the unit values for those observations where consumption is entirely from own-production. The budget shares are computed from *all* households, and zero consumption, own-production consumption, and market consumption are treated identically.

On average, consumers paid 2.4 rupees per kilo for wheat, either as grain or as bread, and almost twice as much per kilo for rice. Meat and edible oils and fats cost more than 14 rupees per kilo. The degree of price variation between households, as measured by the coefficient of variation, is very different for different commodities. For sugar, and for edible oils and fats, where there is a good deal of state price control, there is very little variation in unit values across the sample. For wheat, rice, meat, and other foods, as well as for food as a whole, the cross-sectional standard deviation is between 20 and 30% of the mean. For dairy produce, the coefficient of variation is very large in both sectors. The definition of this group may be less than ideal, since it contains goods with very different unit values. Much of this can be accounted for (and is accounted for) by the estimation procedures below, but there will be difficulties if the composition within the group has a strong spatial component, which is capable of generating a quite spurious relationship between quantity and unit value.

Since we are going to use the spatial price variation to estimate the demand responses, it is worth looking in more detail at the behavior of the unit values. The top panel of Table 3 shows the coefficients obtained by regressing the unit values on dummies for Sindh, Punjab, and North West Frontier Province, omitting Baluchistan, and on dummies for each of the first three quarters in the survey, omitting the fourth. The bottom panel shows *F*-statistics and degrees of freedom for cluster (village) dummy

Table 2: Budget Shares, Unit Values, and Fractions Consuming and Buying

	Rural					Urban				
	share	%c	%b	u.v.	c.v.	share	%c	%b	u.v.	c.v.
Wheat	12.8	96	68	2.4	29	9.0	95	92	2.3	16
Rice	2.7	68	54	5.4	30	2.0	77	76	5.8	26
Dairy	12.7	97	41	6.0	150	9.5	97	90	5.9*	100
Meat	3.7	88	85	14.3	31	5.3	95	94	16.4	34
Oils & Fats	4.1	85	85	14.2	5	4.3	96	96	14.0	4
Sugar	2.9	90	89	8.2	7	3.0	96	96	8.3	6
Other Food	12.2	100	100	4.9	19	10.2	99	98	5.3	28
All Food	51.0	100	100	5.5	29	43.3	99	99	5.9	23

Note: Share is the mean of the percentage budget share over all households, %c is the percentage of people who consume the good, i.e. who buy it, %b, or consume from own production, u.v. is unit value, i.e. the average rupees per kilo spent on the commodity by households buying it, and c.v. is the coefficient of variation of the unit value. Unit values for other food and for total food are computed as geometric means of the unit values of goods purchased with weights equal to the budget shares of the individual goods. (*) One outlier is excluded from the mean unit value of dairy products.

variables, separately for each of the four provinces. The top panel shows how prices vary between provinces and over time, the bottom panel the extent to which they vary from village to village within each province. Note that since every household in each cluster is interviewed at the same time, there is no need to interact cluster with time dummies. For all seven foods, there are pronounced inter-province price differences. Apart from wheat, which is more expensive in Sindh and NWFP than in Baluchistan, all province means are lower than those in Baluchistan. In accord with the calculations in Table 2, the differences are relatively small for oils and fats and for sugar and, to a lesser extent, for wheat, while dairy products show very large provincial differences. Rice is very much cheaper in Sindh than elsewhere. The seasonal differences are much less marked, at least in 1984-85; they show no obvious pattern of price increase over time.

The bottom panel of Table 3 shows that there is typically a great deal of inter-village variation in unit values. All of the *F*-statistics are significant at conventional significance levels. However, the sample sizes are large, and a better indication of the strength of the inter-village variation is whether the *F*-statistics are larger than the logarithm of the sample sizes in each case, (see Schwartz (1978)). For wheat in all

provinces but Baluchistan, the F -tests are smaller than this criterion, and the same happens in a few other cases, rice in Sindh, dairy in Punjab, meat in Punjab and NWFP, and sugar in Punjab. In all other cases, the inter-village spatial variation is large enough to meet even this stringent criterion. In order to save space, we have not included the results for the urban sector corresponding to Table 3. The province effects are broadly similar to those in the rural sector, and once again, the seasonal effects are much less pronounced. The F -statistics for the clusters are again large by conventional criteria, but they are not as large as those for the rural sector. Since city blocks are not as geographically separated as are villages, and since the large cities have many urban blocks each, this result is to be expected.

The model that we use to estimate price elasticities is a standard one in which demand is a function of prices, total consumption, and household demographics. In addition, it is recognized that the observed unit values, while moving with prices, will also vary from household to household because different consumers will typically choose different qualities within the same commodity group. We therefore need two equations, one for the budget shares, and one for the unit values. These are written as follows, see Deaton (1990: equations 1-2):

$$w_{Gic} = \alpha_G^0 + \beta_G^0 \ln x_{ic} + \gamma_G^0 \cdot z_{ic} + \sum_{H=1}^N \theta_{GH} \ln p_{Hc} + (f_{Gc} + u_{Gic}^0) \quad (1)$$

$$\ln v_{Gic} = \alpha_G^1 + \beta_G^1 \ln x_{ic} + \gamma_G^1 \cdot z_{ic} + \sum_{H=1}^N \psi_{GH} \ln p_{Hc} + u_{Gic}^1. \quad (2)$$

The relationship between these equations and standard demand functions is explored later in this section. In equation (1), w_{Gic} is the budget share of good G in the budget of household i , living in cluster c . By analogy with the almost ideal demand system, the share is assumed to be a linear function of the logarithm of total household expenditure, x , and the logarithms of the N prices, as well as of a vector of household characteristics z . The remaining terms represent the residual between the share and its conditional expectation. The first component, f_{Gc} , is a cluster fixed effect for good G , assumed to be orthogonal to the prices, while the idiosyncratic term u^0 represents not only taste variations, but also measurement error in the share. From an econometric point of view, the non-standard feature of equation (1) is that the prices are not observed. Instead, we have data on the unit values v , which are not identical to the prices, but are related to them by equation (2). The logarithm of unit value is a function

Table 3: Price Variation by Province, Quarter, and Village: Rural Pakistan 1984-85

(regressions of unit values on provincial and seasonal dummies ANOVA by village)

	Wheat		Rice		Dairy Products		Meat		Oils & Fats		Sugar		Other Food	
REGRESSION PARAMETER ESTIMATES														
Punjab	-7.2	(6.0)	-8.0	(4.5)	-68.7	(16)	-22.1	(16)	-2.0	(8.5)	-6.6	(22.)	-9.7	(12)
Sindh	7.3	(5.2)	-43.1	(22)	-78.3	(15)	-13.8	(9.5)	-2.6	(10)	-2.9	(9.0)	-11.8	(13)
NWFP	6.3	(4.7)	-2.8	(1.4)	-26.0	(5.4)	-31.1	(21)	-1.3	(5.0)	-3.7	(11.)	-6.3	(6.9)
Quarter 1	-2.1	(2.6)	0.4	(0.4)	2.0	(0.7)	-1.4	(1.6)	-0.2	(1.4)	0.3	(1.7)	0.9	(1.7)
Quarter 2	-2.4	(3.0)	-1.0	(0.8)	-1.5	(0.5)	-3.4	(4.1)	-0.5	(2.9)	-0.2	(1.1)	-4.3	(8.4)
Quarter 3	0.0	(0.0)	-2.0	(1.6)	-3.0	(1.1)	-1.9	(2.2)	-0.3	(1.8)	0.3	(1.5)	-6.9	(13)
ANOVA	<i>F</i>	<i>R</i> ²	<i>F</i>	<i>R</i> ²	<i>F</i>	<i>R</i> ²	<i>F</i>	<i>R</i> ²	<i>F</i>	<i>R</i> ²	<i>F</i>	<i>R</i> ²	<i>F</i>	<i>R</i> ²
Punjab	4.02	0.334	9.14	0.593	3.07	0.385	7.85	0.454	13.7	0.608	7.2	0.420	12.8	0.528
Sindh	5.42	0.526	4.40	0.435	9.44	0.776	8.61	0.475	10.4	0.508	19.9	0.662	14.1	0.571
NWFP	6.06	0.408	7.12	0.572	10.63	0.673	4.46	0.304	13.5	0.541	13.9	0.586	10.8	0.477
Balu'stan	9.14	0.582	8.52	0.610	17.70	0.798	10.97	0.607	7.74	0.500	18.0	0.695	17.9	0.691
All	5.65	0.432	10.83	0.646	6.46	0.590	9.30	0.498	12.39	0.571	14.0	0.588	13.79	0.556

Notes: The figures in brackets are absolute *t*-values. The number of observations in each regression or analysis of variance is the number of rural households who report expenditure on the good in question.

of $\ln x$, so that β_G^1 is the elasticity of quality with respect to total expenditure, and of the characteristics z , since household features may affect quality choice. The unobservable prices appear through the matrix Ψ . In the simplest case, this matrix would be the identity matrix, the other variables would be absent, and prices would equal unit values. The extra generality allows for quality “shading” in response to price change, so that we might expect $\psi_{GG} < 1$, with some non-zero off-diagonal terms. Not surprisingly, the Ψ matrix is not identified without further assumptions, and in Deaton (1987b, 1988), a theoretical model is developed that provides suitable prior information, see also the discussion below. In practice, Ψ is close to the identity matrix, so that a simple procedure works as if the approximation were correct, and to make whatever adjustments are necessary at the end.

Readers who would like to see a complete development of the econometric methodology before turning to the results should consult Deaton (1990). In the current paper, we develop the procedure in an intuitive way, presenting results alongside the discussion of the methodology. To begin, note that the α , β , and γ parameters in equations (1) and (2) can be estimated straightforwardly by OLS if we allow for cluster fixed effects by including dummy variables for each cluster. Since the prices are assumed to be the same in each village, the dummy variables will allow for them, as well as for the fixed effects f_{Ge} , and the cluster means of the idiosyncratic errors. However, we see from Table 1 that there are 757 rural and 648 urban clusters, and most computer programs will not allow regressions with so many dummy variables. However, several estimation programs allow for analysis of variance with continuous regressors and a large number of classes. Alternatively, and more transparently, equations (1) and (2) can be estimated by calculating the mean of each variable for each cluster, and then running a regression using the deviations from the cluster means of all the variables. Standard OLS gives the correct parameter estimates, although the usual standard errors have to be multiplied by $\sqrt{(N-k)/(N-k-C+1)}$ where k is the number of regressors, N the sample size, and C the number of clusters. This correction allows for the fact that $(C-1)$ degrees of freedom are removed when the cluster means are subtracted out. Given the figures in Table 1, the correction is quite small; standard errors should be increased by about four percent.

A selection of the results from this first stage of the estimation are given in Table 4. Household composition effects, the z 's in equations (1) and (2), are modelled by entering, together with the logarithm of total household expenditure, the logarithm

of total household size, and the thirteen *ratios*, n_j/n , where n_j is the number of people on age and sex group j , and n is total household size. There are seven age groups for each sex, making fourteen in all; these sum to unity, so only thirteen need be included in the regression. The Table shows only the coefficients on total expenditure and household size; the demographic ratios are typically important in the share equation - the structure of the household affects demand patterns - but much less so in the unit value equation. The first two rows of the table show the coefficients of the logarithms of total household expenditure and total household size in the budget share equation (1). The rural results are shown in the top panel, and the urban in the lower panel. The third row shows the total expenditure elasticity for the good, calculated according to the formula

$$e_i = 1 - \beta_i^1 + \beta_i^0 w_i^{-1}. \quad (3)$$

Note that expenditure on each good responds to total outlay, not only through the effect on quantity, the usual expenditure elasticity e_i , but also through the quality effect, β_i^1 , which must also be allowed for in (3).

The figures shown in the table are evaluated at the sample means of the budget shares as reported in Table 2. Negative values of β correspond to expenditure elasticities less than unity, and this is the case for all goods except dairy products and meat. Wheat and edible oils and fats are the least elastic of the goods in the table, and the only ones with elasticities less than a half. These figures imply an expenditure elasticity for food as a whole of approximately 0.8 in both sectors. In all cases, the coefficient on household size is of the opposite sign to that on total expenditure; increases in household size, with composition held constant, affect the pattern of demand as do decreases in expenditure. Indeed, in several cases, and most precisely for wheat, the coefficients are close to being equal and opposite, so that, conditional on household composition, it is only *per capita* household expenditure that matters. Changes in household size that are matched by proportional changes in total expenditure have no effect on the share of the budget devoted to wheat. This is only approximately true for the other foods. Note the very close correspondence between the rural and urban estimates; there is no particular reason why this has to be happen, but large differences would have called for some explanation.

Quality elasticities are estimated from the effect of (the logarithm of) total expenditure on the (logarithm of the) unit value. These estimates are shown in line 4 of

Table 4: Income and Household Size Coefficients in Share and Unit Value Regressions
(within cluster regressions)

	Wheat	Rice	Dairy	Meat	Oils & Fats	Sugar	Other Food
RURAL							
<i>w:lnx</i>	-0.068 (56)	-0.006 (7.6)	0.024 (13)	0.009 (14)	-0.024 (42)	-0.004 (8.2)	-0.040 (40)
<i>w:lnhhs</i>	0.070 (49)	0.007 (8.8)	-0.007 (3.2)	-0.006 (8.0)	0.013 (19)	0.002 (4.1)	0.020 (17)
elasticity wrt <i>x</i>	0.374	0.684	1.053	1.095	0.417	0.865	0.593
<i>lnv:lnx</i>	0.095 (15)	0.107 (12)	0.139 (6.4)	0.154 (24)	0.000 (0.4)	0.006 (4.6)	0.076 (21)
<i>lnv:lnhhs</i>	-0.033 (4.3)	-0.062 (6.3)	-0.031 (1.3)	-0.099 (13)	0.000 (0.0)	-0.004 (2.8)	-0.033 (7.9)
URBAN							
<i>w:lnx</i>	-0.053 (53)	-0.004 (6.8)	0.001 (0.7)	0.009 (11)	-0.023 (44)	-0.011 (27)	-0.030 (32)
<i>w:lnhhs</i>	0.054 (47)	0.006 (11)	0.008 (5.0)	-0.000 (0.5)	0.019 (31)	0.009 (20)	0.020 (19)
elasticity wrt <i>x</i>	0.388	0.711	1.001	0.938	0.470	0.648	0.580
<i>lnv:lnx</i>	0.027 (7.6)	0.111 (16)	0.010 (0.9)	0.242 (34)	0.003 (2.6)	-0.000 (0.3)	0.129 (26)
<i>lnv:lnhhs</i>	-0.028 (6.7)	-0.075 (10)	-0.023 (1.9)	-0.156 (19)	-0.002 (2.1)	-0.000 (0.1)	-0.076 (13)

Notes: *w* is the budget share, *lnv* the unit value, *x* total expenditure, and *hhs* household size so that *w:lnx* is the regression coefficient of the regression of *w* on *lnx*, and similarly for the other rows. The elasticity in row 3 is the total expenditure elasticity of *quantity*; the total expenditure elasticity of expenditure on the commodity is the sum of this and the quality elasticity in the next row. Absolute values of *t*-values are given in parentheses.

each panel, and the corresponding coefficient of the logarithm of household size in line 5. With the exception of sugar in the urban sector, where the coefficient is essentially zero, all of these estimates are positive, as is to be expected, but none are very large. The unit value of meat rises with total expenditure, with elasticities of 0.15 in the rural and 0.24 in the urban sector. There are several other cases (rice, other food, wheat and dairy produce in the rural sector) with quality elasticities around 10%. Note that for the two essentially homogeneous goods, sugar and edible oils, the elasticities are zero, as they should be. As was the case for the budget shares, increases in household size typically act in much the same way as decreases in expenditure. This is the major effect of demographics on quality; conditional on total expenditure and household size, age and sex composition have little effect on quality choice.

The estimates in Table 4 are the final estimates for the effects of total expenditure and the demographics. Note that they are derived entirely from *within* cluster information. The next step is to use the *between* cluster information in the data to estimate the price responses. We first use the estimates from the first stage to calculate cluster averages of the shares and unit values “purged” of expenditure and demographic effects. Define:

$$\bar{y}_{Gc}^0 = \frac{1}{n_c} \sum_{i \in c} (w_{Gic} - \tilde{\beta}_G^0 \ln x_{ic} - \tilde{\gamma}_{G \cdot z_{ic}}^0) \quad (4)$$

$$\bar{y}_{Gc}^1 = \frac{1}{n_{Gc}^*} \sum_{i \in c} (\ln v_{Gic} - \tilde{\beta}_G^1 \ln x_{ic} - \tilde{\gamma}_{G \cdot z_{ic}}^1). \quad (5)$$

where n_c is the number of households in cluster c , n_{Gc}^* is the number of households who purchase good G (and thus report unit values) and superimposed tildes indicate estimates from the first, within-cluster stage. As the number of observations in the first-stage regression increases, the estimates will tend to their true values, so that \bar{y}_{Gc}^0 and \bar{y}_{Gc}^1 will tend to the true cluster means, which, by (1) and (2), can be written:

$$\bar{y}_{Gc}^0 = \alpha_G^0 + \sum \theta_{GH} \ln p_{Hc} + f_{Gc} + u_{Gc}^0 \quad (6)$$

$$\bar{y}_{Gc}^1 = \alpha_G^1 + \sum \psi_{GH} \ln p_{Hc} + u_{Gc}^1. \quad (7)$$

where u_{Gc}^0 and u_{Gc}^1 are the cluster means of the errors in (1) and (2). Note that in the construction of (4) and (5), the right hand side variables *do not* have the means excluded;

the cluster means contain the information about the prices that must be excluded from the first stage regressions, but must be included in the between cluster regressions if the price effects are to be identified.

If the matrix Ψ were the identity matrix, and if the cluster means u_{Gc}^0 and u_{Gc}^1 were zero, the columns of the matrix Θ of price effects could be estimated by regressing each y_{Gc}^0 on the matrix of corrected prices y_{Gc}^1 . This is in fact very close to what we do. Given that the quality effects in Table 4 are not large, the correction for the Ψ matrix is not the most important one, but it is necessary to correct for the fact that the sample clusters are not large enough to allow us safely to ignore the averages of the measurement errors. The procedure is a standard errors in variables one, which allows both for the measurement error in the prices, and any possible correlation between the measurement errors in the price and share equations.

To see how this works, define the variance and covariance matrices corresponding to (6) and (7):

$$\begin{aligned} q_{GH} &= \text{cov}(y_{Gc}^0, y_{Hc}^0), & s_{GH} &= \text{cov}(y_{Gc}^1, y_{Hc}^1), \\ r_{GH} &= \text{cov}(y_{Gc}^0, y_{Hc}^1). \end{aligned} \quad (8)$$

By the definition of ordinary least squares, the OLS estimator would be $S^{-1}R$, a feasible version of which is $\tilde{S}^{-1}\tilde{R}$, where the tildes show that the covariances in (8) are evaluated using the estimates in (4) and (5) rather than (6) and (7). The problem with these estimators is that the variance covariance matrix \tilde{S} overestimates the variance covariance matrix of the true prices, because it includes the effects of the measurement error in (7); similarly \tilde{R} is contaminated by any covariances in the measurement error between the two equations. But the variances and covariances of the errors can be estimated from the first stage regressions, and the estimates used to make the correction. Let e_{Gic}^j , $j=1, 2$ be the residual for household i , cluster c , good G , from the first stage share regression ($j = 0$) or unit value regression ($j = 1$). Define:

$$\begin{aligned} \tilde{\sigma}_{GH} &= (n-C-k)^{-1} \sum_c \sum_i e_{Gic}^0 e_{Hic}^0, & \tilde{\omega}_{GH} &= (n_G^*-C-k)^{-1} \sum_c \sum_i (e_{Gic}^1)^2, \\ \tilde{\gamma}_{GH} &= (n_G^*-C-k)^{-1} \sum_c \sum_i e_{Gic}^1 e_{Hic}^0 \end{aligned} \quad (9)$$

where n_G^* is the total number of households in the survey who report purchases of good G . Denote the matrices defined in (9) by $\tilde{\Sigma}$, $\tilde{\Omega}$, and $\tilde{\Gamma}$, and their limits Σ , Ω , and Γ . Then from (6) and (7):

$$S = \Psi M \Psi' + \Omega N_*^{-1}, \quad R = \Psi M \Theta' + \Gamma N^{-1}, \quad (10)$$

where M is the unobservable variance covariance matrix of the true price vector, $N_*^{-1} = \text{plim} C^{-1} \Sigma_c D(n_c^*)^{-1}$, $D(n_c^*)$ is a diagonal matrix formed from the elements of n_{Gc}^* , and N^{-1} is the corresponding matrix formed from the n_c 's. To eliminate the effects of the measurement error, we need to correct OLS by removing the second terms on the right hand side of (10). This leads to the estimator:

$$\tilde{B} = (S - \tilde{\Omega} \tilde{N}_*^{-1})^{-1} (\tilde{R} - \tilde{\Gamma} \tilde{N}^{-1}) \quad (11)$$

where \tilde{N}_*^{-1} and \tilde{N}^{-1} correspond to the sample averages instead of the probability limits. From (10), it is immediate that, taking probability limits as the sample size goes to infinity but with the cluster sizes remaining fixed:

$$\text{plim } \tilde{B} = B = (\Psi')^{-1} \Theta'. \quad (12)$$

To interpret these results, start with the last, equation (12). If Ψ is the identity matrix, \tilde{B} converges to the transpose of the matrix of price responses Θ , which is what we want. If $\Psi \neq I$, then B is all that can be identified without further theory, and we cannot go further with the data alone. From (11) we see that, in the absence of measurement error, in which case the matrices Ω and Γ would be zero, the estimator reduces to the ordinary least squares estimator, as it should. Even if these matrices are not zero, large cluster sizes will make the post-multiplying diagonal matrices small, so that, once again, we approach the OLS estimator. In this case, averaging over clusters is enough to remove the measurement error and the price response matrix can simply be estimated by least squares once the expenditure and demographic effects have been removed at the first stage. The present procedure is more general and allows for the possibility that measurement error is sufficiently severe, or cluster size small enough, so that even the averages are contaminated. This seems wise, since cluster sizes are often in single figures, and the number of purchasers of a good in each cluster will often be only two or three.

A full discussion of how to obtain the variance covariance matrix of β is given in the Appendix of Deaton (1990). The formulae are complex because a complete treatment requires that allowance be made for the facts (i) that the error covariance matrices $\tilde{\Gamma}$ and $\tilde{\Omega}$ in (12) are estimated, not known, and (ii) that the calculated variances and covariances of the “purged” shares and prices also contain estimates of the first stage parameters. However, it seems that these considerations have very little effect on the calculations in practice, and they certainly do not do so in the current case, presumably because the first-stage parameters are so very well-determined. The main contribution to the variance is the sampling variability of the cluster fixed effects f in (1), which are analogous to the disturbances in the between-cluster regression. Of course, the standard OLS variance-covariance matrix formulae have to be modified for the errors in variables case. From equation (A.12) in Deaton (1990), we have:

$$V\{\text{vec}(\tilde{B})\} \approx C^{-1}(P'HP \otimes A^{-1}JHJ'A^{-1}) + C^{-1}(P'HJ'A^{-1} \otimes A^{-1}JA J'A^{-1})K \quad (13)$$

where $A = (S - \Omega N_*^{-1})$, $J = (O_N | I_N)$, and $P' = (I_N | -B')$. The matrices H and Λ are formed from the variance covariance matrices of the measurement errors and the data according to:

$$H = \begin{bmatrix} Q & R' \\ R & S \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \Sigma & \Gamma' \\ \Gamma & \Omega \end{bmatrix} \quad (14)$$

K is the $2N^2 \times 2N^2$ commutation matrix; it is the matrix of ones and zeros with the property that $K\text{vec}(A) = \text{vec}(A')$ for an arbitrary conformable matrix A . Equation (13) is not exact; it ignores the sampling variability of the first-stage estimates, but it has proven accurate in other applications, and the additional terms make no noticeable difference in the current case.

Together with the first stage results, the matrix B yields everything that we need to know. However, it is usual as well as convenient to present the results of demand analysis in terms of elasticities, and a little further work is needed to convert the results to that form. Note first that, since the budget share is the product of quantity and unit value divided by outlay, its derivatives with respect to the logarithms of the prices contains terms representing, not only the response of quantity to price, but also of unit

value with respect to price. If w_G is differentiated with respect to $\ln p_H$, rearrangement gives:

$$e_{GH} = -\psi_{GH} + \theta_{GH} w_G^{-1}. \quad (15)$$

Given appropriate separability assumptions, it is possible to express the responses of unit value to prices, the Ψ matrix, in terms of the total expenditure elasticities of quality. Deaton (1988) derives the formula:

$$\psi_{GH} = \delta_{GH} + \beta_G^1 e_{GH} e_G^{-1} \quad (16)$$

where e_G is the total expenditure elasticity. Equations (15) and (16) can be used to eliminate the matrix of price elasticities, yielding a relationship between Ψ and Θ . But equation (12) also links Ψ and Θ with the estimated matrix B , so that both Θ and Ψ , and thus the matrix of price elasticities E , can be calculated from the estimated parameters. In matrix notation the formulas are:

$$\Theta = B' \{I - D(\xi)B' + D(\xi)D(w)\}^{-1} \quad (17)$$

$$E = \{D(w)^{-1}B' - I\} \{I - D(\xi)B' + D(\xi)D(w)\}^{-1} \quad (18)$$

where w is the vector of budget shares, and the elements of ξ are given by

$$\xi_G = \{(1 - \beta_G^1)w_G + \beta_G^0\}^{-1} \beta_G^1 \quad (19)$$

In order to save space, we report only the estimates of the elasticity matrices, together with their standard errors. These can be derived from the variance covariance matrix of B , (13), using the formula:

$$V\{\text{vec}(E')\} = \{(D(w)^{-1} + ED(\xi)) \otimes G\} V\{\text{vec}(B)\} \{(D(w)^{-1} + D(\xi)E') \otimes G'\} \quad (20)$$

$$G = \{I - D(\xi) + D(\xi)D(w)\}^{-1}.$$

This expression, like that for the variances of the B matrix, ignores the sampling variability of the first-stage estimates. Again, see Deaton (1990) for a full treatment.

We can now summarize the second stage of the estimation as follows. The first-stage parameter estimates are used to calculate equations (4) and (5), giving a “purged” unit value and budget share for each good for each cluster. Clusters that do not have a complete set of unit values are dropped. Note that we are thereby dropping only those clusters where at least one of the goods is bought by no one; it need not be the case that anyone buys all of the goods. Even so, 137 rural clusters are lost, leaving 620 for the estimation; only 10 clusters are lost in the urban sector, leaving 638. It is the large number of missing unit values at the individual level that makes it attractive to average over clusters at this stage. Because prices are assumed to be the same within each cluster, the averaging does not lose price information, and, although a more efficient estimation technique might be based on the individual household data, it would have to deal with the very large numbers of missing values at that level. The first-stage equations are also used to calculate the estimates of the measurement error variances according to (9). We assume that the off-diagonal terms in all three of these matrices are zero. There is no difficulty in principle in generalizing this, although in practice, the off-diagonal terms are typically very small.

A decision also has to be made about the extent of spatial and temporal variation that the prices are asked to explain. The issue is whether there are broad regional taste differences that we do not wish to attribute to regional price differences. To the extent that the inhabitants of, say, the North West Frontier have different tastes from those in Sindh, we should allow (at the least) for dummy variables for those regions. Within provinces, many of the spatial differences in prices will be of long standing, so that consumers have had a great deal of time to adjust to them. In consequence, our procedures are likely to estimate long-run price responses, which would not be expected to be the same as the seasonal responses to short term price changes. We therefore allow for dummy variables for each of the quarters of the survey. The econometric procedures remain as discussed above, but having calculated (4) and (5), these cluster means are first regressed on province and quarterly dummies, and the residuals carried forward for further analysis.

The estimates of the matrix of price elasticities, evaluated at the sample means of the budget shares, are shown in Table 5. The top panel refers to the rural sector, the bottom the urban sector. The final row and column, for non-food, is derived in a

manner that will be discussed below. The estimates are typically well-determined, at least in comparison to estimates from time-series data. All of the diagonal terms, the own-price elasticities, are negative, and several (rice and edible oils in both sectors, and dairy products in the urban sector) are less than -1. There is again a good deal of conformity between the rural and urban estimates of these own-price elasticities. The effects of the sugar and edible oil prices, columns five and six, have relatively large standard errors; recall from Table 2 that these prices show little variation in the sample. Particularly in the urban sector, these two columns contain several estimates that are large in absolute value. For example, wheat, rice and meat are all estimated as being very (positively) elastic to the oil price, and rice (negatively) to the sugar price. These figures have large standard errors and should be treated accordingly.

Apart from the effects of the oil and sugar prices, the cross-price terms are typically not very large. The urban estimates show a large substitution effect of the wheat price on rice purchases, an effect that is much smaller, and not at all significant, in the rural estimates. In the rural matrix, oil and sugar apart, the numerically largest cross-price elasticities are -0.43 for the price of other food on dairy products, and 0.43 for the meat price on rice purchases, and the latter effect is even stronger in the urban estimates.

None of these cross-price effects, except possibly that between wheat and rice, are large enough, or precise enough, to change, by itself, the way in which we might think about taxes. However, the configuration of these elasticities is very different from what would be the case if alternative assumptions had been made. In particular, the assumption of additive preferences, as for example in the linear expenditure system, not only restricts the cross-price elasticities to be small, as indeed is mostly the case here, but also enforces an approximate proportionality between total expenditure and own-price elasticities. This proportionality has dramatic effects for tax policy, because it means that goods that are attractive to tax on efficiency grounds, those with low price elasticities, are those most heavily consumed by the poor, and are thus the least attractive on distributional grounds. There is therefore a balance between equity and efficiency that tends to drive taxes towards uniformity. The estimates in Table 5 show no such proportionality. Table 3 estimates that meat and dairy produce are luxuries and rice a necessity, and yet Table 5 estimates that rice is much more price elastic than is either dairy produce or meat. Taxing rice is therefore doubly unattractive compared with taxing either of the other two goods. We shall return to these issues in more detail in

Table 5: Estimates of Own and Cross-Price Elasticities: Pakistan 1984-85

	Wheat	Rice	Dairy Produce	Meat	Oils & Fats	Sugar	Other Food
RURAL							
Wheat	-0.51 (0.09)	0.08 (0.04)	0.03 (0.02)	0.01 (0.06)	-0.04 (0.29)	0.04 (0.22)	0.09 (0.09)
Rice	0.30 (0.29)	-1.59 (0.13)	-0.11 (0.08)	0.43 (0.20)	-0.18 (0.97)	-0.39 (0.75)	-0.32 (0.30)
Dairy	0.04 (0.12)	0.13 (0.05)	-0.87 (0.03)	-0.15 (0.08)	0.19 (0.40)	0.84 (0.31)	-0.43 (0.13)
Meat	-0.28 (0.20)	0.14 (0.09)	-0.06 (0.05)	-0.57 (0.13)	-1.54 (0.66)	-0.62 (0.51)	0.34 (0.21)
Oils & Fats	0.22 (0.13)	0.23 (0.06)	0.04 (0.04)	0.24 (0.09)	-2.04 (0.44)	-0.31 (0.34)	0.08 (0.14)
Sugar	0.14 (0.17)	0.29 (0.08)	0.04 (0.05)	0.22 (0.11)	-0.99 (0.56)	-0.07 (0.44)	0.16 (0.18)
Other Food	0.18 (0.08)	0.05 (0.03)	0.01 (0.02)	0.11 (0.05)	-0.44 (0.25)	0.21 (0.20)	-0.35 (0.08)
URBAN							
Wheat	-0.86 (0.13)	0.08 (0.05)	-0.07 (0.05)	-0.21 (0.09)	2.26 (0.74)	0.41 (0.35)	-0.05 (0.11)
Rice	1.44 (0.24)	-1.83 (0.10)	-0.45 (0.09)	0.64 (0.17)	1.96 (1.41)	-2.40 (0.66)	0.14 (0.20)
Dairy	-0.22 (0.12)	-0.18 (0.05)	-1.10 (0.04)	0.07 (0.08)	0.58 (0.68)	0.60 (0.32)	-0.14 (0.10)
Meat	-0.20 (0.16)	-0.08 (0.07)	0.21 (0.06)	-0.33 (0.11)	2.66 (0.95)	-0.22 (0.45)	-0.03 (0.14)
Oils & Fats	0.09 (0.10)	0.08 (0.04)	-0.01 (0.04)	-0.09 (0.07)	-1.17 (0.58)	-0.13 (0.27)	-0.04 (0.08)
Sugar	-0.05 (0.15)	0.14 (0.06)	-0.06 (0.05)	-0.14 (0.10)	0.20 (0.85)	-0.81 (0.40)	-0.41 (0.12)
Other Food	0.18 (0.09)	-0.03 (0.04)	-0.00 (0.04)	-0.03 (0.06)	0.06 (0.55)	-0.20 (0.26)	-0.34 (0.08)

Note: The rows refer to the good being affected, and the columns to the price being changed. Hence, row 2, column 7 for the rural sector estimates that a 1% increase in the price of other food will decrease the quantity purchased of rice by 0.32%. The figures in parentheses are standard errors.

Section 3, but for now note only that the results in Table 5 have different implications for tax structures in Pakistan from previous estimates based on models such as the linear expenditure system.

The results in Table 5 are not all that we need to know about demand responses in order to evaluate price reform proposals. The demand for non-foods will generally be affected by price changes in foods, and the demand for food will depend on the prices of non-foods. For most non-foods we do not have data on physical quantities, and so the methods presented above cannot yield the estimates that are required to close the system. However, it is possible to use the results of consumer demand theory to complete the system, at least for a single aggregate of non-foods. The complete system has to satisfy the budget constraint, which yields one set of restrictions, and it must be homogeneous of degree zero, which yields another. We show below that provided we can make a reasonable guess for the quality elasticity of non-foods as a whole, adding up and homogeneity allow us to add another row and column to the matrices of price elasticities in Table 5. In fact, it is possible to do better than this, by using the Slutsky symmetry restriction of demand theory. In the current context, the use of these restrictions is particularly attractive. As we have seen from Table 5, the elasticities in the "sugar" and "oils and fats" columns are not very well determined, an outcome that reflects the low variance of sugar and oils prices in the sample. But other prices do vary a great deal, and symmetry allows us to use, for example, the well-determined effects of rice prices on sugar and oils to measure more precisely the effects of sugar and oils prices on rice. Further, the symmetry restriction is testable, something that is not true of homogeneity when we have data on only a subset of goods. The remainder of this section explains how to complete the system, and how adding up, homogeneity, and symmetry work in a model with quality as well as quantity and price. The econometric extensions are also presented.

A full discussion of the theory of quality and quantity will be presented elsewhere, and only the bare essentials are given here. The basic idea is that quality is an attribute of commodity *aggregates*, so that, for example, a higher unit value for rice means that there are more expensive varieties and fewer cheaper varieties in the rice "bundle" under consideration. There is therefore a group of rice goods, and the way in which rice expenditures are allocated within the group varies systematically with the outlay of the spender. Preferences are assumed to be (weakly) separable between the various commodity groups, so that there exist sub-utility functions for each subgroup,

represented here by the subgroup cost functions, $c_G(u_G, p_G)$ each of which gives the minimum expenditure on group G necessary to reach u_G at group prices given by the vector p_G . Following Deaton and Muellbauer (1980, pp. 130-2), we use these cost functions to define group quantity indices Q_G and group unit value indices v_G by

$$\begin{aligned} Q_G &= c_G(u_G, p_G^0) \\ v_G &= c_G(u_G, p_G) / c_G(u_G, p_G^0) \end{aligned} \quad (21)$$

The vector p^0 is some convenient base vector that is used to measure quantity: it need not necessarily be a price, and quantities can be measured in physical units, or in calories. The quantity index Q_G is a standard one, and for fixed p^0 is a monotone increasing function of u_G . The unit value v_G is linearly homogeneous in the price vector p_G , as should be any sensible price index, but it is also a function of the group utility level u_G and thus of Q_G . It is the dependence of the unit value on group consumption levels that generates the quality effects, and it is a straightforward exercise to use the second part of (21) to prove the relationship between unit value and price used in equation (16) above, at least under the simplifying assumption that prices in the group move proportionately.

The really useful fact about the definitions in (21) is that they can be used to write the consumer's maximization problem as:

$$\begin{aligned} \text{Max } u &= V(Q_1, Q_2, \dots, Q_G, \dots, Q_N) \\ \text{s.t. } \sum v_G Q_G &= x \end{aligned} \quad (22)$$

This reformulation works because utility is a function of the sub-utilities, which are just monotone transformations of the Q 's, and because total expenditure has to be the sum of the costs on each subgroup. Of course, the familiar formula hides the unfamiliar fact that the "prices" v_G are not prices at all, but unit values, which depend on the very Q 's that are the solution to the problem. This makes (22) an unsuitable vehicle for calculation, but it nevertheless shows that it is possible to write down "standard" demand functions, with quantities functions of x and the *unit values*, and that these demand functions satisfy all the functional restrictions with which we are familiar. In particular, we can adopt an "almost ideal" system functional form and write:

$$w_G = \alpha_G + \beta_G \ln(x/v) + \sum_H c_{GH} \ln v_H \quad (23)$$

where v is a (linearly homogeneous) index of all the unit values. Note again that (23) is more complicated than it looks, because the unit values are functions of the quantities, and therefore cannot be econometrically exogenous. The equation must be supplemented with an equation that links the unit values with total outlay and prices, an equation given by (2), which in skeletal form is

$$\ln v_G = \alpha_G^1 + \beta_G^1 \ln x + \sum_H \psi_{GH} \ln p_H \quad (24)$$

The substitution of (24) in (23) yields equation (1), which has been the basis for the estimation. However, the great value of equation (23) is that it tells us exactly what restrictions should be satisfied by the system. Adding-up implies that the β 's in (23) add to zero, as does each of the columns of the C -matrix. Homogeneity is equivalent to the rows of C adding to zero, and symmetry to the symmetry of C .

To translate these statements into statements about the parameters estimated in this paper, we need to specify the index v , and then explicitly make the substitution of (24) into (23). We adapt Deaton and Muellbauer's suggestion and approximate $\ln v$ by the inner product $\bar{w} \cdot \ln v$, where the bar denotes the mean over the sample. We can then match the parameters in (1) and (2) with the parameters in (23) and (24), whereby

$$\begin{aligned} \Theta \Psi^{-1} &= (C - \beta \bar{w}') = D', \text{ say} \\ \beta &= \beta^0 - D' \beta^1. \end{aligned} \quad (25)$$

The matrix D is the matrix B defined in (12), but, since it corresponds to a complete system of goods that together exhaust the budget, has one more row and column than B ; hence the change of notation. Note also that the vectors β^0 , β^1 and \bar{w} in the expressions above refer to the full system, and therefore have one more element than their counterparts earlier in the paper.

Equation (25) yields the restrictions on D , and hence on B , that are required to complete the system and to impose symmetry. Adding-up is

$${}_{\iota}' D' = 0, \quad {}_{\iota}' \beta^0 = 0, \quad (26)$$

where ${}_{\iota}$ is the vector of units. Homogeneity requires that

$$D'(\iota - \beta^1) + \beta^0 = 0, \quad (27)$$

while the symmetry restriction is that

$$D'(I - \beta^1 \bar{w}') + \beta^0 \bar{w}' = (I - \bar{w} \beta^1')D + \bar{w} \beta^0'. \quad (28)$$

The adding-up and homogeneity restrictions are used to complete the system, that is to add the final rows and columns to the B matrix and thence to the elasticity matrix. The symmetry restriction, by contrast, generates testable restrictions on the matrix D and thus ultimately on B .

The econometric analysis proceeds as follows. We already have an estimate of the B -matrix, in this case 7×7 , as well as an estimate of V , the 49×49 variance covariance matrix of the estimates of $\text{vec} B'$. Adding-up and homogeneity together imply that D can be derived from B by means of

$$D' = \begin{bmatrix} I_7 \\ -\iota_7' \end{bmatrix} B' \begin{bmatrix} I_7 & -\frac{\iota_7 - \beta^1}{1 - \beta_N^1} \end{bmatrix} - \frac{1}{1 - \beta_N^1} \begin{bmatrix} 0 & -\beta^0 \\ 0 & -\beta_N^0 \end{bmatrix} \quad (29)$$

or, in an obvious notation,

$$\text{vec} D' = (J_2 \otimes J_1) \text{vec} B' - \text{vec} J_3 \quad (30)$$

which offers an immediate means of calculation for the full matrix, as well as of its variance covariance matrix. Note that we require estimates of the last elements β_N^0 and β_N^1 , corresponding to the non-food category. Since the elements of β^0 add to zero over the complete set of goods, the former is easily estimated, but the quality elasticity of non-food cannot be estimated without (the non-existent) data on unit values for non-food. Here, we simply assume a value for β_N^1 of 0.10; given the general unimportance of the quality estimates in this study, this need for this obviously arbitrary choice is unlikely to be a serious difficulty. Note too that we are effectively assuming that the first stage parameters, β^0 and β^1 are known; they are treated as constants in the calculation of D and its standard errors. Ideally, the first-stage sampling variability should be taken into account, but, just as in the evaluation of the standard errors for B , and for the same reason, it is ignored.

The symmetry restriction (28) can be conveniently written:

$$L(I-K)G[(I-\bar{w}\beta^{1'})\otimes I]\text{vec}D' + (\bar{w}\otimes I)\beta^0 = 0, \quad (31)$$

where G is a 49×64 matrix of ones and zeros which delivers a vector containing the top left 7×7 elements of an 8×8 matrix in vec form, K is the 49×49 commutation matrix, which converts $\text{vec}A$ into $\text{vec}A'$, and L is a 21×49 matrix which selects from $\text{vec}A$ the elements of the lower left triangle in the original matrix. This expression is simpler than it looks. The restriction $L(I-K)A = 0$ for some matrix A requires that the lower left triangle of the transpose should be identical to the lower left triangle of the original, *i.e.* that the original should be symmetric. It is written in this form rather in the simpler $A = KA$ because this last imposes redundant restrictions, including for example that the diagonal elements equal themselves, which would result in singularities when calculating the variances and covariances of the restrictions. The G matrix is necessary because the restrictions of adding-up, homogeneity and symmetry are not independent. The D matrix satisfies adding-up and homogeneity by construction, so that symmetry need only be applied to the top 7×7 submatrix of the left hand side of (28).

Equation (30) is substituted into equation (31) to generate a “standard” set of linear restrictions

$$R \text{vec}B' = r. \quad (32)$$

A restricted estimate of B is then calculated according to

$$\text{vec}(B^{*'}) = \text{vec}(\hat{B}') + VR'(RVR')^{-1}(r - R\text{vec}(\hat{B}')) \quad (33)$$

with variance covariance matrix $(RVR')^{-1}$. A Wald test for the symmetry restriction is available from

$$W = (r - R\text{vec}(\hat{B}'))'(RVR')^{-1}(r - R\text{vec}(\hat{B}')). \quad (34)$$

The results are shown in Table 6 for both sectors. The Wald tests for the symmetry restriction take values 58.04 in the rural sector and 80.03 in the urban sector, both of which would be distributed as χ^2_{21} if the null hypothesis were true. Although these tests indicate rejection at any conventional significance level, they are based on large numbers of observations, 620 rural villages, and 638 urban clusters. They are not

Table 6: Estimates of Own and Cross-Price Elasticities: Pakistan 1984-85
(symmetry contained estimates for the full system)

	Wheat	Rice	Dairy Produce	Meat	Oils & Fats	Sugar	Other Food	Non- Food
RURAL								
Wheat	-0.51 (0.08)	0.07 (0.03)	0.08 (0.02)	-0.04 (0.04)	0.06 (0.04)	0.04 (0.04)	0.01 (0.05)	-0.08 (0.13)
Rice	0.30 (0.15)	-1.53 (0.12)	-0.04 (0.07)	0.30 (0.10)	0.38 (0.09)	0.34 (0.18)	-0.34 (0.13)	-0.10 (0.28)
Dairy	0.00 (0.02)	-0.01 (0.02)	-0.89 (0.03)	-0.02 (0.02)	0.02 (0.01)	0.01 (0.01)	0.01 (0.02)	-0.17 (0.05)
Meat	-0.22 (0.14)	0.21 (0.07)	-0.08 (0.05)	-0.61 (0.13)	0.17 (0.10)	0.13 (0.09)	0.20 (0.13)	-0.91 (0.29)
Oils & Fats	0.19 (0.13)	0.26 (0.06)	0.16 (0.04)	0.19 (0.09)	-1.59 (0.40)	-0.29 (0.25)	0.02 (0.13)	0.65 (0.51)
Sugar	0.12 (0.17)	0.31 (0.07)	0.10 (0.05)	0.18 (0.11)	-0.44 (0.35)	0.03 (0.40)	0.10 (0.17)	-1.27 (0.62)
Other Food	-0.01 (0.06)	-0.07 (0.03)	0.07 (0.02)	0.08 (0.04)	-0.01 (0.04)	0.03 (0.04)	-0.40 (0.08)	-0.28 (0.12)
Non- Food	-0.10 (0.03)	-0.01 (0.02)	-0.05 (0.01)	-0.07 (0.02)	-0.09 (0.04)	-0.09 (0.04)	-0.13 (0.03)	-0.68 (0.10)
URBAN								
Wheat	-0.69 (0.12)	0.22 (0.04)	-0.03 (0.04)	-0.10 (0.06)	0.05 (0.05)	-0.02 (0.05)	0.13 (0.07)	0.04 (0.18)
Rice	0.97 (0.17)	-1.85 (0.10)	-0.50 (0.08)	0.28 (0.12)	0.22 (0.09)	0.19 (0.09)	-0.07 (0.14)	0.04 (0.32)
Dairy	-0.08 (0.04)	-0.11 (0.02)	-1.11 (0.04)	0.10 (0.03)	-0.00 (0.02)	-0.02 (0.02)	-0.02 (0.03)	0.24 (0.08)
Meat	-0.22 (0.10)	0.10 (0.04)	0.18 (0.05)	-0.32 (0.10)	-0.07 (0.06)	-0.09 (0.06)	-0.03 (0.09)	-0.49 (0.22)
Oils & Fats	0.10 (0.10)	0.11 (0.04)	0.05 (0.04)	-0.05 (0.07)	-1.52 (0.51)	-0.07 (0.24)	-0.02 (0.08)	0.92 (0.52)
Sugar	-0.07 (0.14)	0.13 (0.06)	-0.03 (0.05)	-0.15 (0.10)	-0.11 (0.34)	-0.72 (0.39)	-0.41 (0.12)	0.71 (0.53)
Other Food	0.10 (0.06)	-0.01 (0.03)	0.01 (0.03)	0.00 (0.05)	-0.02 (0.04)	-0.12 (0.04)	-0.34 (0.08)	-0.21 (0.13)
Non- Food	-0.06 (0.03)	-0.00 (0.01)	0.02 (0.01)	-0.05 (0.02)	0.04 (0.04)	0.02 (0.03)	-0.09 (0.02)	-0.97 (0.08)

Note: The rows refer to the good being affected, and the columns to the price being changed. Hence, row 2, column 7 for the rural sector estimates that a 1% increase in the price of other food will increase the quantity of rice purchased by 0.38%. The figures in parentheses are standard errors.

impressive when compared with critical values that take account of sample size; the Schwartz test has critical values of 135.0 and 135.6 respectively, which are comfortably larger than the test statistics here. More convincing perhaps is the extent to which the results in Table 6 overcome the problems in Table 5. For wheat, rice, dairy produce, and to a large extent meat, the results are very close to those in the previous table. However, for oils and fats and for sugar, where the previous estimates were quite cross-price elasticities in Table 5 in these two imprecise, the estimates change a great deal, and are now much more precisely estimated. Further, the several large columns no longer appear in the symmetry-constrained estimates. This is exactly what we had hoped that the imposition of symmetry would accomplish, and the results are a good indication of the usefulness of even very simple theory in supplementing relatively uninformative data.

The own-price elasticities are very similar in the two tables, and those for sugar and oils are still imprecisely estimated; symmetry cannot help here. In the rural sector, the own-price elasticity of sugar is *positive*, but the standard error is large. The estimates in the non-food row should be treated rather more seriously than those in the non-food column. Since the budget shares of the eight goods add to unity, we have data on the share of non-food, and the estimates of the effects of food prices on non-foods have essentially the same status as the estimates of the effects of food prices on foods. However, there are no unit value data for non-food, so that the estimates of the effects of the non-food price on foods are determined by the theoretical restrictions, rather than by the data. The standard errors in the last column tend to reflect this imprecision, and these elasticities should not be treated too seriously. There are several important cross-price elasticities that survive the imposition of symmetry. There are well-determined substitution effects between rice and wheat, between rice and meat, rice and sugar, and between rice and oils, and there is a significant negative effect, only some of which is an income effect, between wheat and meat. These cross-price effects appear in both urban and rural samples, and we feel that a good deal of credence should be given to them.

The Analysis of Tax and Price Reform

This section looks at the implications of the estimates for the analysis of a number of possible price reforms in Pakistan. The general procedure is the standard one, of looking at the distributional consequences of price changes, together with their effects on economic efficiency, and trying to identify directions of price reform that can improve efficiency without deleterious distributional consequences, or which can meet stated distributional aims of protecting the poor, but without avoidable efficiency costs. For reasons that will become apparent, the policy options that we examine are somewhat artificial ones, but the effects that they illustrate would be important in any more thoroughgoing analysis. We are also concerned not to use the results of the previous sections mechanically, but instead to identify those implications for policy that are credible and robust, and downweight those that may be due to sampling variability, or are otherwise doubtful.

The theory of price and tax reform in developing countries is well-developed in the literature, see for example the introductory chapters in Newbery and Stern (1987), and the survey paper by Dreze and Stern (1987). These principles have been applied to tax reform in Pakistan by Ahmad and Stern (1990, 1991), and we shall follow the general framework of that article, if not the details. In the standard case, where there are no quality effects, and everyone faces the same price, the analysis runs as follows. Suppose that there is a proposal to increase the consumer tax on good i . The (local) consequences of this change can be assessed by looking at the derivatives of consumer welfare with respect to the change, together with the effects on government revenue. The compensation required by agent h for tax-induced price change Δp_i is $q_i^h \Delta p_i$. But we are typically not indifferent as between different gainers and losers, particularly in LDCs where there are very limited instruments for redistribution, so these individual compensations must be weighted according to weights θ^h that are proportional to the social marginal values of income to each agent. The social cost of the tax increase is therefore given by the derivative

$$\frac{\partial W}{\partial t_i} = \sum_h \theta^h q_i^h \quad (35)$$

The tax change will also have an effect on government revenue, directly through the additional taxes raised on the good, and indirectly through the own- and cross-price

substitution effects that are induced by the price increase. If R is total government revenue, then the revenue derivative is given by

$$\frac{\partial R}{\partial t_i} = \sum_h q_i^h + \sum_h \sum_k t_k \frac{\partial q_k^h}{\partial t_i} \quad (36)$$

The ratio of (35) to (36), typically denoted λ_i , measures the costliness of raising additional revenue through good i . If λ_i is large, additional units of revenue are obtained at high social cost, as would happen, for example, if the good is heavily taxed, is highly price elastic, and most heavily consumed by the poor. By contrast, goods with low λ_i values are attractive candidates for raising revenue.

As written, equations (35) and (36) take no account of distortions elsewhere in the economy, nor of the resource allocation effects of tax changes that will typically result if prices do not reflect opportunity costs. To take account of these effects, write the consumer price p_i as $s_i + t_i$, where, for the moment, s_i is interpreted as a base price that is unaffected by the tax change. Since the tax change does not alter the total amount spent by consumers, (36) can be rewritten as

$$\frac{\partial R}{\partial t_i} = \sum_h \sum_k s_k \frac{\partial q_k^h}{\partial t_i} \quad (37)$$

so that the tax cost ratios λ_i become

$$\lambda_i = \frac{\sum_h \theta^h q_i^h}{-\sum_h \sum_k s_k \frac{\partial q_k^h}{\partial t_i}} \quad (38)$$

Under appropriate (but non-trivial) assumptions, Dreze and Stern (1987) show that (38) has an attractive alternative interpretation that allows a substantial generalization over the pure revenue focus of (35) and (36). In particular, if *actual* taxes are ignored, the vector s is taken as a vector of *shadow* prices, *i.e.* prices that reflect the relative social opportunity costs of the goods, so that the vector t is a vector of shadow tax rates, then (38) captures all of the effects of the tax changes through the general equilibrium system, and is therefore of quite general applicability. Indeed, the denominator of (38) represents minus the resource costs, *i.e.* the resource benefits of the tax increase, while

its numerator is, as before, the welfare cost, so that the λ_i 's are simply cost-benefit ratios at the "right," *i.e.* shadow prices.

The implementation of (38) requires the consumption of the various commodities, which we have from the survey, the price derivatives, which have been estimated in Section 1, social income weights, which are discussed below, and shadow prices. Ahmad, Coady and Stern (1988) have estimated a complete set of shadow prices for Pakistan based on an 86 x 86 input output table for 1976, and Ahmad and Stern (1990, 1991) use these shadow prices in their tax reform experiments. However, these prices embody a very large number of assumptions, about shadow prices for land, labor, and assets, as well as, most crucially, judgments about which goods are tradeable (further divided into importables and exportables) and which goods are non-traded. These assumptions all have a role in determining the shadow prices, and while the procedure itself is unimpeachable, the use of the prices for the current exercise makes it very difficult to isolate the role played in the reform proposals by the different elements in the story. Instead, we shall work with an illustrative set of shadow prices for our eight goods. These prices are illustrative in the sense that they do not claim to take into account all the ingredients that should ideally be included in shadow prices. However, they *are* based on the current Pakistan situation, and help us see how important features of that situation enter into the analysis of price reform.

The crucial commodities are wheat, rice and sugar. Rice and sugar sell at internal prices that are substantially lower than their world prices at the official exchange rate. According to the *Pakistan Economic Survey*, Government of Pakistan (annual), the border price for both rice and wheat was approximately 40% above the world prices in the mid-80's. Both are tradeable goods, and for both we work with accounting ratios (shadow prices divided by consumer prices) of 1.4. In the case of sugar, there is a heavily protected domestic sugar refining industry, and the border price of refined sugar is some 60% of the domestic price. Oils and fats are also protected by a complex system of taxes and subsidies with the result that the border price is perhaps 95% of the domestic price. For the other four goods in the system, we shall work with accounting ratios of unity. Although there are tariffs and taxes on many non-food items, which would call for accounting ratios less than unity, many other items are non-traded, so that their shadow prices depend on the shadow price or conversion factor for labor, which, given the pricing of wheat and rice, is likely greater than unity. Our experiments are therefore conducted under the supposition that the ratio of shadow prices to consumer

prices for the eight goods (wheat, rice, dairy, meat, oils and fats, sugar, other food, non-food) are given by the vector (1.4, 1.4, 1.0, 1.0, 0.95, 0.6, 1.0, 1.0).

It is important to be precise about the policy experiments that are being considered, and about those that are not. The instruments that are being considered here are instruments that increase or decrease the *consumer* price of the good, as would be the case, for example, if there were a value added tax, or if the good were being imported over a tariff. Such instruments may be available in Pakistan for sugar or for oils. However, to the extent that the rice and wheat subsidies are maintained by export taxes, the experiments described here *do not* correspond to what would happen if the export tax were changed. An increase in an export tax *increases* consumer welfare at the same time as it *increases* government revenue, so that if only these two effects were considered, as is the case in (35) and (36), it would pay to increase the tax indefinitely. Of course, export taxes are limited by the supply responses of farmers, as well as by international demand when the country has some monopoly power, as is the case for Pakistan with basmati. An export tax is an instrument that alters not only consumer prices but also producer prices, and it cannot be analyzed without looking at the supply responses. The calculations here, which look only at the demand side, correspond to a different hypothetical experiment, in which producer and consumer prices are separated. This would suppose that the government procures the total output of rice and wheat at one set of prices, that it sells the commodities either to foreigners at world prices, or to domestic consumers at a third set of prices. It is the consequences of changing these domestic prices that can be examined using the apparatus described above. Of course, these are very artificial experiments, which would not be feasible in practice. Farmers who grow wheat cannot be charged one price for consumption and paid another for procurement. However, more realistic policies will have the same effects as those described here, plus additional effects that work through the supply side. What we have to say here is only a part of the story, but it is a part that has to be understood if good policy decisions are to be made.

Before calculating the results, we need to adapt the basic model of tax reform to a world where people pay different prices, and where quality as well as quantity is an object of choice. The simplest assumption, and that adopted here, is that taxes on each good are *ad valorem*, so that the tax paid is a constant proportion of price, so that the tax per kilo is higher for higher priced varieties, as well as at locations where rice is more expensive. It is easy to think of cases where this is not true, where taxes are fixed

at rates per quintal, and where transport margins are the same irrespective of the total value of the load. Nevertheless, the assumption yields substantial simplifications, as we shall see. If the tax rate on good i is τ_i , then the tax paid by consumer h on good i is $\tau_i v_i^h q_i^h$ where $v_i^h = v_i^h / (1 + \tau_i)$ is the ex-tax unit value. Hence, the compensation payable for an increase $\Delta \tau_i$ is $v_i^h q_i^h \Delta \tau_i$ so that the numerator of the λ_i ratio (35) is replaced by

$$\frac{\partial W}{\partial \tau_i} = \sum_h \theta^h v_i^h q_i^h = \sum_h (x^h/n^h)^{-\epsilon} x^h w_i^h \quad (39)$$

where we have adopted the standard practice that the social weight given to additional income for h is proportional to the level of *per capita* household expenditure, x/n , raised to the power of $-\epsilon$, for some $\epsilon > 0$. A value of ϵ of unity implies that additional income is twice as valuable to someone with half the income, with higher values implying a greater focus on the poor, and lower values less. Social welfare accounting is done for individuals, not households, but (39) is nevertheless correct under the inevitable assumption that household consumption levels are equally shared among household members.

To derive the revenue effects of changes in τ_i , it helps to decompose unit value into the product of p_i and quality, μ_i . Revenue raised from consumer h , R^h , is then

$$R^h = \sum_k \tau_k \mu_k^h p_k^h q_k^h \quad (40)$$

where p_k^h is the ex-tax price faced by the consumer. We can then differentiate with respect to τ_i , remembering to take into account the shading of qualities in response to price, and using the facts that

$$\frac{\partial p_i^h}{\partial \tau_i} = \frac{p_i^h}{1 + \tau_i} \quad \frac{\partial \ln \mu_i}{\partial \ln p_j} = \psi_{ij} - \delta_{ij} \quad (41)$$

it is possible to show that

$$\frac{\partial R^h}{\partial \tau_i} = \frac{1}{1 + \tau_i} \sum_k \frac{\tau_k}{1 + \tau_k} x^h (\theta_{ki} - \delta_{ki} w_i^h) \quad (42)$$

which can be evaluated given actual or shadow tax rates, the data from the survey, and the estimates of the parameters. One final formula will be useful. Define the aggregate budget shares, the shares of each expenditure in aggregate consumers' expenditure, by

$$\tilde{w}_i = \frac{\sum_h x^h w_i^h}{\sum_h x^h} \quad (43)$$

and the “socially representative” budget shares by

$$w_i^\epsilon = \frac{\sum_h (x^h/n^h)^{-\epsilon} x^h w_i^h}{\sum_h x^h} \quad (44)$$

The cost-benefit ratios λ_i are the ratios of (39) to (42) and can be written

$$\lambda_i = \frac{w_i^\epsilon / \tilde{w}_i}{1 + \frac{\tau_i}{1+\tau_i} \left[\frac{\theta_{ii}}{\tilde{w}_i} - 1 \right] + \sum_{k \neq i} \frac{\tau_k}{1+\tau_k} \frac{\theta_{ki}}{\tilde{w}_i}} \quad (45)$$

The numerator of (45) is a pure distributional measure for good i ; it can be interpreted as the relative shares of the market representative individual (the representative agent) and the socially representative individual, whose income is lower the higher is the inequality aversion parameter ϵ . This measure is modified by the action of the terms in the denominator. The first of these (apart from 1) is the tax factor multiplied by the elasticity of expenditure on good i with respect to its price, quality and quantity effects taken together, see equation (15) above. This term measures the own-price distortionary effect of the tax. If it is large and negative, as would be the case for a heavily taxed price elastic good, the term will contribute to a large λ_i -ratio and would indicate the costliness of raising further revenue by that route. The last term is the sum of the tax factors multiplied by the cross-price elasticities, and captures the effects on other goods of the change in the tax on good i , again with quantity and quality effects included. From a theoretical point of view, this decomposition is trivial, but when we look at the results, it is useful to separate the own- and cross-price effects since the former are likely to be more reliably measured than the latter.

Table 7 shows the efficiency effects of raising taxes on each of the goods, distinguishing between the terms in the denominator of (45); these are calculated for the rural sample only. The urban results are similar, and one sector is sufficient to illustrate how the procedure works. The first column shows the tax factors $\tau_i/(1+\tau_i)$ calculated from the accounting ratios discussed above; these are the shadow taxes calculated from comparison of the world and domestic prices. Wheat and rice carry shadow subsidies, oils and sugar, shadow taxes. The second column shows the own-price elasticities for

Table 7: Efficiency Aspects of Price Reform: Rural Sector

	Tax Factor	e_{ii}	Own	Cross	Total
Wheat	-0.67	-0.61	0.40 (0.08)	-0.04 (0.03)	1.36 (0.08)
Rice	-0.67	-1.80	1.20 (0.09)	-0.17 (0.13)	2.02 (0.14)
Dairy	0.00	-1.01	0.0 -	-0.05 (0.02)	0.95 (0.02)
Meat	0.00	-0.70	0.0 -	-0.00 (0.12)	1.00 (0.12)
Oils	0.05	-1.67	-0.08 (0.02)	-0.49 (0.16)	0.43 (0.16)
Sugar	0.29	0.05	0.01 (0.12)	-0.41 (0.16)	0.61 (0.19)
Other Food	0.00	-0.41	0.0 -	0.06 (0.06)	1.06 (0.06)
Non-Food	0.00	-0.75	0.0 -	0.00 (0.03)	1.00 (0.03)

Notes: The tax factor is the share of the shadow tax in the tax-inclusive price, e_{ii} is the own-price elasticity of quality times quantity, own and cross are the second and third terms in the denominator of (45), and total is the denominator itself. Figures in parentheses are standard errors.

quantity and quality taken together. Because the quality effects are not large, these elasticities are close to those listed in Table 6, but they are conceptually distinct. Taxes are *ad valorem* so that quality shading in response to tax increases depresses revenue just as do decreases in quantity. The third column shows the own-price distortions and corresponds to the middle term in the denominator of (45). Goods that bear no shadow taxes do not appear in this column since there are no distortionary effects of small tax increases at zero tax rates. The large effects are on wheat and rice, with the latter much larger because its own-price elasticity is much larger. The standard errors for both these terms are small relative to their estimated magnitudes. By themselves, these own price terms will generate small cost-benefit ratios, particularly for rice. Subsidies are being paid, they are distortionary, particularly for rice, and it is desirable to reduce them, at least from this limited point of view.

The cross terms in the next column are typically smaller and, as expected, have larger standard errors. The large terms are for rice, for oils, and for sugar; all are negative and all act so as to *decrease* the attractiveness of raising those prices. By

checking back through the full matrix of price elasticities, it is possible to track down the specific terms that are responsible for these total cross-price effects. For rice, the interaction is with wheat. As we have seen, an increase in the price of rice is attractive on own-price efficiency grounds, but it will also switch demand into wheat, which itself bears a subsidy, so that the attractions of raising the price are (partially) reduced. The oils and sugar terms are very similar; both are substitutes for wheat and rice, so that increasing either price generates demands for the rice and wheat subsidies which are socially costly. In addition, increases in the oils price generate a fall in the demand for sugar, which bears a shadow tax, so that its opportunity cost is well below its price, an effect which again makes it less attractive to raise the prices of oils. The final column in the Table presents the sum of both effects, plus one according to (45), together with the estimated standard error. According to these, and recall that nothing yet has been said about distributional issues, rice is the commodity whose price should be raised, and sugar and oils the commodities whose prices should be lowered. The wheat subsidy too is distortionary, and there is an efficiency case for raising the price. The standard errors on the oils and sugar terms are relatively large, reflecting the fact that they come from cross-price effects, but they are not large enough to threaten the main thrust of these conclusions.

Equity effects are incorporated in Table 8 for a range of reasonable values of the distributional parameter ϵ . The first panel corresponds to $\epsilon=0$, so that there are no distributional concerns, and the cost-benefit ratios are simply the reciprocals of the last column in Table 7. Rice is a good candidate for a price increase, oils and fats for a price decrease. As we move through the table to the right, the distributional effects modify this picture to some extent. The "equity" columns in the table, which are the relative budget shares of an increasingly poor individual relative to the market representative individual, move away from luxuries and towards necessities as the distributional parameter increases. Wheat, the basic staple, attracts a larger weight as we move down the income distribution, as do oils and fats. Other food, which contains pulses, has large weights when ϵ takes the extreme value of 2. Wheat is a target for price increases on efficiency grounds, although not as much so as rice, but the equity effects tend to make it less attractive, and its λ_i -value becomes relatively large as we move to the right in the table. However, for oils and fats, and to a lesser extent for sugar, which were the main candidates for price decreases, the equity effects only strengthen the conclusions. Oils and fats in particular figure relatively heavily in the budgets of low income consumers,

Table 8: Equity and Cost-Benefit Ratios for Tax Increases

	$\epsilon=0$		$\epsilon=0.5$		$\epsilon=1.0$		$\epsilon=2.0$	
	Equity	λ	Equity	λ	Equity	λ	Equity	λ
Wheat	1.00	0.73	1.11	0.81	1.29	0.95	2.00	1.47
Rice	1.00	0.49	1.06	0.52	1.18	0.58	1.65	0.82
Dairy	1.00	1.05	1.03	1.08	1.12	1.18	1.51	1.62
Meat	1.00	1.00	1.00	1.00	1.05	1.06	1.37	1.38
Oils	1.00	2.33	1.09	2.54	1.26	2.93	1.90	4.43
Sugar	1.00	1.64	1.03	1.68	1.11	1.83	1.51	2.48
Other Food	1.00	0.94	1.06	1.00	1.18	1.11	1.72	1.63
Non-Food	1.00	1.00	0.98	0.98	1.04	1.04	1.43	1.43

and reducing their price is desirable for both equity and efficiency reasons. Rice remains the best candidate for consumer tax increases; it is not consumed by the very poor, and its subsidy is the most distortionary.

In conclusion, let us emphasize once again that these policy implications do not apply directly to the tax instruments currently in use in Pakistan. Our findings suggest that raising revenue from an increase in the price of rice to consumers would be desirable on grounds of efficiency and equity. But the main instrument for controlling the price of rice is the export tax, a decrease in which would certainly increase the consumer price of rice, but it would also *decrease* not *increase* government revenue, and it would change supplies of rice, effects which we are in no position to consider here. Nevertheless, the fact that rice and wheat are substitutable for one another in consumption, and that there are further substitution effects between rice, wheat, sugar and edible oils, would have to be taken into account in a reform of any export taxes, just as they are taken into account here. Indeed, perhaps the most useful lesson of the analysis in this paper is how important it is to *measure* the way in which price changes affect the pattern of demand. The Pakistani substitution patterns between rice, wheat, sugar, and edible oils are not consistent with additive preferences and so cannot be accommodated within a model like the linear expenditure system. In countries where some prices are far from opportunity cost, cross-price effects of tax changes will often be more important than the effects on the good itself, so that these effects must be measured in a flexible way. Nor could additive preferences accommodate the pattern of total expenditure and own-price elasticities that characterize demand patterns in Pakistan. Oils and fats are a

necessity, but have a high own-price elasticity, so that the ratio of the own-price to income elasticity is many times greater than for a rice, which has high own-price and income elasticities. Additive preferences require that this ratio be (approximately) uniform over goods. Yet it is this ratio that is the principal determinant of how the balance between equity and efficiency ought to be struck.

References

- Ahmad, S.E., and N.H. Stern, (1991), *The theory and practice of tax reform for developing countries*, Cambridge. Cambridge University Press.
- Ahmad, S.E., and N.H. Stern, (1990), "Tax reform and shadow prices for Pakistan," *Oxford Economic Papers*, **42**, 135-159.
- Ahmad, S.E., Coady, D., and N.H. Stern, (1988), "A complete set of shadow prices for Pakistan: illustrations for 1975-6," *Pakistan Development Review*, **27**, 7-43.
- Ahmad, S.E., Ludlow, S., and N.H. Stern, (1988), "Demand responses in Pakistan: a modification of the linear expenditure system for 1976," *Pakistan Development Review*, **27**, 293-308.
- Alderman, H., (1988), "Estimates of consumer price response in Pakistan using market prices as data," *Pakistan Development Review*, **27**, 89-107.
- Deaton, A.S., (1981), "Optimal taxation and the structure of preferences," *Econometrica*, Vol. 49, pp. 1245-60.
- Deaton, A.S., (1987a), "Econometric issues for tax design in developing countries," Chapter 4 in Newbery and Stern (1987).
- Deaton, A.S., (1987b), "Estimation of own- and cross-price elasticities from survey data," *Journal of Econometrics*, **36**, 7-30.
- Deaton, A.S., (1988), "Quantity, quality, and spatial variation of price," *American Economic Review*, **78**, 418-430.
- Deaton, A.S., (1990), "Price elasticities from survey data: extensions and Indonesian results," *Journal of Econometrics*, **44**, 281-309.
- Deaton, A.S., and J. Muellbauer, (1980a), "An almost ideal demand system," *American Economic Review*, **70**, 312-326.
- Deaton, A.S., and J. Muellbauer, (1980b), *Economics and consumer behavior*, New York. Cambridge University Press.
- Dreze, J.P., and N.H. Stern, (1987), "The theory of cost-benefit analysis," in A. Auerbach and M. Feldstein, eds., *Handbook of Public Economics*, Amsterdam. North-Holland.
- Government of Pakistan, (annual), *Pakistan Economic Survey*, Islamabad. Ministry of Finance.
- Government of Pakistan, (monthly), *Monthly Statistical Bulletin*, Islamabad. Federal Bureau of Statistics.

- Newbery, D.M.G., and N.H. Stern, (1987), *The theory of taxation for developing countries*, Oxford. Oxford University Press for The World Bank.
- Schwartz, G., (1978), "Estimating the dimension of a model," *Annals of Statistics*, **6**, 461-464.

Appendix: Alternative Methods of Estimation

The methodology used in this paper, and described in the Section 2, is designed to extract price and expenditure elasticities from survey data in the form in which they are typically available. It allows for the possibilities that there are village level fixed effects in demand patterns, effects that may be correlated with average village incomes, that unit values are not prices, and that there may be non-independent measurement error in the reported unit value and quantity data. Handling these various issues imposes costs in terms of relative complexity, and the question arises as to whether there may be short-cuts that ignore some of the problems that are less serious in practice. In fact, the calculations in section 1 are not very difficult to carry out, and it is only the calculation of the standard errors that requires careful and customized programming. As always, the heaviest work is in the preparation and cleaning of the basic data, not in the calculation of the estimates.

This section has two purposes. The first is to provide some alternative estimates that are calculated using standard regression methods, and examines how much difference there is between these and the methodology of this paper. The second purpose is to try out alternative data, not methodologies. We follow Alderman (1988) and reestimate the model using not the unit values from the survey, but the urban regional price data that is independently collected by the Federal Bureau of Statistics and published in the *Monthly Statistical Bulletin*. Households are matched to the survey point that is geographically closest to them, and assigned the price for that area in the quarter during which they were interviewed. The model is essentially the same as that used in the paper, but the prices are genuine prices, not unit values. The disadvantage of the method is that there are relatively few collection points, and they are all urban, so that they may provide only very poor indications of the prices facing the households in the survey.

Note that the two sets of alternative estimates do not have the same status. For the models using the alternative price data, we would like to obtain the same answers as before, since there is nothing incorrect in using the alternative prices, except for the possible effects of measurement error. But if the issues raised in Section 2 are real, the short-cut methodologies are essentially incorrect and they should give different answers. Of course, measurement error, village taste differences, or quality effects may not be very important in a particular application, and the estimates may turn out to be similar.

This is no guarantee that in other applications the same would be true, although it is to some extent possible to highlight the sources of difference, and to be aware of situations in which short-cuts would be dangerous. Even so, the alternative estimates given here are only indicative of the possibilities. A full understanding of the alternative estimators would require more analysis and Monte Carlo experimentation than can be included in this paper, although see Deaton (1990) for some limited evidence.

There are eight different sets of estimates in the columns of Table A1. In order to keep the comparisons manageable, we confine ourselves to the rural sector, and present only the results for the total expenditure elasticity and the own-price elasticity, the former in the top of the table, and the latter in the bottom. The first column repeats the results of Section 2, and is taken from Table 5. The last two columns, to the right of the double line, are the results from using the alternative price data, and will be discussed below.

Columns (2) and (3) come from the following specification:

$$\ln q_i = \alpha_{0i} + \alpha_{1i} \ln x + \alpha_{2i} \ln n + \sum \eta_{ij} \ln v_j + \gamma_i \cdot z + \text{province and seasonal dummies} + u_i \quad (\text{A1})$$

where q_i is the quantity purchased of good i , z is the same vector of household demographics used in the first-stage regression as in Section 2, and the unit values are included on the right hand side as if they were prices. Equation (A1) is attractive in its simplicity; the parameters are elasticities, and a double logarithmic regression is perhaps the simplest way to estimate them. However, there are a number of immediate problems of application. The most immediate is that households who do not purchase the good must be excluded from the regression; the logarithm of zero does not exist. For the analysis of tax reform, the exclusion of non-purchasing households is undesirable since we are interested in the effects of price change on total or average demand, and we are not concerned whether that change takes place at the intensive or extensive margins. From an econometric point of view, the exclusion of zero purchases also reduces the sample size, particularly if all the unit values are included on the right hand side, since in that case only those households can be included who are recorded as purchasing *all* of the goods. This is a very severe restriction. Table 2 shows the fractions of households buying *each* good, but a much smaller fraction buy *all* of them and only

1,666 of the 9,119 rural households do so. The regressions in column (2) report results from this very severely selected sample, while in column (3), the selection problem is lessened by including only the own-price on the right hand side, effectively assuming *a priori* that the cross-price terms are negligible or at least, uncorrelated with the other variables.

Columns (4) and (5) are obtained from regressions identical to (26) but with the budget share as the dependent variable in place of the logarithm of the quantity. The functional form is thus essentially the same as that used in Section 2, but the regressions are run directly without allowance for village fixed effects, measurement error, or any distinction between unit value and price. Although all zero shares can now be included on the left hand side, the selection problem recurs through the unit values. Once again, column (4) reports estimates on the restricted sample of 1,666 households with all the unit values included, and column (5) those estimates with only the own unit value, so that the selection fractions are those in Table 2. Finally, column (6) uses the same model as columns (4) and (5), but all the data are averaged over each cluster prior to the estimation. Included in the regressions are all those clusters where each good is purchased by at least one household; of the 757 rural clusters, 620 satisfy this condition. This is the same selection that is enforced upon us in Section 2.

There is another possible technique which is not examined here, but which is sometimes used and with which we experimented early in the research. One of the problems of estimating a model on the household data comes from the potential spurious correlation between measured quantities and measured unit values, a correlation that can be generated by measurement error in either or both. One possibility is to replace the *individual* unit value by the average of all the unit values in the village; since prices are supposed to be the same for all households in the village, this is clearly legitimate. However, the procedure does not resolve the spurious correlation problem, since the village mean is still correlated with the individual unit values used to compute it. It can therefore be modified one step further by using for each household the village mean unit value computed using all unit values in the village *except* that of the household itself. While this removes any spurious correlation between quantities and unit values, or between shares and unit values, the “leave-out” mean, like the original mean, is still an error-ridden estimate of the true and unobservable average village unit value. Attenuation bias is therefore not removed. In the models used in Section 2, where the share is the dependent variable, the spurious correlation problem is typically less serious

Table A1: Alternative Estimates of Total Expenditure and Price Elasticities (Rural Pakistan 1984-85)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TOTAL EXPENDITURE ELASTICITIES								
Wheat	0.37	0.31	0.28	0.48	0.35	0.34	0.44	0.38
Rice	0.68	0.63	0.66	0.70	0.61	1.03	0.85	0.91
Dairy produce	1.05	0.68	0.85	0.81	0.92	1.03	1.12	0.99
Meat	1.10	0.91	0.98	1.04	1.05	1.35	1.32	1.48
Oils & Fats	0.42	0.44	0.37	0.44	0.42	0.38	0.43	0.42
Sugar	0.86	0.72	0.79	0.76	0.82	1.22	0.95	1.12
Other Food	0.59	0.57	0.64	0.66	0.70	0.81	0.74	0.89
OWN-PRICE ELASTICITIES								
Wheat	-0.51	-0.75	-0.93	-0.76	-0.89	-0.72	-1.16	-1.18
Rice	-1.59	-1.07	-1.09	-1.37	-1.41	-1.94	-2.18	-2.25
Dairy produce	-0.87	-0.92	-0.94	-0.91	-0.91	-0.98	-1.02	-1.00
Meat	-0.57	-0.45	-0.35	-0.31	-0.27	-0.71	-0.52	-0.60
Oils & Fats	-2.04	-0.00	-0.89	+0.04	-0.65	-1.62	+2.88	+3.57
Sugar	-0.07	-0.59	-0.78	-0.31	-0.62	-0.22	+0.76	+0.82
Other Food	-0.35	-0.32	-0.54	-0.43	-0.64	-0.42	-0.91	-0.95

Notes: Column (1) was calculated using the methodology of Section 1, and the estimates are those in Table V. Column (2) comes from a double logarithmic regression of $\ln q$ on $\ln x$, the logarithms of the unit values of all the goods, and the other variables. Column (3) is the same as (2), but only the "own" unit value is included in the regression. Column (4) shows elasticities calculated from the results of regressing the *shares* on $\ln x$, all unit values, and the other variables, while (5) is the same as (4) but omitting all but the own unit value. Column (5) is the same as regression (3), but uses cluster/village means for all variables. Columns (6) and (7) correspond to columns (4) and (5) but use the urban prices from the *Monthly Statistical Bulletin* in place of the unit values from the survey.

than the attenuation problem. Computation using the "leave-out" method is almost as complicated as the full-method, and so it has little to commend it, and we do not examine it further.

Columns (7) and (8) in the table correspond to the same regressions as columns (4) and (6), that is they use (26) with shares as dependent variables, but the prices from the urban centers are used in place of the unit values. Column (7) uses the individual data, and column (8) the data averaged by cluster. There are no missing values for these

regressions, so that column (7) uses all 9,119 rural households, and column (8) all 757 rural clusters.

There are a variety of econometric issues that can account for the differences in the columns of Table A1. Columns (2) and (3) use a double logarithmic functional form that is different from the share-log form of the other columns. Because of the missing value (zero) problem, these regressions are also estimated on different samples from the others, and the samples vary from commodity to commodity. Column (1) obtains its total expenditure elasticities from within-village estimation, and so allows for village taste differences that can be correlated with village income or village demographics. If these effects are important, all the other estimates will be biased, since they pool within and between village information. Columns (2) through (6) treat unit values as if they were prices, and make no allowance for the effects of quality choice on unit values. All columns except the first assume that the total expenditure elasticity of *expenditure* on each good is identical to the total expenditure elasticity of *quantity*. Columns (1), (2) and (3) measure the latter, and columns (4) through (8) the former. The last seven columns also make no allowance for measurement error. In the last two columns, which use the urban prices, the measurement error has the standard effects, of attenuating the coefficient on the price, which in this context implies that the own-price elasticities are biased towards minus one. For columns (2) through (6), the effects of the measurement error are more complicated. Given the way that unit values are measured, it is likely that the reported logarithms of each unit value, which is the logarithm of expenditure minus the logarithm of quantity, will be spuriously negatively correlated with the reported logarithm of quantity. For columns (2) and (3), this will generate an attenuation bias towards zero on the price term, together with a bias towards minus one from the spurious correlation, and the net effect cannot be signed theoretically. For columns (3) through (6), it is more likely that the measurement errors of the dependent variables, the shares, and the log unit values will be uncorrelated, so that the attenuation bias will dominate, and the own-price elasticities will be biased towards minus one. Of course, all of these results suppose that the specification in Section 2 is correct, which may or may not be the case. Even so, one cannot reverse the supposition, and treat any of the other models as the correct specification. Although the assumptions behind column (1) may be invalid, the problems of measurement error and quality variation are real enough, and they are ignored by the models in the other columns. Note too that all

results about bias are results about expectations or probability limits, and need not be true for any one realization.

Begin with the upper panel of Table A1, which contains the total expenditure elasticities. Most of these are remarkably robust across the different models, presumably because there is a great deal of variation in total expenditure, and the slope of the Engel curve is very well-determined and robust to a wide range of misspecifications. All of the wheat expenditure elasticities are between 0.28 and 0.48. Given that there is a quality elasticity of ten percent for wheat, see Table 4, column (1) gives an expenditure elasticity of 0.48, which is close to the 0.44 estimate in column (7) using Alderman's method. Apart from column (4), the other figures are too low. For rice, there is again good agreement between column (1) plus the quality elasticity of 0.11 and column (7). The other figures are not very different, although the estimates that ignore the within-village variance, columns (6) and (8), are clearly too high. Similar results apply to the other expenditure elasticities. If the fourth row of Table 4 is added to the first column of Table A1, the results are close to those in Column (7), so that the two "defensible" sets of estimates are in good agreement. Averaging over clusters in column (8) sometimes makes some difference, but is perhaps not of major importance. The biggest deviation is in column (2), where the double logarithmic specification on the heavily selected sample generates an elasticity for expenditure on dairy produce of 0.68, compared with an estimate that is greater than unity in the first column.

The price elasticities vary a good deal more. The price information in the data, although plentiful for most goods, is less plentiful than the information on outlay, and is less able to overcome the variations in specification. As is to be expected, there is most variability in estimates where there is least variability in the price, for edible oils and fats, and to some extent for sugar. The estimates using the urban prices generate *positive* price elasticities for these two goods, and the other methods show large differences, especially for sugar. We know from Section 2 and Table 5 that the price elasticities of these two categories are estimated with large standard errors, and this clearly applies to the estimates from the other models too. However, the imprecision is often not apparent from the calculated standard errors of these models, and several of the estimates in columns (2) through (8) would appear to be estimated quite precisely, if the regression results were to be treated seriously. It is difficult to draw any firm conclusions from the results for the other commodities. The averaging over clusters does not seem to have very marked effects, perhaps because most of the price variation is

between rather than within villages. Rice appears to be quite price elastic, whether we use the unit values or the urban prices, although the latter make it more so. It is perhaps also difficult to believe that wheat, which is the staple food, has a price elasticity (numerically) in excess of unity, although there may well be some substitution between wheat and rice. Dairy produce and meat have similar price elasticities in both approaches, although for other food, the estimate in column (1) is again smaller than that in Column (7). Apart from rice, where the effect is the other way, these results are perhaps consistent with an attenuation bias towards minus one in column (7). The double logarithmic specification in columns (3) and (4) does not seem to be very reliable, and sometimes produces markedly different answers depending on whether the cross-price effects are included. The selectivity problems that are involved in making this choice seem to render this a very unsatisfactory model in practice. Columns (4) and (5), which are the short-cut methods closest to column (1), produce the results that are closest, except for the obvious problems with oils and fats and sugar. But these specifications too suffer from the selectivity issue, albeit by less than the double logarithmic formulation.

The estimation of price elasticities is not an easy task, and it is perhaps too much to expect that there should be simple ways of providing good estimates. Some of the short-cut methods produce reasonable estimates for some commodities, but there is no guarantee that they would do so in other applications. In the absence of the full analysis there would be no way of knowing whether the approximations are adequate or catastrophic, and their use involves the analyst in a number of difficult and unsatisfactory compromises. The full analysis in column (1) may not produce the right answer, but we know that the other methods *cannot* deliver it, and will provide acceptable approximations only if the problems they ignore happen not to be serious in a particular application. The simplicity comes at too high a price.

Distributors of World Bank Publications

ARGENTINA
Carlos Hirsch, SRL
Galeria Guemes
Florida 165, 4th Floor-Ofc. 453/465
1333 Buenos Aires

**AUSTRALIA, PAPUA NEW GUINEA,
FIJI, SOLOMON ISLANDS,
VANUATU, AND WESTERN SAMOA**
D.A. Books & Journals
648 Whitestone Road
Mitcham 3132
Victoria

AUSTRIA
Gerold and Co.
Graben 31
A-1011 Wien

BAHRAIN
Bahrain Research and Consultancy
Associates Ltd.
Esterad Building No. 42
Diplomatic Area
P.O. Box 2750
Manama Town 317

BANGLADESH
Micro Industries Development
Assistance Society (MIDAS)
House 5, Road 16
Dhanmondi R/Areas
Dhaka 1209

Branch offices:
156, Nur Ahmed Sarak
Chittagong 4000

76, K.D.A. Avenue
Kulna

BELGIUM
Jean De Lannoy
Av. du Roi 202
1050 Brussels

CANADA
Le Diffuseur
C.P. 85, 1501 B rue Ampère
Boucherville, Québec
J4B 5E6

CHINA
China Financial & Economic
Publishing House
& Da Fo Si Dong Jie
Beijing

COLOMBIA
Infoenlace Ltda.
Apartado Aereo 34270
Bogota D.E.

COTE D'IVOIRE
Centre d'Édition et de Diffusion
Africaines (CEDA)
04 B.P. 541
Abidjan 04 Plateau

CYPRUS
MEMRB Information Services
P.O. Box 2098
Nicosia

DENMARK
Samfundslitteratur
Rosenørns Allé 11
DK-1970 Frederiksberg C

DOMINICAN REPUBLIC
Editora Taller, C. por A.
Restauración e Isabel la Católica 309
Apartado Postal 2190
Santo Domingo

EGYPT, ARAB REPUBLIC OF
Al Ahram
Al Galas Street
Cairo

The Middle East Observer
41, Sherif Street
Cairo

EL SALVADOR
Fusades
Avenida Manuel Enrique Araujo #3530
Edificio SISA, 1er. Piso
San Salvador

FINLAND
Akateeminen Kirjakauppa
P.O. Box 128
Helsinki 10
SF-00101

FRANCE
World Bank Publications
66, avenue d'Iéna
75116 Paris

GERMANY
UNO-Verlag
Poppelsdorfer Allee 55
D-5300 Bonn 1

GREECE
KEME
24, Ippodamou Street Platis Plastiras
Athens-11635

GUATEMALA
Librerías Piedra Santa
5a. Calle 7-55
Zona 1
Guatemala City

HONG KONG, MACAO
Asia 2000 Ltd.
48-49 Wymindham Street
Wingning Centre
2nd Floor
Central Hong Kong

INDIA
Allied Publishers Private Ltd.
751 Mount Road
Madras - 600 002

Branch offices:
15 J.N. Heredia Marg
Ballard Estate
Bombay - 400 038

13/14 Asaf Ali Road
New Delhi - 110 002

17 Chittaranjan Avenue
Calcutta - 700 072

Jayadeva Hostel Building
5th Main Road Gandhinagar
Bangalore - 560 009

3-5-1129 Kachiguda Cross Road
Hyderabad - 500 027

Prarthana Plaza, 2nd Floor
Near Thakore Baug, Navrangpura
Ahmedabad - 380 009

Patiala House
16-A Ashok Marg
Lucknow - 226 001

INDONESIA
Pt. Indira Limited
Jl. Sam Ratulangi 37
P.O. Box 181
Jakarta Pusat

ISRAEL
Yozmot Literature Ltd.
P.O. Box 56055
Tel Aviv 61560
Israel

ITALY
Licosa Commissionaria Sansoni SPA
Via Duca Di Calabria, 1/1
Casella Postale 552
50125 Florence

JAPAN
Eastern Book Service
37-3, Hongo 3-Chome, Bunkyo-ku 113
Tokyo

KENYA
Africa Book Service (E.A.) Ltd.
P.O. Box 45245
Nairobi

KOREA, REPUBLIC OF
Pan Korea Book Corporation
P.O. Box 101, Kwangwhamun
Seoul

KUWAIT
MEMRB Information Services
P.O. Box 5465

MALAYSIA
University of Malaya Cooperative
Bookshop, Limited
P.O. Box 1127, Jalan Pantai Baru
Kuala Lumpur

MEXICO
INFOTEC
Apartado Postal 22-860
14060 Tlalpan, Mexico D.F.

MOROCCO
Société d'Etudes Marketing Marocaine
12 rue Mozart, Bd. d'Anfa
Casablanca

NETHERLANDS
De Lindeboom / In-Or-Publikaties
P.O. Box 202
7480 AE Haaksbergen

NEW ZEALAND
Hills Library and Information Service
Private Bag
New Market
Auckland

NIGERIA
University Press Limited
Three Crowns Building Jericho
Private Mail Bag 5095
Ibadan

NORWAY
Narvesen Information Center
Book Department
P.O. Box 6125 Etterstad
N-0602 Oslo 6

OMAN
MEMRB Information Services
P.O. Box 1613, Seeb Airport
Muscat

PAKISTAN
Mirza Book Agency
65, Shahrah-e-Qaid-e-Azam
P.O. Box No. 729
Lahore 3

PERU
Editorial Desarrollo SA
Apartado 3824
Lima

PHILIPPINES
International Book Center
Fifth Floor, Filipinas Life Building
Ayala Avenue, Makati
Metro Manila

PORTUGAL
Livraria Portugal
Rua Do Carmo 70-74
1200 Lisbon

SAUDI ARABIA, QATAR
Jarir Book Store
P.O. Box 3196
Riyadh 11471

MEMRB Information Services
Branch offices:
Al Alas Street
Al Dahna Center
First Floor
P.O. Box 7188
Riyadh

Haji Abdullah Alireza Building
King Khalid Street
P.O. Box 3969
Dammam

33, Mohammed Hassan Awad Street
P.O. Box 5978
Jeddah

**SINGAPORE, TAIWAN,
MYANMAR, BRUNEI**
Information Publications
Private, Ltd.
02-06 1st Fl., Poi-Fu Industrial
Bldg.
24 New Industrial Road
Singapore 1953

SOUTH AFRICA, BOTSWANA
For single titles:
Oxford University Press
Southern Africa
P.O. Box 1141
Cape Town 8000

For subscription orders:
International Subscription Service
P.O. Box 41095
Craighall
Johannesburg 2024

SPAIN
Mundi-Prensa Libros, S.A.
Castello 37
28001 Madrid

Librería Internacional AEDOS
Consell de Cent, 391
08009 Barcelona

SRI LANKA AND THE MALDIVES
Lake House Bookshop
P.O. Box 244
100, Sri Chittampalam A.
Gardiner Mawatha
Colombo 2

SWEDEN
For single titles:
Fritzes Fackboksföretaget
Regeringsgatan 12, Box 16356
S-103 27 Stockholm

For subscription orders:
Wennergren-Williams AB
Box 30004
S-104 25 Stockholm

SWITZERLAND
For single titles:
Librairie Payot
6, rue Grenus
Case postale 3312
CH 1211 Geneva 11

For subscription orders:
Librairie Payot
Service des Abonnements
Case postale 3312
CH 1002 Lausanne

TANZANIA
Oxford University Press
P.O. Box 5299
Dar es Salaam

THAILAND
Central Department Store
306 Silom Road
Bangkok

**TRINIDAD & TOBAGO, ANTIGUA,
BARBUDA, BARBADOS,
DOMINICA, GRENADA, GUYANA,
JAMAICA, MONTserrat, ST.
KITTS & NEVIS, ST. LUCIA,
ST. VINCENT & GRENADINES**
Systematics Studies Unit
#9 Watts Street
Curepe
Trinidad, West Indies

UNITED ARAB EMIRATES
MEMRB Gulf Co.
P.O. Box 6097
Sharjah

UNITED KINGDOM
Microinfo Ltd.
P.O. Box 3
Aiton, Hampshire GU34 2PG
England

VENEZUELA
Librería del Este
Aptdo. 60.337
Caracas 1060-A

LSMS Working Papers (continued)

- No. 51 *Child Anthropometry in Côte d'Ivoire: Estimates from Two Surveys, 1985 and 1986*
- No. 52 *Public-Private Sector Wage Comparisons and Moonlighting in Developing Countries: Evidence from Côte d'Ivoire and Peru*
- No. 53 *Socioeconomic Determinants of Fertility in Côte d'Ivoire*
- No. 54 *The Willingness to Pay for Education in Developing Countries: Evidence from Rural Peru*
- No. 55 *Rigidité des salaires: Données microéconomiques et macroéconomiques sur l'ajustement du marché du travail dans le secteur moderne (in French only)*
- No. 56 *The Poor in Latin America during Adjustment: A Case Study of Peru*
- No. 57 *The Substitutability of Public and Private Health Care for the Treatment of Children in Pakistan*
- No. 58 *Identifying the Poor: Is "Headship" a Useful Concept?*
- No. 59 *Labor Market Performance as a Determinant of Migration*
- No. 60 *The Relative Effectiveness of Private and Public Schools: Evidence from Two Developing Countries*
- No. 61 *Large Sample Distribution of Several Inequality Measures: With Application to Côte d'Ivoire*
- No. 62 *Testing for Significance of Poverty Differences: With Application to Côte d'Ivoire*
- No. 63 *Poverty and Economic Growth: With Application to Côte d'Ivoire*
- No. 64 *Education and Earnings in Peru's Informal Nonfarm Family Enterprises*
- No. 65 *Formal and Informal Sector Wage Determination in Urban Low-Income Neighborhoods in Pakistan*
- No. 66 *Testing for Labor Market Duality: The Private Wage Sector in Côte d'Ivoire*
- No. 67 *Does Education Pay in the Labor Market? The Labor Force Participation, Occupation, and Earnings of Peruvian Women*
- No. 68 *The Composition and Distribution of Income in Côte d'Ivoire*
- No. 69 *Price Elasticities from Survey Data: Extensions and Indonesian Results*
- No. 70 *Efficient Allocation of Transfers to the Poor: The Problem of Unobserved Household Income*
- No. 71 *Investigating the Determinants of Household Welfare in Côte d'Ivoire*
- No. 72 *The Selectivity of Fertility and the Determinants of Human Capital Investments: Parametric and Semiparametric Estimates*
- No. 73 *Shadow Wages and Peasant Family Labor Supply: An Econometric Application to the Peruvian Sierra*
- No. 74 *The Action of Human Resources and Poverty on One Another: What We Have Yet to Learn*
- No. 75 *The Distribution of Welfare in Ghana, 1987-88*
- No. 76 *Schooling, Skills, and the Returns to Government Investment in Education: An Exploration Using Data from Ghana*
- No. 77 *Workers' Benefits from Bolivia's Emergency Social Fund*
- No. 78 *Dual Selection Criteria with Multiple Alternatives: Migration, Work Status, and Wages*
- No. 79 *Gender Differences in Household Resource Allocations*
- No. 80 *The Household Survey as a Tool for Policy Change: Lessons from the Jamaican Survey of Living Conditions*
- No. 81 *Patterns of Aging in Thailand and Côte d'Ivoire*
- No. 82 *Does Undernutrition Respond to Incomes and Prices? Dominance Tests for Indonesia*
- No. 83 *Growth and Redistribution Components of Changes in Poverty Measure: A Decomposition with Applications to Brazil and India in the 1980s*
- No. 84 *Measuring Income from Family Enterprises with Household Surveys*

The World Bank

Headquarters

1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

Telephone: (202) 477-1234

Facsimile: (202) 477-6391

Telex: WUI 64145 WORLDBANK

RCA 248423 WORLDBK

Cable Address: INTBAFRAD
WASHINGTONDC

European Office

66, avenue d'Iéna
75116 Paris, France

Telephone: (1) 40.69.30.00

Facsimile: (1) 40.69.30.66

Telex: 640651

Tokyo Office

Kokusai Building
1-1, Marunouchi 3-chome
Chiyoda-ku, Tokyo 100, Japan

Telephone: (3) 3214-5001

Facsimile: (3) 3214-3657

Telex: 26838



Deaton/Grimard