

Efficiency and Equity of a Marginal Tax Reform:
Income, Quality and Price Elasticities for Mexico*

Alessandro Nicita[⊗]

World Bank Policy Research Working Paper 3266, April 2004

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at <http://econ.worldbank.org>.

Keywords: Demand analysis, Household behavior, Mexico, Optimal taxation, Poverty.
JEL Classification: D1, H21, Q11, R2

* I wish to thank Jaime de Melo, Marcelo Olarreaga, Guido Porto, Isidro Soloaga, David Tarr, and the participants at the 2003 LACEA conference for helpful comments.

[⊗] Development Research Group, Trade, World Bank, 1818 H Street, NW, Washington DC 20433, mail stop MC3-303. Email: anicita@worldbank.org

Abstract

This paper investigates the effects of a marginal tax reform on household welfare in Mexico. It estimates the extent to which Mexican households react to changes in prices, and uses the estimates are then used to simulate changes in social welfare resulting from marginal tax reform. Results indicate that poorer households tend to have significantly larger income and price elasticities. Hence, to correctly evaluate the effects of economic policies on the poor, it is important to estimate elasticities that reflect the behavioral responses of the poor rather than of the entire population.

The results from the micro simulation suggest that since poorer households rely mostly on maize, legumes and vegetables to fulfill their caloric intake, any price reform that reduces the price of those products will have a larger effect on the welfare of poor households. In particular, reducing the taxes on maize, alcoholic beverages and vegetables would be both more equitable and more efficient in terms of social welfare. Meanwhile, a reduction in the tax on legumes, sugar, and oils and fats, while inefficient, would contribute to reduce inequality. Finally, a decrease in the price of meat, wheat, and dairy products, while benefiting higher-income households, would have only a marginal impact on poorer households.

1. Introduction

This paper investigates the efficiency and equity of the present price structure in Mexico. It estimates the behavioral responses of Mexican households to prices and income changes using a series of household surveys collected between 1989 and 2000, and then computes the efficiency and equity of the price structure, therefore providing information and policy guidance on the welfare effects of tax rate movements.

In estimating changes in social welfare, the knowledge of households' responses to prices is important for at least two reasons. First, there is a direct relationship between changes in prices and households' welfare (McCulloch, Winters and Cirera, 2002). Second, is the link between the prices of taxed goods and government revenues is especially critical in developing countries where taxes are collected through an indirect system that taxes consumer goods (Ahmad and Stern, 1991).

While behavioral responses are usually estimated through the analysis of time-series data, it is often the case that the scarce availability of data series for developing countries prevents the use of this standard methodology. Therefore, in order to calculate effectively how prices and income are reflected in the patterns of consumption, this paper adopts a methodology developed by Deaton (1987, 1988, 1990) and Deaton and Grimard (1992), which introduces a technique based on the variation of prices across space to identify behavioral responses. This paper extends Deaton's methodology in two directions. First, it estimates the elasticities differentiated by income quintile and by rural and urban areas.¹ Second, it applies Deaton's methodology to be used with panel data assuming each time-spatial location as a separate cluster, and thereby uses a series of six household surveys collected over a period of 12 years from 1989 to 2000.

The resulting estimates of income, price and cross price elasticities are utilized in a simulation exercise to investigate the efficiency and equity of the price structure in

¹ Especially for developing countries with high inequality, households' responses to changes in prices likely vary according to the income level of the household. Therefore, the use of aggregate price elasticities may be misleading.

Mexico. The results suggest that the efficiency and equity of the price structure can be improved through small movement in the tax of some key products. In particular, a reduction in the taxes levied on maize-based products would be beneficial both on the basis of equity and efficiency. Moreover, an increase in the taxes levied on oils and fats, pulses and sugar would be beneficial in terms of efficiency but would increase inequality.

Mexico was chosen as the country of analysis for several reasons, primarily the high quality and quantity of household surveys. In addition, Mexico implemented a considerable number of economic reforms that likely caused large changes in prices, which are necessary for the estimation of price elasticities. Finally, despite Mexico's status as a middle-income country, more than one-third of the population lives below the poverty line and there is a large degree of inequality (Ianchovicina, Nicita, Soloaga 2001). These features make Mexico a good candidate for estimating elasticities by income level and assessing the impact of a marginal tax reform.

The rest of the paper is organized as follows. Section 2 describes the consumption patterns of Mexican households. Section 3 presents the methodology for estimating household behavioral responses. Section 4 presents and discusses the elasticity estimates. Section 5 produces the results of a micro simulation to identify welfare-improving tax reforms. Section 6 concludes. Appendix A describes the data and appendix B gives further details on the micro simulation.

2. Consumption in Mexico

This paper focuses on 10 key commodity groups that constitute the bulk of the diet of Mexican households.² The commodity groups are aggregated over different qualities and varieties.³ Table 1.a reports yearly average budget shares for product aggregates disaggregated by income quintile and urban/rural areas.

² The ten product groups: wheat, maize, alcohol and tobacco, meat, legumes, dairy products, oil and fats, other vegetables, fruit, and sugar. In total, those product aggregates account for about one third of the expenditures basket of the average household.

³ For example, meat includes beef, chicken and pork as well as different qualities of cut.

Mexican households on average spend about 40% of their income on food products. Maize, meat, dairy products and vegetables constitute the most important product groups in the household basket and account for about half of the food expenses or 20% of the total expenditure of Mexican households. Slightly less important in terms of total expenditures are wheat-based products, oils and fats, legumes and fruits. Finally, alcohol and tobacco, and sugar each account for less than 1% of total expenditures. The richer households allocate slightly less than one-fourth of income to purchase food products. This percentage is much higher for the poorest households (about 50%). For poor households, maize-based products (7.1%) constitute the largest of the food expenses. Important are also meat (7.0%), dairy products (6.0%) and vegetables (5.3%). Legumes with 4.7% of total expenditures constitute an important source of nutrients for poorer households. Differences in urban and rural areas are relatively small and probably are driven by income more than urban/rural differentiation. Table 1.b shows the importance of each of the food categories in total food consumption across the five income deciles. The expenditure share of food products that are considered inferior, such as maize, legumes and vegetables, rapidly fades with the increase in income. Meanwhile, products perceived as luxury, those with a higher cost per calorie, such as meat, dairy products, and fruit, are consumed more in higher income households.

In summary, table 1.a and 1.b reveal three stylized facts of consumption patterns: 1) Household consumption of food products is about 50% for poor households, while for the richest it is about 25%. 2) The consumption basket of poor households is biased towards maize, legumes and vegetables, while that of rich households is biased toward meat and dairy products. 3) Differences between urban and rural areas seems to be driven by the different levels of income.

3. The Measurement of the Elasticities

As mentioned in the introduction, time series of household surveys are rarely available, particularly for developing countries, making it more difficult to estimate price and

income elasticities. To circumvent the data constraints, Deaton (1988) developed a technique for estimating price elasticities using cross-sectional household surveys when there is spatial variation in prices. The data requirements to estimate consumers' behavior are limited to household expenditures by commodity, quantity of each commodity purchased, spatial localization and household characteristics. These variables are usually collected in most household surveys. Deaton's method assumes no price variations within a spatial location (cluster), and therefore the variation in unit values reflects only quality differentials and measurement errors. Because of available data across time, this paper makes use of both the spatial and time dimensions to identify clusters. This extension circumvents the fact that since many developing countries have data on household surveys that cover several years but are limited in the number of geographical locations, Deaton's methodology may be inadequate due to the scarcity of geographic clusters in the case of a single household survey.

In this paper I assume that prices do not vary within each year in the same way that Deaton assumes that they do not vary within each geographic location. This assumption implies that the change in prices between years is similar across clusters. Once changes in demand across years are corrected using fixed effects estimators, then the model can be estimated considering each time-spatial location as a separate cluster. This method results in a much larger number of clusters, therefore yielding more precise elasticity estimates.

The essence of this technique can be explained with a one good model.⁴ The model is composed of two basic equations. The simplest possible model of demand links the quantity consumed to prices and to the purchasing power of the household (income). Under the assumption that prices do not vary within clusters and that prices are not observable (but unit values are) the model consists of a demand function:

$$\ln q_{hct} = \alpha_1 + \beta_1 \ln x_{hct} + \gamma_1 z_{hct} + \theta_1 \ln p_{hct} + f_c + u_{1hct} \quad (1)$$

⁴ In the multi-good model the own price elasticities are estimated in the same way. The hurdle is given by the introduction of estimates of cross price elasticities.

which is a double logarithm demand function, and

$$\ln v_{hct} = \alpha_2 + \beta_2 \ln x_{hct} + \gamma_2 z_{hct} + \theta_2 \ln p_{hct} + u_{2hct} \quad (2)$$

which is the unit value equation, where h denotes the household, c spatial cluster and t the time cluster. The variables are labeled as:

- q_{hct} = quantity purchased of the good by household h in cluster c ,
- x_{hct} = total expenditures,
- z_{hct} = vector of household characteristics,
- p_{hct} = unobserved cluster price,
- f_c = unobserved fixed cluster effect capturing taste variation across clusters,
- v_{hct} = unit value,
- u_{hct} = residual error terms.⁵

The estimation procedure discussed in detail below involves three stages and follows Deaton (1989). The first stage uses within-cluster information on households' demand, income and products' unit values to obtain estimates of total expenditure and quality elasticities⁶, as well as estimates of error-measurement variances and covariances. The second stage uses the estimates from the first stage in netting out the effect of the total expenditure, quality, and household characteristics and therefore calculates the "corrected" budget shares and unit values. A regression of "corrected" budget shares on "corrected" unit values, averaged by cluster, produces an estimate of the ratio of the responses to price of the budget share and the unit value. Finally, in the third stage, the

⁵ One of the novelties of this paper with respect to Deaton's methodology resides in the fact that idiosyncratic time components are controlled for by taking a fixed effect approach in estimating equations 1 and 2. Hence, the error term has the form: $u_{hct} = \omega_t + \eta_{hct}$, where ω_t is the fixed effect for year t and η_{hct} is an i.i.d. term with mean zero and variance σ .

⁶ Quality elasticities refer to the change in the quality of purchased products corresponding to a change in household income.

effect of price on the budget share is extracted from the ratio by using the theory linking quality and quantity elasticities.

Household characteristics consist of the logarithm of household size, a series of age-gender ratios,⁷ a dummy variable for skilled workers, and a rural indicator. Since time and regional differences in tastes may result in demand not being the same across time and across region, the covariates are augmented with regional and time dummies. In the first stage, cluster means are subtracted from the variables in equations (1) and (2) and then those equations are estimated using OLS to obtain estimated parameters $\hat{\beta}_i$. These estimates measure the influence of income and household characteristics on demand and quality.

The size of the quality effect can be easily estimated using the methodology of Prais and Houthakker (1955). The underlining assumption is that, *ceteris paribus*, the richer the household, the higher the quality of the products that the household consumes. Because unit values not only vary with the choice of quality, but also with actual market prices, prices theoretically should be included in the estimation. The fact that prices are not observed in household surveys makes it impossible to obtain directly parameter estimates for the price variable. Nevertheless, it is possible to estimate consistently the non-price parameters if we assume that all households in the same cluster face the same vector of prices. Therefore the estimation of equation (2) is conducted by adding dummies for each cluster.⁸ A further problem with the estimation of equations (1) and (2) arises in the case when households do not purchase a particular category of goods. Because the logarithmic specification can be used only to describe the behavior of those households who purchase positive amounts, the logarithmic form can exclude a large part of the households for narrowly defined product groups. To avoid this problem, Deaton and Muellbauer (1980)

⁷ The gender and age composition of the household is likely to be reflected in the consumption basket of the household. The gender-age ratios are constructed as the number of a determinate category of individuals over total household members. In the estimation, individuals are divided by gender and by three age groups (children, adults, elderly), resulting in six ratio variables.

⁸ This is equivalent to estimating the regression of deviations from cluster means

suggest estimating the model substituting budget shares for the logarithm of quantity purchased.

Removing the cluster means from (1) and (2) and substituting the logarithm of quantity consumed with budget shares (w) gives the estimating equations:

$$w_{hct} - \bar{w}_{hct} = \beta_1 (\ln x_{hct} - \ln \bar{x}_{hct}) + \gamma_1 (z_{hct} - \bar{z}_{hct}) + (u_{1hct} - \bar{u}_{1hct}) \quad (3)$$

$$\ln v_{hct} - \ln \bar{v}_{hct} = \beta_2 (\ln x_{hct} - \ln \bar{x}_{hct}) + \gamma_2 (z_{hct} - \bar{z}_{hct}) + (u_{2hct} - \bar{u}_{2hct}) \quad (4)$$

Estimation of (3) and (4) yields first stage estimates $\hat{\beta}_i$. Using these estimates, one can define the budget shares and unit values purged by the effect of total expenditure. To do so, the following two variables are constructed:

$$\hat{y}_{1hct} = w_{hct} - \hat{\beta}_1 \ln x_{hct} \quad (5)$$

$$\hat{y}_{2hct} = \ln v_{hct} - \hat{\beta}_2 \ln x_{hct} \quad (6)$$

Corresponding to those estimates are the true values of y_{1c} and y_{2c} which we would have obtained had the first-stage parameters been known rather than estimated. These can be written from (1) and (2) and averaging equations over household by clusters:

$$y_{1ct} = \alpha_1 + \theta_1 \ln p_{ct} + f_c + u_{1ct} \quad (7)$$

and

$$y_{2ct} = \alpha_2 + \theta_2 \ln p_{ct} + u_{2ct} \quad (8)$$

wherein y_{1ct} and y_{2ct} the absence of the h subscript indicates cluster averages.

In the second stage, between-cluster variation of budget shares and unit values is used to estimate price elasticities. Nevertheless, in equations (7) and (8) the θ coefficients cannot

be estimated directly because the cluster prices p_{ct} are not observed. To solve this problem, the ratio $\phi = \theta_1 / \theta_2$ is obtained from the covariance of \hat{y}_{1ct} and \hat{y}_{2ct} . To correct for measurement error underlining the first-stage estimation, first the variance of u_{2ct} and the covariance of u_{1ct}, u_{2ct} need to be estimated using the residual from the first stage estimation:

$$\hat{\sigma}_{11} = \frac{e_2' e_2}{(n-k-C-1)} \text{ and } \hat{\sigma}_{01} = \frac{e_1' e_2}{(n-k-C-1)} \quad (9)$$

where e_1 and e_2 are the residuals from the first stage regression and the denominator is the number of degrees of freedom.

Next the variance of y_{2ct} and covariance of y_{1ct}, y_{2ct} are computed from (7) and (8). For a large number of clusters the formulas used are:

$$\text{cov}(\hat{y}_{1ct}, \hat{y}_{2ct}) = \theta_1 \theta_2 m + \sigma_{01} / n_{ct} \quad \text{and} \quad \text{var}(\hat{y}_{2ct}) = \theta_2^2 m + \sigma_{11} / n_{ct} \quad (10)$$

where m is the large-sample inter-cluster variance of the logarithm of prices, and n_{ct} is the number of households per cluster.

Solving for $\phi = \theta_1 / \theta_2$ we obtain:

$$\hat{\phi} = \frac{\text{cov}(\hat{y}_{1ct}, \hat{y}_{2ct}) - \hat{\sigma}_{01} / n_{ct}}{\text{var}(\hat{y}_{2ct}) - \hat{\sigma}_{11} / n_{ct}} \quad (11)$$

where the terms $\hat{\sigma}_{01}$ and $\hat{\sigma}_{11}$ are designed to correct for the part of the between-cluster variances and covariance that comes from measurement and econometric error in the underlying first-stage equations. The estimator $\hat{\phi}$ is a standard errors in the variable

estimator and is composed of the OLS estimator (covariance over variance), plus a term that corrects for the effect of measurement error underlying the first-stage equations.⁹

The final step in Deaton's methodology is to derive a consistent set of expenditures (proxy for income) ε_{ex} and price elasticities ε_{pr} from $\hat{\phi}$ and the following relationships: (a) between the price and income elasticities of unit value (θ_2 and β_2); (b) between price and expenditure shares (θ_1 and β_1); and (c) between price and quantities (ε_{pr} and ε_{ex}). In Deaton's model, the theoretical relationship between quality, price and expenditure relationship is given by:

$$\theta_2 = 1 + \beta_2 \varepsilon_{pr} / \varepsilon_{ex} \quad (12)$$

Now, from (1) and (2), the expenditure elasticity is given by β_1 / w , but since the budget share is unit values times quantity divided by total expenditure, we have:

$$\varepsilon_{ex} + \beta_2 = (\beta_1 / w) + 1 \quad (13)$$

Similarly, the price elasticity is:

$$\varepsilon_{pr} + \theta_2 = (\theta_1 / w) \quad (14)$$

Substituting (13) and (14) into (12) and rearranging, we have:

$$\theta_2 = 1 - \frac{\beta_2 (w - \theta_1)}{\beta_1 + w} \quad (15)$$

Substituting $\phi = \theta_1 / \theta_2$ into (15) and solving for θ_1 yields:

⁹ As a clarification, the covariance ($\hat{y}_{1ct}, \hat{y}_{2ct}$) is calculated directly from the \hat{y}_{1ct} and \hat{y}_{2ct} corrected from the effects of total expenditures as in (5) and (6).

$$\theta_1 = \frac{\phi}{1 + (w - \phi)\xi} \quad (16)$$

where

$$\xi = \frac{\beta_2}{\beta_1 + w(1 - \beta_2)} \quad (17)$$

Similarly, θ_2 is obtained by substituting (16) into (15). Finally, with the parameters θ_1 and θ_2 recovered, the price elasticities will be calculated from (14).

This single good model can be extended in a multi-good case, so to estimate cross-price elasticity. To do so, ones needs to add additional price terms to the budget share equation (equation 3) and model quality under the separability assumption. Describing the whole method is out of the scope of this paper as the algebra for this exercise would require the introduction of a good deal of new notation. The interested reader can refer to Deaton, 1997.¹⁰

4. Income, Quality and Price Elasticities for Mexican Households

Income, quality and price elasticities are produced for 10 product groups: wheat, maize, alcohol and tobacco, meat, legumes, dairy products, oils and fats, vegetables, fruits and sugar. As seen in the second section, the importance of these product groups is illustrated by the fact that they represent about one-third of the total expenditure basket of the average Mexican household, and more than 40% of the total expenditure of the households in the first income quintile. The estimated elasticities are presented for the totality of the households, differentiated by income quintile and by rural and urban areas. Bootstrapped¹¹ standard errors for the price elasticities are estimated by making 500

¹⁰ Deaton goes into this topic from equation 5.56 to equation 5.102, in chapter 5.3.

¹¹ The bootstrap involves repeated re-estimation of a parameter using random samples with replacement from the original data. The idea behind this technique is to treat the sample as if it were the whole population. The statistic of interest can then be calculated from each of these samples. The average of the statistics over all of these samples is known as the bootstrap estimator. With repeated estimates of the same

draws from the cluster (second stage) data, and then finding the length of the interval that is symmetric around the bootstrapped mean and that contains 95.4 percent of the bootstrapped mean.¹²

4.1 First-Stage Estimates: Income and Quality Elasticities

Table 2 presents the income elasticity estimates and quality elasticity estimates using the methodology illustrated in section 3.

Overall income elasticities vary from 0.386 to 0.861. As expected, product aggregates that constitute the diet of poorer households have lower income elasticities. In particular, food products perceived as necessities, such as legumes and maize, result in the lower income elasticities (0.386 and 0.422). More expensive and appealing products such as meat, dairy products, and fruit, have higher elasticities respectively with 0.850, 0.716 and 0.861. Wheat, vegetables, oils and fats and sugar are reasonably elastic with their income elasticities ranging from 0.464 to 0.630.

Table 2 also reports the estimates for quality elasticities which vary from 0.008 (sugar) to 0.262 (alcohol and tobacco). Sugar, oil and fats, have very low quality elasticities indicating homogeneity within the product aggregates and a similar variety and quality consumption across households. The remaining categories of goods have larger quality elasticity estimates suggesting a wider range of choices and quality for the consumers and a preference of richer households towards the more expensive. Because of the wide range of quality available to consumers, alcoholic beverages and tobacco products resulted in the highest quality elasticity.

statistic its (bootstrapped) standard error can be calculated. The accuracy of the bootstrap estimator increases with the number of samples drawn. In order to construct confidence intervals, this paper follows the methodology proposed by Hall (1995) which bootstraps the distribution of the t-values rather than the estimate itself.

¹² That is the fraction of a random variable normally distributed that lies within two standard deviations from the mean.

Table 2 also reports elasticity estimates by income quintiles.¹³ Not surprisingly, in most cases income elasticities diminish with the increase in income while quality elasticities increase. The fact that the lowest quintile has the highest elasticities indicates that any change in income for the households within the first quintile affects the quantity purchased in a much larger way than any change in income for the richer quintiles. In particular, for the lowest quintile the only two product categories that have income elasticities lower than one are legumes and alcohol and tobacco.¹⁴ For all the other product aggregates, including maize based products, the income elasticity is more than one. Far from being a luxury good, this estimate suggests that additional income is spent on maize because this is the most efficient product to increase caloric intake.

Examining the income elasticities across quintiles, meat and dairy products, arguably the most expensive goods in terms of caloric content, have high income elasticities. Not surprisingly, legumes, vegetables and cereals, which are the bulk of the diet of poorer households, have lower elasticities. The income elasticities for the two most important cereals, maize and wheat, follow a downward pattern as income increases. But while maize products have an income elasticity that quickly becomes close to zero as income increases, the income elasticity of wheat also decreases but at a much slower pace. This reflects the fact that, as seen in the budget shares, wheat-based products constitute the cereal of choice for richer households, while maize-based products are the usual alternative of poorer households. Product aggregates for more appealing and expensive products such as meat, dairy, and fruits have quite high elasticities for the lower quintiles and in many cases these account as real luxury products for the poorer households. Furthermore, the elasticities for these products are still substantially above zero, even for the richer households. Quality elasticities by quintiles indicate that poorer households have a more homogeneous consumption (lower quality elasticities) for most of the products and especially for maize (0.102) and wheat (0.073). With the increase of

¹³ As a clarification, the elasticities are calculated based on income within quintile. This means that income elasticities need to be interpreted as the effect that the variation of income within the quintile has on the purchase of the products. Quality elasticities are to be interpreted as how the income differential within the quintile affects the quality choice of the product aggregate.

¹⁴ As a cautionary note, elasticities for alcohol and tobacco need to be regarded carefully since they are subjected to higher measurement error and higher quality differential.

income, the households tend to diversify their consumption to a wider quality of goods within the product aggregates. For example, the quality elasticity of maize grows to 0.160 for the second quintile up to 0.283 for the richest fifth quintile.

Table 2 also reports estimates differentiated by urban and rural areas.¹⁵ Urban households tend to have lower income elasticities and higher quality elasticities. This is likely due to the fact that, on average, urban households tend to be richer than rural ones. For the same reason, the opposite pattern is found in the analysis of quality elasticities, with rural households having lower quality elasticities estimates than those of urban households.

4.2 Second-Stage Estimates: Price Responses

Table 3 gives the own price elasticity estimates for the 10 product categories.¹⁶ All the price elasticities are negative and in several cases their absolute value is above one. In particular, wheat, meat, dairy and fruit products have the highest price elasticity. The price elasticity for sugar is even higher but insignificantly different from zero. This is due to the fact that this elasticity is poorly determined because its price varies little. The same applies for oils and fat. As expected, maize based products, which constitute the base of the diet of poor households have a low price elasticity (0.741 in absolute value).

Legumes, an important staple for Mexican households, also has in a very low elasticity (0.161) but not significantly different from zero. Looking at wheat, the likely substitute for maize, its price elasticity is higher (1.288) than the one of maize (0.741).

Price elasticities across income quintiles indicate that, in most cases, changes in prices affect consumption patterns only for the poorer households. For the richest fifth quintile no product aggregate resulted with a price elasticity significantly different from zero. Also, for the third and fourth quintiles few products resulted in price elasticities significantly different from zero. Vice versa, for the first quintile only alcohol and tobacco, oils and fats, fruit and sugar have an elasticity estimate not significantly

¹⁵ Further differentiation of urban and rural households by income quintile failed to produce significant estimates.

¹⁶ The price elasticities are symmetry-constrained estimates purged from the quality effect.

different from zero.¹⁷ Similarly for the second quintile, legumes, oils and fats, fruit and sugar resulted in price elasticities not significantly different from zero. Comparing the price elasticities for cereals, they are substantially higher for the first quintile. Maize based products have a price elasticity of 0.860 for the first quintile, and it stabilizes around 0.6 for the second quintile to the fourth. Wheat resulted in having quite a high price elasticity for all the poor households with elasticities (in absolute values) consistently more than 1 for the first three quintiles. With wheat-based products being substantially more expensive than maize-based products, these high elasticities may suggest that poorer households have strong substitutability between these two cereals. Similar patterns of high price elasticity for poorer households can be found in all the product groups. The price elasticity for meat goes from 2.232 for the first quintile to 0.886 for the third, while dairy products resulted in a more constant price elasticity of around 1.1 for all the quintiles.

The comparison of price elasticities between urban and rural areas reveals substantially higher price elasticity in rural areas for meat, cereals, and vegetables.¹⁸ Vice versa, elasticity for dairy products are estimated to be slightly higher for urban areas, but not significantly so. Overall, these results suggest that price responses are driven mostly by income rather than urban/rural differentiation.

4.3 Second-Stage Estimates: Cross-price Elasticities

Significant cross-price elasticity estimates could be obtained only in the case where all households were included in the analysis. Therefore, estimates are produced for the totality of households without distinguishing across quintiles or between urban and rural areas.¹⁹ In many cases cross-price elasticities estimates were not significantly different from zero. Nevertheless, the analysis was able to detect a few product groups that are substitutes or complementary for Mexican households. Table 4 reports the complete

¹⁷ Those products, with the exception of fruits, have little variance in their price and therefore fail to produce a precise estimate.

¹⁸ Although only the elasticity for maize is significantly different between urban and rural areas.

¹⁹ Cross-price elasticities were estimated in all cases, but in every case besides that of the total households their large standard error resulted in no cross-price elasticity different from zero.

matrix of symmetry constrained cross-price elasticities for the Mexican households. Since the inclusion of the bootstrapped standard error would make the table difficult to read, the values for which the estimate is greater than twice the bootstrapped standard error are reported in bold.²⁰

The results show a number of important patterns. Maize, meat, dairy products and legumes are substitutes. Maize-based products and vegetables are complementary, as are legumes and wheat, and vegetables and oils and fats. The estimation fails to capture any substitutability effect between wheat and maize. Their cross-price elasticities have the correct sign but it is not significantly different from zero.

5. Microsimulation: Marginal Tax Reform and Social Welfare

Using the estimates of quality, price and cross-price elasticities, it is possible to measure the welfare gains due to small changes in the tax rates of every good. In doing so, this paper utilizes the marginal tax reform approach which takes the existing tax system and identifies directions of reform at the margin (Ahmad and Stern, 1984).²¹ Since the methodology reflects the one of Deaton (1997), I briefly illustrate the idea behind this methodology while referring to the appendix for a detailed discussion.

In the simplest form, taxes affect social welfare through two channels: a movement in the tax rate of a particular good i will cause a change in welfare ($\partial W / \partial t_i$) but also a change in government revenues ($\partial R / \partial t_i$).²² The ratio of those two measures, known as the Pigovian ratio:

$$\frac{\partial W / \partial t_i}{\partial R / \partial t_i} = \lambda_i \quad (18)$$

²⁰ In table 4, the change in the price of products listed in the columns affects purchases of the products listed in rows.

²¹ This approach is different from the traditional optimal tax literature which attempts to derive the tax rates that minimize welfare losses for the collection of a given revenue.

²² Price and cross-price elasticities enter into the calculation of the government revenues and therefore the benefits on the tax reforms.

gives the marginal social cost (λ_i) of raising one unit of revenue from an increase in the tax of good i . When λ_i is above one (i.e. the cost of the tax is higher than its benefit) then social welfare would be improved through a decrease of the tax. The opposite is true when λ_i is below one.

Using Roy's identity, it can be shown that the welfare effect of a tax-induced marginal price change is given by the sum of each household's consumption of the good i .

$$\frac{\partial W}{\partial t_i} = -\sum_{h=1}^H \eta_h q_{ih} \quad (19)$$

where η_h is the social marginal utility of money in the hands of household h .²³

$$\eta_h = \frac{\partial V}{\partial u_i} \frac{\partial \psi_h}{\partial x_h} = \frac{\partial W}{\partial x_i} \quad (20)$$

If the households are valued differently, for example giving more importance (weight) to poor households, then it is possible to estimate the effect of the tax reform with respect to the poor instead of to the average household.²⁴ More precisely, the weight given to particular households depends on the degree of inequality aversion that society is willing to put in the calculation (so higher weight for poorer households will identify a tax reform that is inequality reducing). For this scope, using Atkinson's social welfare function, the numerator of equation 18 can be written as:

²³ Social welfare is defined as the sum of each households welfare as in $W = V(u_1, u_2, \dots, u_n)$, with $u_h = \psi(x_h, p)$

²⁴ Similarly, one could give more weight to households of a particular geographic region or ethnic group.

$$\frac{\partial W}{\partial t_i} = -\sum_{h=1}^H \eta_h q_{ih} = -\frac{1}{H} \sum_{h=1}^H \left(\frac{x_h}{n_h} \right)^{-\varepsilon} q_{ih} \quad (21)$$

where the term x_h/n_h is the per capita expenditure of household h . The higher the inequality aversion parameter ε , the higher is the weight given the lower per capita expenditure households in the computation of λ_i . In the case that ε is zero, all households have the same weights; therefore the effect of tax reform will be calculated on the average household.

To see how price elasticities enter the calculation of changes in social welfare, the denominator of equation 18 can be written as:

$$\partial R / \partial t_i = 1 + \frac{t_i}{1+t_i} \left(\frac{\theta_{ii}}{\tilde{w}_i} - 1 \right) + \sum_{k \neq i} \frac{t_k}{1+t_k} \left(\frac{\theta_{ki}}{\tilde{w}_i} \right) \quad (22)$$

where θ_{ii} is the own price and θ_{ij} the cross price elasticity and \tilde{w}_i is the budget share for the average households.

The usual practice, which is followed in this paper, is to analyze the results of the various cost-benefit ratios λ_i calculated for a range of values of ε .²⁵ Results are given in table 5.

The first column gives the OECD estimates of the Mexican tax structure.²⁶ It represents the tax (or subsidy) in percentage terms that each of the goods is subject to. These are calculated taking into account tariffs, taxes, subsidies and other transfers, all of which contribute in setting the domestic price of the product. OECD estimates suggest that government intervention causes higher prices for all goods, with the exception of

²⁵ For the exact formula used in the calculation see the appendix.

²⁶ Producer and Consumer Support Estimates, OECD Database 1986-2002

vegetables. Products such as sugar, alcoholic beverages, dairy, and meat bear taxes between 25 and 40 percent. Distortionary measures are lower in the case of fruit, legumes and maize (between 10 and 20 percent). Finally, minimal distortions are estimated for wheat and oils and fats.²⁷

The remaining columns in table 5 give two indicators for the four values of the aversion parameter ε : the relative budget shares of an increasingly poor household (w^ε) relative to the market-representative household (\tilde{w}) and the cost benefit ratio λ_i . From the ratio w^ε/\tilde{w} , it is possible to infer the importance of the good in the consumption basket of poor households relative to the average households. Maize, oils and fats, legumes and sugar become increasingly more important in the basket of poor households. On the other hand, fruit and meat products lose importance. Budget shares of the other food products also increase for poorer households but not in a remarkable way.²⁸

The effect of a small change in the tax rate can be inferred from the analysis of the cost benefit ratio λ . In the case of no aversion to inequality ($\varepsilon = 0$), the most efficient way to increase social welfare would be to decrease the price of maize and alcoholic beverages ($\lambda > 1$). On the contrary, sugar, oils and fats, and legumes are targets for price increase ($\lambda < 1$).

The importance of reducing the taxes on maize to favor social welfare and reduce inequality, as well as in the case of alcoholic beverages and vegetables is evident when values of the ratio λ_i are compared across the different values of the aversion parameter ε . However an increase in social welfare is not always accompanied by reductions in inequality, as in the case of sugar and legumes, and marginally for oils and fats. In these cases an increase in the tax would produce welfare gains for the average households, while the tax increase instead diminishes welfare with respect to poorer households, hence increasing inequality.

²⁷ The last category in the table, necessary for symmetry in the estimation, represents the expenditure of the household on all other goods.

²⁸ This reflects the fact the richer the household, the more its income is spent on non-food products.

In summary, the most important food product for tax reform is maize, which figures relatively heavily in the budgets of low-income households. A reduction of its price is desirable for both equity and efficiency reasons. Legumes, alcoholic beverages and vegetables are also candidates for price decrease on the basis of efficiency and equity. On the other hand, taxes could be increased in the case of fruits, oils and fats, meat and non-food products without substantially decreasing welfare of poor households.

6. Conclusions

This paper adapts the Deaton methodology to derive income, quality, own-price and cross-price elasticities for 10 aggregated goods and for five levels of income of Mexican households using six household surveys collected between 1989 and 2000. Those estimates are then fed into a micro simulation model to estimate the effect of a marginal tax reform.

The results on the elasticity estimates indicate that there is a substantial difference between estimates obtained when households are differentiated by level of income. Poorer households tend to have larger income elasticities, suggesting a level of income barely able to support the basic necessities of the household. Similarly, price elasticities are usually larger for poorer households, likely because of more stringent budget constraints. For richer households, response to change in prices tends to fade. Results also found that differences between urban and rural areas are small and likely to be caused by differences in income more than differences in behavioral parameters between rural and urban households.

The results from the micro simulation suggest that since poorer households rely mostly on maize, legumes and vegetables to fulfill their caloric intake, any price reform that reduces the price of those products will have a larger effect on the welfare of poor households. Furthermore, a reduction in the taxation on maize, alcoholic beverages and vegetables would be both more equitable and more efficient in terms of social welfare.

Meanwhile, a decrease in the tax on legumes, sugar, and oils and fats, while inefficient, will have a redistributive welfare effect from the richer to the poorer households. Finally, a decrease in the price of meat, wheat, and dairy products, while benefiting higher-income households, would have only a marginal impact on poorer households.

References:

- Ahmad Ehtisham and Nicholas Stern (1991), "The theory and practice of tax reform in developing countries". Cambridge University Press. New York.
- Deaton, A and J. Muellbauer (1980), "Economics and Consumer Behavior" Cambridge University Press. New York.
- Deaton, Angus (1987), "Estimation of Own- and Cross-Price Elasticities from Household Survey Data", *Journal of Econometrics*, 36:7-30.
- Deaton, Angus (1988), "Quality, Quantity and Spatial Variation of Price", *American Economic Review*, 78:418-430.
- Deaton, Angus (1990), "Price Elasticities from Survey Data. Extensions and Indonesian Results", *Journal of Econometrics*, 44:281-309.
- Deaton, A. and F. Grimard (1992), "Demand Analysis and Tax Reform in Pakistan", *LSMS working paper*. 0253-4517; no. 85.
- Deaton, Angus (1997), "The analysis of household surveys : a microeconomic approach to development policy", Johns Hopkins University Press. Baltimore.
- Hall, Peter (1995), "Methodology and Theory for the Bootstrap", in Robert F. Engle and Daniel L. McFadden, eds., *Handbook of Econometrics*, 4, Amsterdam, Elsevier, 2341-81.
- Ianchovichina, E., A. Nicita, and I. Soloaga "Trade Reform and Poverty: The Case of México," *World Economy*; 25, 945-972.
- McCulloch N., A. Winters and J. Cicera (2001), "Trade Liberalization and Poverty: A Handbook".CEPR.

Table 1a: Average budget shares of Mexican households

Budget shares	Average	Quintile1	Quintile2	Quintile3	Quintile4	Quintile5	Urban	Rural
Wheat	2.9%	3.6%	3.5%	3.0%	2.6%	1.6%	2.7%	3.2%
Maize	4.0%	7.1%	5.4%	3.8%	2.5%	1.1%	3.4%	5.5%
Alcohol Tobacco	0.8%	1.1%	0.8%	0.8%	0.7%	0.6%	0.8%	0.9%
Meat	8.9%	7.0%	10.3%	10.7%	9.9%	6.4%	9.0%	8.5%
Legumes	2.1%	4.7%	2.8%	1.7%	1.1%	0.5%	1.9%	2.8%
Dairy	5.8%	6.0%	6.9%	6.6%	5.8%	3.8%	5.8%	5.9%
Oil Fats	1.4%	2.6%	1.8%	1.3%	0.9%	0.4%	1.2%	2.0%
Vegetables	3.6%	5.3%	4.4%	3.7%	3.1%	1.7%	3.5%	4.0%
Fruit	1.4%	1.1%	1.5%	1.6%	1.6%	1.3%	1.5%	1.3%
Sugar	0.9%	2.1%	1.1%	0.7%	0.4%	0.2%	0.8%	1.2%
Other food	7.7%	9.5%	8.6%	8.1%	7.2%	5.0%	7.3%	8.4%
Total Food	39.5%	50.1%	47.1%	42.0%	35.8%	22.6%	37.9%	43.7%
Non Food	60.5%	49.9%	52.9%	58.0%	64.2%	77.4%	62.1%	56.3%

Source: Author's calculation using INEGI households surveys 1984-2000.

Table 1b: Share of expenditure in the 10 food groups as percentage of total food expenditures.

Budget shares	Quintile1	Quintile2	Quintile3	Quintile4	Quintile5
Wheat	7.2%	7.4%	7.1%	7.3%	7.1%
Maize	14.2%	11.5%	9.0%	7.0%	4.9%
Alcohol Tobacco	2.2%	1.7%	1.9%	2.0%	2.7%
Meat	14.0%	21.9%	25.5%	27.7%	28.3%
Legumes	9.4%	5.9%	4.0%	3.1%	2.2%
Dairy	12.0%	14.6%	15.7%	16.2%	16.8%
Oil Fats	5.2%	3.8%	3.1%	2.5%	1.8%
Vegetables	10.6%	9.3%	8.8%	8.7%	7.5%
Fruit	2.2%	3.2%	3.8%	4.5%	5.8%
Sugar	4.2%	2.3%	1.7%	1.1%	0.9%
Other food	19.0%	18.3%	19.3%	20.1%	22.1%

Source: Author's calculation using INEGI households surveys 1984-2000.

Table 2: Quality and income elasticity

<i>Quality Elasticities</i>	Average	1st Quintile	2nd Quintile	3th Quintile	4th Quintile	5th Quintile	Rural	Urban
Wheat	0.12	0.07	0.16	0.16	0.16	0.12	0.09	0.13
Maize	0.18	0.1	0.16	0.14	0.16	0.28	0.16	0.19
Alcohol Tobacco	0.24	0.2	0.73	0.45	0.35	0.07	0.15	0.28
Meat	0.11	0.13	0.14	0.13	0.1	0.09	0.08	0.12
Legumes	0.08	0.07	0.09	0.09	0.11	0.09	0.07	0.08
Dairy	0.11	0.06	0.16	-0.05	0.17	0.19	0.1	0.12
Oil Fats	0.02	0	-0.05	-0.01	0.04	0.08	0.02	0.02
Vegetables	0.1	0.08	0.11	0.12	0.17	0.09	0.09	0.1
Fruit	0.2	0.14	0.19	0.3	0.21	0.16	0.18	0.21
Sugar	0.01	0.01	0	0.04	0.02	0.01	0.01	0.01

<i>Income Elasticities</i>	Average	1st Quintile	2nd Quintile	3th Quintile	4th Quintile	5th Quintile	Rural	Urban
Wheat	0.63	1.04	0.71	0.67	0.37	0.32	0.7	0.6
Maize	0.42	1.05	0.33	0.14	-0.1	0.07	0.51	0.34
Alcohol Tobacco	0.75	0.75	0.27	0.13	1.02	0.79	0.86	0.73
Meat	0.85	1.5	1.08	0.72	0.44	0.36	1.04	0.76
Legumes	0.39	0.93	0.04	0.24	0.04	0.26	0.46	0.38
Dairy	0.72	1.28	0.78	0.74	0.32	0.25	0.81	0.65
Oil Fats	0.57	1.15	0.27	0.48	0.23	0.39	0.67	0.56
Vegetables	0.57	1.05	0.54	0.32	0.21	0.33	0.64	0.54
Fruit	0.86	1.43	1.14	0.85	0.63	0.43	1.03	0.79
Sugar	0.46	0.88	0.12	0.36	0.2	0.44	0.55	0.48

Note: Estimates shown in boldface are greater in absolute value than twice their bootstrapped standard error. (5% significance level)

Source: Author's calculation using INEGI households surveys 1984-2000.

Table 3: Own-price elasticities

<i>Price Elasticities</i>	2nd						Rural	Urban
	Average	1st Quintile	Quintile	3th Quintile	4th Quintile	5th Quintile		
Wheat	-1.29	-1.61	-1.34	-1.26	-0.94	-0.87	-1.46	-1.18
Maize	-0.74	-0.86	-0.54	-0.61	-0.65	-0.25	-1.33	-0.79
Alcohol Tobacco	-0.91	-0.98	-0.91	-0.86	-0.93	-0.5	-0.9	-0.93
Meat	-1.16	-2.23	-1.05	-0.89	-0.8	-0.76	-1.26	-0.87
Legumes	-0.16	-0.74	-0.43	-0.08	-0.79	-1.72	-0.47	0.12
Dairy	-1.16	-1.19	-1.11	-1.31	-1.15	-1.1	-1.07	-1.28
Oil Fats	-0.89	-0.39	-1.26	-0.79	-1.27	-3.47	-1.26	-0.4
Vegetables	-0.91	-0.91	-1.16	-0.87	-0.85	-0.34	-1.1	-0.8
Fruit	-1.33	-2.46	-0.84	-0.99	-1.49	-0.96	-1.23	-1.27
Sugar	-1.79	-2.32	-2.04	-1.86	-1.31	-2.41	-0.62	-1.94

Note: Estimates shown in boldface are greater in absolute value than twice their bootstrapped standard error. (5% significance level)

Source: Author's calculation using INEGI households surveys 1984-2000.

Table 4: Cross-price elasticities (average household)

	Wheat	Maize	Alcohol Tobacco	Meat	Legumes	Dairy	Oil Fats	Vegetab.	Fruit	Sugar	Others
Wheat	.	0.05	0.02	-0.03	-0.39	0.40	0.13	-0.04	-0.17	-0.01	0.08
Maize	0.04	.	0.11	1.25	-0.02	0.33	0.22	-0.37	-0.15	-0.15	1.29
Alcohol Tobacco	0.05	0.51	.	0.05	-0.26	0.14	-0.08	0.06	0.13	-0.02	-0.70
Meat	-0.01	0.55	0.01	.	0.35	0.36	0.16	0.12	0.18	-0.09	0.41
Legumes	-0.51	-0.04	-0.10	1.58	.	0.05	-0.10	0.35	0.03	-0.13	-1.24
Dairy	0.19	0.21	0.02	0.58	0.01	.	0.19	-0.13	0.00	0.23	-1.14
Oil Fats	0.28	0.61	-0.04	1.11	-0.16	0.81	.	-1.22	-0.44	-0.24	-0.07
Vegetables	-0.03	-0.41	0.02	0.33	0.20	-0.21	-0.47	.	-0.14	-0.36	1.44
Fruit	-0.36	-0.45	0.08	1.20	0.03	-0.01	-0.44	-0.37	.	0.35	-0.09
Sugar	-0.02	-0.63	-0.01	-0.91	-0.30	1.51	-0.37	-1.43	0.54	.	4.02
Others	0.00	-0.02	0.00	-0.02	0.02	0.05	0.00	-0.02	0.00	-0.02	.

Note: The rows shows commodity being affected and the columns the commodity whose price changes.

Estimates shown in boldface are greater in absolute value than twice their bootstrapped standard error (5% significance level).

Source: Author's calculation using INEGI households surveys 1984-2000.

Table 5: Micro simulation: marginal tax reform

Product Group	Tax Equivalent	$\varepsilon = 0$		$\varepsilon = 0.5$		$\varepsilon = 1$		$\varepsilon = 2$	
		w^ε/\tilde{w}	λ	w^ε/\tilde{w}	λ	w^ε/\tilde{w}	λ	w^ε/\tilde{w}	λ
Wheat	0.04	1.00	1.04	1.19	1.24	1.34	1.39	1.37	1.42
Maize	0.13	1.00	2.04	1.40	2.91	1.84	3.77	2.25	4.60
Alcohol Tobacco	0.40	1.00	2.35	1.03	2.43	1.07	2.52	1.38	3.24
Meat	0.25	1.00	1.16	1.13	1.32	1.18	1.37	0.94	1.09
Legumes	0.20	1.00	0.79	1.52	1.21	2.17	1.72	3.27	2.60
Dairy	0.30	1.00	1.23	1.15	1.42	1.23	1.52	1.08	1.33
Oil Fats	0.06	1.00	0.49	1.37	0.67	1.78	0.87	2.43	1.19
Vegetables	0.00	1.00	1.18	1.24	1.47	1.46	1.72	1.74	2.05
Fruit	0.20	1.00	1.04	1.03	1.07	0.99	1.03	0.70	0.73
Sugar	0.40	1.00	0.51	1.45	0.75	2.02	1.04	3.26	1.68
Other Goods	0.00	1.00	0.99	0.93	0.92	0.87	0.86	0.83	0.83

Note: The table reports relative budget share and cost benefit ratios calculated for four different levels of the inequality aversion parameter.

Source: Author's calculation using INEGI households surveys 1984-2000.

Appendix A: Marginal Tariff reform and the poor.

The purpose of this appendix is to illustrate how tax reform (small departures from the existing tax structure) affects household welfare. In doing so, it follows Deaton (1997).

In the simplest form, social welfare can be written as:

$$W = V(u_1, u_2, \dots, u_N) \quad (\text{A.1})$$

where N is the number of people (households) in the economy for each of which the welfare levels are given by the indirect utility function:

$$u_h = \psi(x_h, p) \quad (\text{A.2})$$

where x_h is total expenditures and p is a vector of prices.

The effect of the tax reform on social welfare W through the movement in the prices of goods can be obtained differentiating A.1 with respect to the tax change (with the chain rule):

$$\frac{\partial W}{\partial t_i} = \sum_{h=1}^H \frac{\partial V}{\partial u_h} \frac{\partial u_h}{\partial p_i} \quad (\text{A.3})$$

which can be written using Roy's identity ($\partial u_h / \partial p_i / \partial u_h / \partial x_h = -q_{ih}$) as:

$$\frac{\partial W}{\partial t_i} = -\sum_{h=1}^H \eta_h q_{ih} \quad (\text{A.4})$$

Equation A.4 represents the social cost of the tax increase (represented by the increase in the cost of the expenditure baskets of the household), where η_h is the social marginal utility of money in the hands of household h :

$$\eta_h = \frac{\partial V}{\partial u_i} \frac{\partial \psi_h}{\partial x_h} = \frac{\partial W}{\partial x_i} \quad (\text{A.5})$$

Besides by changes in the cost of the expenditure basket, social welfare is determined also by government public expenditures, which in turn, are a function of taxes. Government revenues are determined by the sum over all goods of the tax payments and subsidy costs. Hence:

$$R = \sum_{i=1}^M \sum_{h=1}^N t_i q_{ih} \quad (\text{A.6})$$

where M is the set of goods i . The effect of a small change in the tax (or subsidy) will have an effect on the government revenue R and thereafter on the level of individual welfare.

Taking the derivative of the revenue with respect of the tax change leads to:

$$\frac{\partial R}{\partial t_i} = \sum_{h=1}^N q_{ih} + \sum_{h=1}^N \sum_{i=1}^M t_i \frac{\partial q_{ih}}{\partial p_i} \quad (\text{A.7})$$

which represents the outcome of the tax reform through its retributive effects.

Equation A.7 can be further simplified. Assuming that consumer prices are determined by world prices (p_i^0) and tax or subsidies (t_i)²⁹ $p_i = p_i^0 + t_i$ the household' budget constrain can be written as:

$$x_h = \sum_{i=1}^M (p_i^0 + t_i) q_{hi} \quad (\text{A.8})$$

where q_{hi} is quantity consumed.

Since the total expenditure of each household is unaffected by the tax increase, differentiating A.6 with respect to tax and holding world prices constant gives:

$$\sum_{h=1}^H q_{ih} + \sum_{h=1}^H \sum_{i=1}^M t_i \frac{\partial q_{ih}}{\partial p_i} = - \sum_{i=1}^M p_i^0 \frac{\partial q_{hi}}{\partial p_i} \quad (\text{A.9})$$

In this setup, the two equations A.4 and A.9 represent the social benefits and social costs of a tax increase. The costs are born by any individual that purchases the goods, and the benefit is represented by the additional government revenue.

A measure of the effects of the tax reform can be given by the ratio of (negative) costs to benefits which is:

$$\lambda_i = \frac{\text{Costs}}{\text{Benefits}} = \frac{\sum_{h=1}^H \eta_h q_{ih}}{\sum_{h=1}^H q_{ih} + \sum_{h=1}^H \sum_{j=1}^M t_j \frac{\partial q_{jh}}{\partial p_i}} \quad (\text{A.10})$$

Equation A.10 represents the social cost of raising one unit of revenue by increasing the tax on good i . Hence, if λ_i is large, social welfare would be improved by reducing the tax (either because the tax is taxing those with higher weights in the social welfare function,

²⁹ As in the case of a small economy.

or because the tax is distortionary, or both). On the contrary, good with low λ_i are those that are candidates for an increase in tax. When all the ratios are the same, taxes are optimally set and there is no scope for beneficial reform.

In practice, when calculating the denominator of λ_i (the marginal social cost of revenue from the tax on each of the products) one needs to take into account both quality and quantity responses to price changes. This has been performed by Deaton (1997) for which λ_i becomes:

$$\lambda_i = \frac{w_i^\varepsilon / \tilde{w}_i}{1 + \frac{t_i}{1+t_i} \left(\frac{\theta_{ii}}{\tilde{w}_i} - 1 \right) + \sum_{k \neq i} \frac{t_k}{1+t_k} \left(\frac{\theta_{ki}}{\tilde{w}_i} \right)} \quad (\text{A.11})$$

where the numerator is a pure distributional measure of the good i which can be adjusted by modifying the “inequality aversion” parameter ε , therefore giving more weight to the poorer households (when $\varepsilon = 0$ there are not distributional concerns). In this case the outcome will be optimal in equity rather than in efficiency. The aversion parameter can also be interpreted as the relative shares of the market-representative individual and the socially-representative individual whose income is lower the higher the inequality aversion parameter ε . The two terms in the numerator are defined as:

$$w_i^\varepsilon = \left(\sum_{h=1}^H \left(\frac{x_h}{n_h} \right)^{-\varepsilon} x_h w_{ih} \right) / \sum_{h=1}^H x_h \quad \text{and} \quad \tilde{w}_i = \left(\sum_{h=1}^H x_h w_{ih} \right) / \sum_{h=1}^H x_h \quad (\text{A.12})$$

The denominator of equation A.11 is the tax factor multiplied by the income elasticity of good i with respect to its price, quality and quantity effects taken together.³⁰ If this term is large (and negative) then it would be costly to further increase revenue through taxes in this product. Finally, the last term is the cross-price effect which captures the effect on other goods due to the change in tax on good i .

³⁰ Therefore the first term of the denominator measures the own-price distortionary effect of the tax.

Equation A.11 derives from the fact that one can choose to treat individuals equally or to give a particular group of individuals more weight. For example, if more weight is put on poor individuals, it is possible to obtain the social cost of the tax with respect to the poor. Similar exercises can be done with respect to geographic location or ethnic group.

Appendix B: Data appendix

The estimates are based on the series of Mexican National Household Income and Expenditure Surveys (ENIGH) collected by the Insituto Nacional de Estadistica, Geografia e Informatica (INEGI) between 1984 and 2000. Data was collected for the years 1984, 1989, 1992, 1994, 1996, 1998, and 2000. All of the surveys are representative at the national level and were similarly constructed. The design survey is stratified at the state level. The analysis looks specifically at ten categories of food products which are aggregates over varieties. The concordance between the household expenditures and the product categories follows the one developed by Ianchovicina, Nicita and Soloaga (2002).

The application to Mexico relies on seven household surveys that categorize the households into 32 different geographical locations corresponding to the federal states of Mexico. Assuming that price differences within each federal state and survey year are due only to quality issues, measurement error and inflation, the households are classified into clusters according to their geographical location and to the year of survey.³¹ Therefore, the households are categorized into 224 clusters with an average of 329 observations per cluster.

Household questionnaires ask respondents to report the physical quantity actually purchased along with the expenditures. The ratio of expenditure and quantity determines the unit value that, with some caution, can be interpreted as a measure of price. Figure A.1 reports the value of the logarithm of unit value for the ten goods from 1984 to 2000. The graphics show a downward trend in the logarithm of the unit values of meat

³¹ Only the 2000 household survey classifies the households in a more disaggregated geographical location.

products, oil and fats, vegetables and dairy products.³² Unit values for fruits, sugar and wheat are constant, while maize-based products appear to be the only product that experienced an increase in the unit value.³³

In order to be suitable for elasticity estimates, the data must show a certain degree of variability across spatial locations and time. Even if unit values are not the same as prices because of the quality issue³⁴ and measurement errors³⁵, from inspection of figure 1 and 2 it is sufficiently clear that the unit values, and likely prices, for most of the products show a substantial variation across cluster.

Figure A.2 shows the logarithm of unit values of the ten product groups across the 32 federal districts. Sugar, meat, oil and fats seem to have the most stable unit value across districts. Meanwhile, there is a higher variance in the unit value for alcohol and tobacco, dairy products, fruit, legumes, maize, vegetables and wheat. Overall, there seems to be enough spatial and time variation to support the methodology undertaken in this paper.

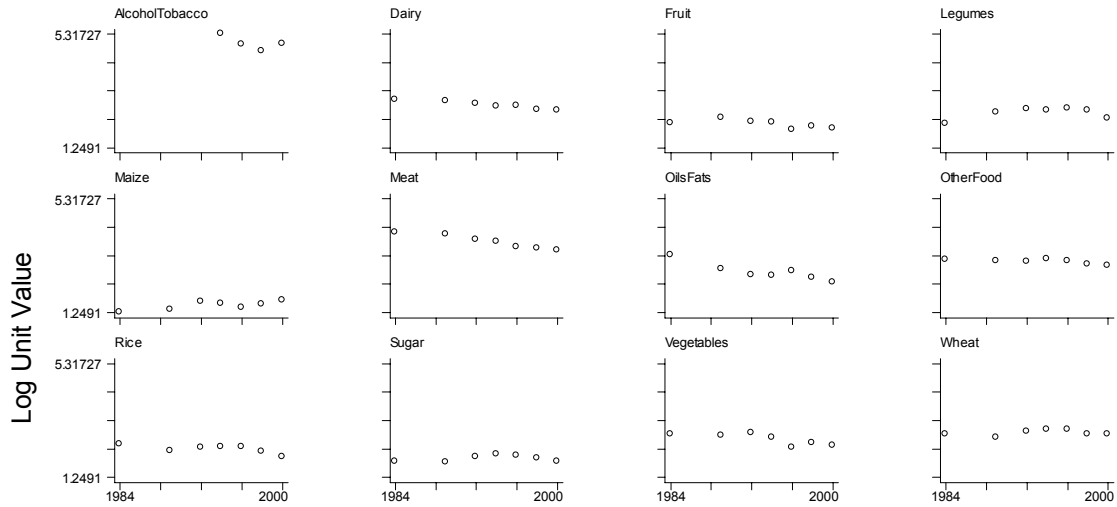
³² The unit values are adjusted by the inflation rate of food products and are reported in year 2000 terms.

³³ The unit values of alcoholic beverages and tobacco products are calculated only starting from the year 1994 because of a different classification in the first three household surveys.

³⁴ Two unit values may result different even if the underlining price is the same because the goods purchased are of different quality.

³⁵ Data on household surveys is usually subject to measurement error. This may be due to the construction of the questionnaire, processing of the data, misreporting and survey design.

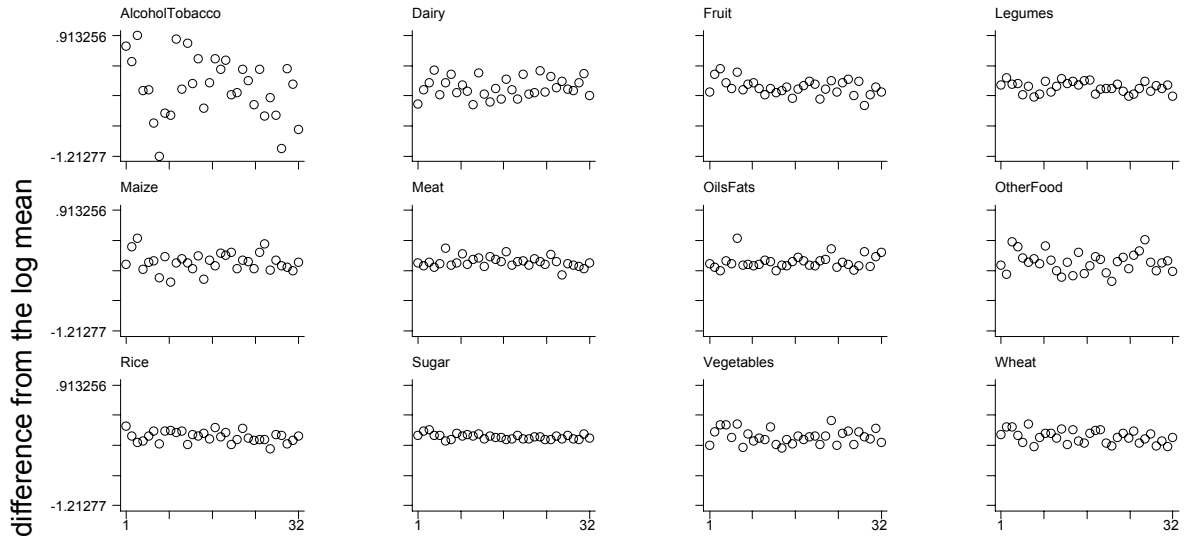
Figure A.1: Trend of log of unit value 1984-2000



Year
Graphs by good

Source: Author's calculation using INEGI households surveys 1984-2000.

Figure A.2: Average log unit value by district



Federal District Graphs by good

Source: Author's calculation using INEGI households surveys 1984-2000.