

**UNIVERSITY OF WAIKATO**

**Hamilton  
New Zealand**

**Improved Modelling of Spatial Cost of Living Differences  
in Developing Countries:  
A Comparison of Expert Knowledge and Traditional Price Surveys**

**John Gibson and Trinh Le**

**Working Paper in Economics 8/18**

June 2018

*Corresponding Author*

**John Gibson**

School of Accounting, Finance  
and Economics

University of Waikato

Private Bag 3105

Hamilton

New Zealand, 3240

Tel: +64 (7) 838 4289

Email: [jkgibson@waikato.ac.nz](mailto:jkgibson@waikato.ac.nz)

**Trinh Le**

Motu Economic and Public Policy Research

Wellington

New Zealand

Email: [trinh.le@motu.org.nz](mailto:trinh.le@motu.org.nz)

### **Abstract**

Most developing countries lack spatially disaggregated price data, despite the importance of spatial transactions costs in these settings. We experimented in Vietnam with a new way of obtaining disaggregated price data, using local expert knowledge to derive the mean and variance for prices of 64 items in over 1000 communities. We use these prices to calculate regional cost-of-living indexes. These provide a better approximation to benchmark multilateral price indexes calculated from traditional market price surveys than do two no-price methods, based on using food Engel curves to derive deflators and based on unit values (survey group expenditure over group quantity).

### **Keywords**

expert knowledge  
inequality  
prices  
regional cost-of-living

### **JEL Codes**

D12, E31, O15

### **Acknowledgments**

We are grateful for assistance from Valerie Kozel, Nguyen Tam Giang, Bonggeun Kim and Geua Boe-Gibson, along with many staff from the General Statistics Office of Vietnam.

All remaining errors are those of the authors.

## I. Introduction

Reliable data on differences in real welfare across space are rare in developing countries. These are places for which it is implausible to assume that prices are the same everywhere, with high internal transport costs and retail chains that set prices on a national basis yet to emerge. Statistical agencies mostly focus on the temporal Consumer Price Index (CPI), which lets one compare changes in, but not levels of, prices over space. For example, China, India and Indonesia do not collect market price data in rural areas to match to their household income and expenditure surveys (Gibson and Rozelle, 2005). While urban prices are collected for the CPI, these are a poor guide to prices prevailing in the countryside (Deaton and Dupriez, 2011). Even countries that ostensibly have a national CPI, such as Vietnam, gather prices at convenient locations that are more accessible than the markets where the average rural household shops and so misstate the typical experience of relative price changes (Winters et al, 2004). Moreover, because the CPI focuses on temporal consistency, the item specifications used by statistical offices in different areas of a country may vary, limiting use of these prices for making consistent spatial comparisons (Gibson *et al.* 2017).

The aversion to collecting rural price data is not surprising. It can be hard for outsiders to find, understand and study markets in many poor rural areas. Markets may meet intermittently, at different places on different days, and often at early hours. Perhaps because it already is logistically difficult to manage the traditional part of the data collection effort (household expenditures), survey agencies are reluctant to add another part of the survey (for collecting prices) with its own set of complications that may cause a decline in overall survey quality. The problems are likely to be most apparent in places with poor infrastructure and low population densities, which are exactly where prices will vary most over space and where, therefore, nominal income or expenditure data are the least useful for revealing spatial differences in real living standards.

The lack of spatial prices forces some analysts to use a food Engel curve method, to derive the deflator that gives nominal incomes in different regions the same real standard of living (based on having the same food share). This adapts a method from Hamilton (2001), for temporal comparisons, where Engel curves are used to back out the implied true price index and real income growth over time. For example, Gong and Meng (2008) use Engel curves to examine regional price differences for urban households in China. Almås *et al.* (2018) use a similar approach to calculate state-level deflators for India, which imply much greater spatial variation in poverty rates than what official data show. However, a comparison with multilateral price indexes shows that time-space deflators from food Engel curves may be quite distorted (Gibson *et al.* 2017).

Another no-price method relies on unit values (expenditures on a survey group divided by the group quantities). These are obtained as a byproduct of expenditure surveys, for some foods (those whose metric quantities are asked about) and perhaps for fuels. Unit values are often used to calculate the regional costs of a food poverty line, which is a Laspeyres-type price index. There are several problems with unit values; they are available only for purchasers, they reflect reporting errors in quantities and expenditures, and they will vary with the quality choices that households make over the items within a survey group (Deaton, 1989). In particular, a key concern for understanding regional differences is that unit values should refer to a higher quality mix for a survey group the further one moves from the point of surplus production. The Alchian-Allen effect, *aka* ‘shipping the good apples out’ means that spatial (transport), temporal (storage) and other transactions costs alter the relative price of quality over time and space (Gibson and Kim, 2015). Therefore, regional averages of unit values will not refer to a constant quality mix within each survey group, and so will misrepresent spatial cost of living differences.

In light of these problems with no-price methods, and recognizing that statistics agencies find it hard to implement spatial cost-of-living surveys in rural areas of developing countries, we experimented in Vietnam with a new approach to obtaining disaggregated price data. We use the expert knowledge of local residents, who see local prices in their everyday market transactions. We developed a price questionnaire for the General Statistics Office (GSO) of Vietnam to field in a random sample of 1049 communes (one-eighth of all communes in Vietnam) in 2010.<sup>1</sup> Focus groups with the Women’s Union in each commune provided data on the lowest price the item sells for (a), the typical price (b), and the highest price (c), for 64 consumer items. To ensure that reports in all communes referred to the same quality for each item, photographs of the items were shown to these expert informants. With the data from (a), (b), and (c), we use a triangular distribution to estimate the mean and variance of the local prices for each of these items. A benchmark to assess the accuracy of the spatial price indexes based on these expert informant data comes from a market price survey carried out in stores and markets in the same communes at the same time. We also have unit values for 30 food groups, from the Vietnam Household Living Standards Survey (VHLSS) carried out in the same communes by another department of the GSO.

The idea to use expert informants is not new. In early stages of the World Bank’s Living Standards Measurement Study (LSMS) it was proposed to interview groups of housewives about prices rather than use a full market price survey (Saunders and Grootaert, 1980). Discussion within the LSMS program was critical of this ‘novel but risky’ idea (Wood and Knight, 1985) and it was not implemented. However, the idea was used in two other surveys, neither of which is on the scale of what we report here, in either spatial coverage or

---

<sup>1</sup> Vietnam’s communes are the lowest level administrative unit, averaging about 10,000 people or 2,500 households.

commodity detail. The Indonesia Family Life Survey (IFLS1 and 2) had opinions on local prices for 38 items obtained from key informants in local women's groups, located in the 320 enumeration areas (EAs) where the IFLS was first fielded. Later waves of the survey had one key informant per EA provide a different set of prices to those obtained from a market price survey but no comparisons between the two types of price data have been reported. In the Papua New Guinea (PNG) Household Survey, respondents were shown photographs of representative items from ten food and tobacco groups and asked to report the local prices. The same households did a consumption recall survey, which provided unit values. Comparisons across the 118 EAs where the survey was fielded suggested that picture-aided price opinions were more accurate proxies than were the unit values (Gibson and Rozelle, 2005).

The results for Vietnam suggest that price data from local experts better approximate a benchmark multilateral price index calculated from traditional market price surveys than do two no-price methods, based on using food Engel curves to derive deflators and based on using unit values. For example, the mean correlation between prices and unit values over the 1049 communes is just 0.33, yet the corresponding correlation for the prices from local experts is 0.73 (and the gap for the median food group is even larger). Spatial price indexes from the country-product-dummy method that use prices provided by the local experts are much closer to the benchmark index from market price surveys than are regional deflators derived from Engel curves. We use the sum of squared differences (SSD) to summarize discrepancies from the benchmark index; the SSD is about 14 for indexes using the data from the local experts but is about 400 times larger for spatial deflators from Engel curves (with SSDs ranging from 4500 to 5800). The Engel curves especially overstate the cost-of-living in some of the poorest rural areas and understate in some rich urban areas. Thus, Engel curve deflators make spatial inequality look worse than it actually is, raising the Gini index for real per capita expenditures to 0.48 compared with a benchmark Gini index of 0.41 using the deflators from either the market price surveys or from the local experts.

The rest of the paper is structured as follows. Section II describes the survey experiment, while methods of using the price data and calculating the spatial deflators are set out in Section III. The comparative performance of Engel curve deflators versus spatial deflators that use data from expert knowledge is reported in Section IV. In Section V we introduce a key spatial feature of prices, the Alchian-Allen effect, and show that the prices obtained from local experts illustrate this effect about as well as do the prices from the traditional price survey. This effect matters because it undermines use of unit values as a proxy for prices. We then give food group level comparisons between the benchmark prices, the expert knowledge prices, and the unit values, and report food price indexes in Section VI. The discussion and conclusions are in Section VII.

## II. The Price Survey Experiment

In 2010 we designed a spatial price survey for 64 items, to be fielded by the Prices Department of the General Statistics Office (GSO) of Vietnam. A key concern was to maintain consistency of item specification across areas; prior spatial deflators for the biennial poverty line calculations attempted to use prices collected for the CPI but faced the problem that local statistics offices in various provinces were using different item specifications (for example, Hanoi brand beer in the north and Saigon brand beer in the south). Concerns about the regional poverty profile derived from these CPI-based deflators induced the World Bank to fund this new spatial cost of living survey. In order to ensure consistency across space, when the surveyors went to stores and markets they used detailed photographs of each of the 64 target specifications. These same pictures were used in the focus groups with the expert informants. Therefore, this design is a hybrid of what was done previously in PNG, with photographs shown to individual households, and in Indonesia with price opinions obtained at community level from women's groups, without the aid of photographs.

Figure 1 presents examples of these photographs for four items: two are specifications for the fresh fish and shrimp group, for tiger shrimp of 7-10 cm length and shrimp of 3-5 cm length. The price survey and the expert opinions are for price per kilogram, but the size range shown (and use of a matchbox in the picture as a scale indicator) was to help the surveyors and the expert informants report price for a particular grade of what is a heterogeneous product. For major components of household budgets, such as rice, pork, beef, chicken, fish and shrimp, oils and fats, and outdoor meals, the price survey included multiple specifications and these have previously been used to show how price of a higher quality item relative to a lower quality item within a food group varies over space, that is, the Alchian-Allen effect (Gibson and Kim, 2015). The other two pictures are for two types of outdoor street meals; a typical breakfast meal (beef noodle soup) and a typical lunch or dinner meal (rice, pork, tofu, and vegetables). For the street meals, the price was for the plate as displayed in the picture.<sup>2</sup> Most of the items in the survey were for foods, because of the focus on repricing regional food poverty lines, but the survey also covered major non-foods and also basic services, such as haircuts, puncture repairs, tailoring, and local school fees.

The list of the 64 items in the price survey is given in Appendix Table 1. We also report how many communes (of 1049 in total) have price data for each item. For the market price survey, price readings were required from two vendors for the specified item, and a third reading was made for the (next) most popular item in the market. The price of this extra specification allowed us to impute prices for the target specification in communes where the

---

<sup>2</sup> Given the difficulty of quantifying street meals (for example, weighing is of little help because of their diverse ingredients) the use of pictures provides an advantage, because a spatial price index for street meals should be anchored on the same items (the pictured meals) which is not guaranteed with traditional surveys that do not quantify street meals.

target item was unavailable. We obtained 54,200 observations on local prices from the expert informants, and 54,600 observations on the prices of the target items, using traditional price surveys. Thus, using local experts provided 99.3 percent of the data obtained with the traditional approach. If all 64 items were found in each commune we would have 67,140 observations, so the price surveys obtained 81.3 percent of this target. For the market price survey (but not for the expert opinions), a further 7,860 observations were imputed (from regressions using the third reading on the non-target item to predict price of the target item).<sup>3</sup> The country-product-dummy method we outline in Section III can deal with the missing prices for the remaining seven percent of item-commune combinations.

**Figure 1: Examples of the Photographs Used in the Price Surveys**

Tiger shrimp (7-10cm long)



Shrimp (3-5 cm long)



Breakfast (beef noodle soup)



Lunch/dinner (rice, pork, tofu & vegetables)



<sup>3</sup> The spatial deflators we report below are robust to including or excluding the imputed prices.

The commodity weights for forming spatial price indexes, and the food budget shares and covariates for estimating the Engel curves come from the 2010 VHLSS. This multi-topic living standards survey was fielded in the same communes (and in others without the price survey) at the same time, but by a different GSO department. This timing overlap ensures that budget shares needed to calculate price indexes relate to the same period as the prices. Moreover, the food unit values from the VHLSS should also refer to prices from the same period as what the price survey observed.<sup>4</sup> The consumption module of the VHLSS is applied to three households per commune (a larger sample get an income-only questionnaire) and uses a 30-day recall of purchases and consumption from own-production and gifts for 54 food and drink groups, a 30-day recall for 28 frequently purchased non-food items and an annual recall for 36 other items. The mapping of the price survey items to the expenditure groups in the VHLSS that provide the budget shares needed for the spatial price indexes is reported in Gibson *et al.* (2017).

### III. Methods

Our survey of local expert knowledge provided us with data on (a) the lowest price that the item sells for locally, (b) the typical price and (c) the highest price. With these three values we estimate a triangular distribution, which has a mean:

$$\mu = \frac{a + b + c}{3}$$

and the variance of the local price for each of the 64 items is:

$$\sigma^2 = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18}.$$

The inverse of the variance can be used to weight the mean of the commune-level prices obtained from the Women’s Union informants when these are aggregated to regional averages. In other words, more weight goes on the commune-level reports that have less dispersion.<sup>5</sup>

In order to derive spatial price indexes from these price data, we use two of the three broad approaches to empirically approximating the true cost of living index (COLI), which is

---

<sup>4</sup> The 2010 VHLSS introduced a fixed recall period (the last 30 days) for food expenditures and quantities which aids the comparison of unit values and contemporaneous prices. In prior years, the VHLSS asked about a notional ‘usual month’ that does not correspond to any particular month of the year when prices might have been observed. Thus, the comparisons of spatial deflators based on prices with those based on unit values that are reported in this paper would not have been possible for Vietnam prior to 2010.

<sup>5</sup> For all but two items (namely, bricks and men’s haircuts) the weights were small (a median value of  $1.8 \times 10^{-5}$ ) and so the inverse-variance weighted average prices provides just a modest gain in performance, in terms of matching with price indexes from the market price surveys.



the ratio of minimum expenditure at alternative prices to minimum expenditure at base prices holding the standard of living constant (Dumagan and Mount, 1997; Breur and von der Lippe, 2011). The first approach is to use a price index with known biases, such as the Laspeyres, that gives an upper bound to the COLI because it ignores consumer substitution in response to relative price changes. The second approach is to use a superlative index formula, such as the Törnqvist, which is closer to the true COLI (due to less substitution bias) if preferences are homothetic, but which has an income bias if preferences are not homothetic. Proponents of the Engel curve approach to spatial deflation, such as Almås *et al.* (2018), point out that evidence of falling food shares as incomes rise suggests that preferences are non-homothetic and so this income bias issue is potentially relevant. The third approach, and the one that we do not follow in this paper, is to econometrically estimate demand equations for a set of goods, from which the theoretical expenditure functions that are the numerator and denominator of the COLI can be derived.

We use the weighted country-product-dummy (WCPD) method to form spatial deflators because this approach allows both fixed-weight and variable-weight price indexes to be calculated that are analogous to Laspeyres and Törnqvist indexes (Gibson *et al.* 2017). Thus, we bypass the issue of whether spatial price deflators should be based on homothetic or non-homothetic indexes by using both types. We also are guided by the sort of price indexes that a statistics office in a developing country would likely use if they had disaggregated spatial price data available. These offices are familiar with fixed-weight indexes, like Laspeyres, for their temporal deflation, and this type of index avoids the income bias if preferences are non-homothetic. However, for spatial deflation a statistics office may want a variable-weight superlative index, like a Törnqvist, because substitution bias is likely a bigger concern over space than over time because relative prices do not vary much over the short to medium term (Van Veelen and Van der Weide, 2008).

### **Weighted Country Product Dummy (WCPD) Method**

The WCPD method provides multilateral price indexes, which allow simultaneous comparison of multiple regions (or time periods), and can handle substitution effects, democratic weights, and reversibility. These features matter to spatial comparisons, given (i) there is no natural base, unlike for temporal comparisons; (ii) the cost of living of the typical person in a region (rather than of the typical dollar, as in a plutocratic inflation index) is usually of interest; and, (iii) the likely greater importance of substitution bias over space compared to over time. We also note that Deaton *et al.* (2004) recommend the WCPD method (along with the Eltetö, Köves and Szulc or EKS method) for use as a deflator when measuring household living standards. A further feature of the WCPD method is that with appropriate choice of expenditure or quantity weights one can derive several bilateral price indexes, including those of Dutot, Jevons, Törnqvist, and Walsh (Diewert, 2005), and also a multilateral system that is an expenditure-share weighted geometric form of the Geary-Khamis index widely used for international purchasing power parity comparisons (Rao, 2005).

The WCPD works as follows: for  $J$  regions,  $K$  goods, and  $T$  periods the relationship between the prices of goods in different regions and periods is assumed to follow:<sup>6</sup>

$$P_{k,j,t} = \rho_{j,t} \eta_k u_{k,j,t} \quad (1)$$

where  $\rho_{j,t}$  is the price level in region  $j$  and period  $t$  relative to the base region/period,  $\eta_k$  is the price level of good  $k$  relative to the base good, and  $u_{k,j,t}$  is a random disturbance term. The price parameters ( $\rho_{j,t}$  and  $\eta_k$ ) in equation (1) can be directly estimated in a log-linear regression model, using the  $K \times J \times T$  prices from a spatially disaggregated price survey:

$$\sqrt{w_{k,j,t}} \ln p_{k,j,t} = \hat{\phi} + \sum_{j=1}^J \ln \rho_{j,0} \sqrt{w_{k,j,0}} D_{j,0} + \sum_{t=1}^T \sum_{j=0}^J \ln \rho_{j,t} \sqrt{w_{k,j,t}} D_{j,t} + \sum_{k=1}^K \eta_k \sqrt{w_{k,j,t}} D_k + u_{k,j,t} \quad (2)$$

where the weight  $w_{k,j,t}$  for good  $k$  in region  $j$  and period  $t$  is described below,  $D_{j,t}$  is a dummy variable for region  $j$  and period  $t$ ,  $D_k$  is a dummy for good  $k$  and  $\hat{\phi}$  is the intercept plus the coefficient for the omitted base category dummies.

Our first benchmark price index has variable weights:  $(s_{kj,t} + s_{k0,1})/2$  where  $s_{kj,t}$  is the average budget share of item  $k$ , in region  $j$ , and time  $t$ , and  $s_{k0,1}$  is the average budget share for item  $k$  in the base period/region (the urban sector of the Red River region, that includes Hanoi). We refer to this price index as WCPD-vw (for variable-weight), which gives estimated deflators:

$$\rho_{j,t} = \left[ \sum_{k=1}^K \left( \frac{s_{kj,t} + s_{k0,1}}{2} \right) \ln \left( \frac{p_{kj,t}}{p_{k0,1}} \right) \right] \quad (3a)$$

The WCPD-vw allows for substitution because it uses budget shares from both the base region and the current region (or period), but it exactly measures the cost of living only for homothetic preferences. Therefore we also use a fixed-weight index that does not rely on homothetic preferences, but is subject to substitution bias, by using  $s_{k0,1}$  as the weight for all periods and regions. The time-space deflators for the WCPD-fw (for fixed-weight) index are:

$$\rho_{j,t} = \left[ \sum_{k=1}^K (s_{k0,1}) \ln \left( \frac{p_{kj,t}}{p_{k0,1}} \right) \right] \quad (3b)$$

---

<sup>6</sup> We include the time index for generality. In the application below the data were collected within a sufficiently short period (August and September 2010) that we ignore any temporal effects.

Intuitively, WCPD-fw is a Laspeyres-like index but it is not exact.<sup>7</sup> However, the deflator in equation (3b) does not depend on homothetic preferences (so there is no income bias) and is like the sort of price index that a statistics office may use if they had disaggregated spatial price data.

### Engel Curve Method

In the original formulation of Hamilton (2001), for assessing bias in temporal deflators, the budget share of food at home for household  $i$  in region  $j$  and time period  $t$ ,  $w_{i,j,t}$  is treated as a linear function of the logarithm of real household income, a relative price term and control variables:

$$w_{i,j,t} = \phi + \gamma(\ln P_{F,j,t} - \ln P_{N,j,t}) + \beta(\ln Y_{i,j,t} - \ln P_{j,t}) + \mathbf{X}'\theta + u_{i,j,t} \quad (4)$$

where  $P_{F,j,t}$ ,  $P_{N,j,t}$ , and  $P_{j,t}$  are the true but unobserved prices of food, non-food, and all goods,  $Y$  is total expenditure (a permanent income proxy), the  $\mathbf{X}$  are control variables and  $u$  the disturbance. A set of temporal dummy variables are then added to the specification in equation (4), to look for ‘drift’ in the Engel curve, after all incomes have supposedly been put on a common temporal basis by using the CPI to deflate them. Under certain assumptions that are discussed by Hamilton (2001), and by the many subsequent papers that apply the method to temporal data, the coefficients on the temporal dummy variables (scaled by the coefficient on income) indicate the CPI-bias.

If this method is adapted to space-time deflation (we include the time dimension for generality, even though we ignore time in the application below), the estimating equation becomes:

$$w_{i,a,j,t} = \hat{\phi} + \gamma(\ln P_{F,a,j,t}^* - \ln P_{N,a,j,t}^*) + \beta \ln Y_{i,a,j,t} + X'\theta + \sum_{j=1}^J \delta_{j,0} D_{j,0} + \sum_{t=1}^T \sum_{j=0}^J \delta_{j,t} D_{j,t} + u_{i,a,j,t} \quad (5)$$

with the extra subscript for sub-regional area  $a$ , the starred terms are nominal price indexes for food and non-food,  $D_{j,t}$  is a time-space dummy set to 1 for region  $j$  and period  $t$ , and the intercept  $\hat{\phi}$  includes the coefficient on a single omitted dummy,  $D_{0,0}$ .<sup>8</sup> The coefficients on the time-space dummy variables are the key parameters; by showing how food shares vary over time periods and across regions for households with the same nominal income, one can derive time-space deflators for real incomes. Under assumptions about the role of relative price effects

<sup>7</sup> Selvanathan (1991) shows how an appropriately weighted linear regression lets one calculate a Laspeyres index from the WCPD method; the difference is that our WCPD models are log-linear.

<sup>8</sup> The sub-regional area,  $a$  is needed because, otherwise, the relative price effect is identified from the same regional and temporal variation as the time-space dummies and perfect collinearity will result.

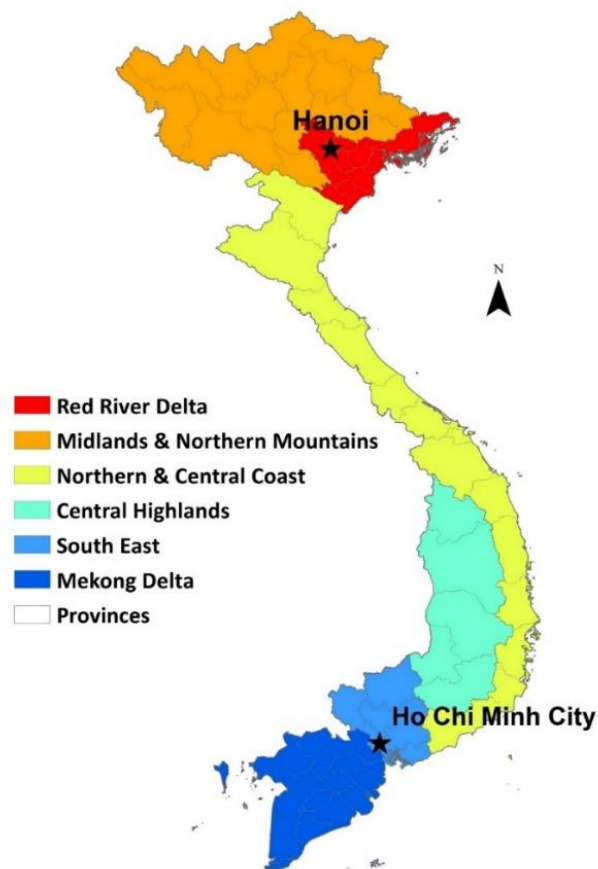
(discussed more fully in Gibson et al, 2017), and with the main identifying assumption that the food share can indicate welfare and that nothing relevant omitted from the Engel curve equation is correlated with the dummies, the index for the price level in region  $j$  and time period  $t$  is calculated as:

$$P_{j,t} = \exp(-\delta_{j,t}/\beta). \quad (6)$$

#### IV. Comparisons in terms of Spatial Price Indexes

The spatial deflators are estimated for Vietnam's six broad regions (see Figure 2), with the cost of living allowed to vary between urban and rural sectors within regions. The estimates that we report are for an expenditure aggregate (and a food share) that excludes housing and durable goods. While these two budget components are important, having a combined budget share of almost one-fifth, there was only a single durable good in the price survey (a Samsung 21 inch television) and it was considered too difficult to consistently survey prices for housing. In unreported results we do allow for regional differences in housing costs, based on hedonic regressions of self-reported dwelling values (values are used because almost no rents are reported in VHLSS), and the results are very similar to what we report for the non-durables, non-housing expenditure aggregate.

**Figure 2: The Six Regions of Vietnam Used for Calculating Spatial Price Indexes**



In Table 1 we report the spatial deflators using the WCPD-vw approach, with the results that use the market price survey in column (1) and those that use the opinions about prices from the local experts in column (2). According to the benchmark data from the market price surveys, the rural Mekong Delta has the lowest cost of living index, at 85.4, while the urban Red River (which includes Hanoi) has the highest cost of living, with an index value of 100. In all regions except the Red River and the South East (which includes Ho Chi Minh City) there are only small differences in the cost of living between urban and rural sectors. This reflects the fact that Hanoi and Ho Chi Minh City are much larger than the other urban areas, and so have more differentiation from their rural hinterland, than is the case for the other regions that only have smaller cities.

**Table 1: Spatial Deflators from Expert Informant Prices and from Engel Curves Compared to using Benchmark Price Surveys for the WCPD-vw Method**

	Market Price Survey (1)	Expert Informants (2)	Squared Differences [(1)-(2)] <sup>2</sup>	Engel Curve Deflator (4)	Squared Differences [(1)-(4)] <sup>2</sup>
Urban Red River	100.0	100.0	0.0	100.0	0.0
Urban Mid-Northern Mountains	97.3	96.4	0.8	113.0	246.7
Urban North-Central Coast	89.9	87.7	4.8	80.3	92.2
Urban Central Highlands	90.9	91.1	0.0	65.6	641.9
Urban South East	97.9	98.0	0.0	58.9	1523.0
Urban Mekong Delta	88.4	88.7	0.1	77.1	128.5
Rural Red River	90.3	89.7	0.3	108.7	339.2
Rural Mid-Northern Mountains	95.1	95.2	0.0	137.9	1831.2
Rural North-Central Coast	87.5	85.3	5.0	91.8	18.1
Rural Central Highlands	89.2	90.8	2.7	95.7	42.4
Rural South East	91.4	91.0	0.2	65.0	696.4
Rural Mekong Delta	85.4	85.0	0.2	100.0	213.8
<b>Sum of Squared Differences</b>			<b>14.1</b>		<b>5773.5</b>

*Notes*

The WCPD-vw method is based on equation (3a) in the text, and the Engel curve deflators are based on equation (6). The commune-level means of the expert informant prices are weighted by their inverse variances, based on the triangular distribution.

Similar spatial patterns appear when the price index is calculated using the data from the expert informants (with all other aspects of equation (3a) kept the same). The correlation between the index values in columns (1) and (2) is 0.98 and, using the price data from the local expert informants ranks the four region-sectors with the highest cost of living, and the two with the lowest cost of living, the same as when using the data from the traditional price survey. The squared differences between the two sets of index values are shown in column (3); these show that it is mainly the Central Coast region where the price opinions suggest a lower cost of living than what the market price survey shows. The overall sum of squared differences (SSD) is

14.1.<sup>9</sup> If we had not weighted the commune-level mean prices obtained from the expert informants by their inverse variance, the SSD would be somewhat higher, at 20.7, so asking about low, typical, and high prices and then using a triangular distribution provides some advantage over just asking about the typical local price (which would then not allow a variance to be calculated).

The spatial deflators that are derived from the food Engel curve are reported in column (4) and these provide a quite different picture of regional cost of living variation in Vietnam. With these deflators, it is the Mid-Northern Mountains region – and particularly the rural sector – that appears to have the highest cost of living. Poverty maps from small-area estimation methods show that poverty is increasingly concentrated in Vietnam’s Northern Mountains (World Bank, 2012) so it is surprising that prices and the cost of living would be so high in such a region (and would seem to be even higher if we included housing in the food shares).<sup>10</sup> Also surprising is the position of the rural Mekong Delta as having the same cost of living as the urban Red River region (which includes Hanoi and Hai Phong). The rural Mekong Delta is Vietnam’s rice bowl, with surplus rice moving out of this region to feed the rest of the country (as discussed in more detail in Section V, below), and trade normally moves goods from lower cost to higher cost regions.

The correlations between the benchmark price index in column (1) and the deflator derived from the Engel curve estimates in column (4) is just 0.13 (and the rank correlation is 0.05).<sup>11</sup> We would not reject independence of the column (4) values from the benchmark values (at  $p=0.68$ ). The squared deviations from the benchmark price index are reported in column (5) and these show that the two areas with the biggest discrepancy are the rural Mid-Northern Mountains and the urban South East. The Engel curve shows low cost of living in a nominally rich area (around Ho Chi Minh City) and high cost of living in a nominally poor area (the rural Northern Mountains). Thus, spatial inequality seems far higher – with a Gini coefficient for per capita expenditures of 0.48 if using the Engel curve deflator – while using

---

<sup>9</sup> If we used the full expenditure aggregate, including durables and housing, the SSD would be 12.0. The hedonic analysis of regional differences in dwelling values provides results that are shared by both the traditional price survey deflators in column (1) and the expert informant deflators in column (2) and so tends to make the SSD values smaller.

<sup>10</sup> With housing and durables included the Engel curve deflator for the rural Mid-Northern Mountains rises to 142. The Balassa-Samuelson effect would suggest that housing, and other non-traded goods, would be cheapest in the poorest areas, which is one reason why using spatial price deflation should result in lower spatial inequality.

<sup>11</sup> Results of the Engel curve regressions (and the WCPD regressions) are available from the authors. The Engel curve regressions include as covariates the log of household expenditure (excluding durables and housing), the log of the relative price of food, household size, four demographic ratios (for shares of children, youth, elderly and migrants in the household), the gender, age, sector of activity and education of the household head, and prices for two types of street meals, which are a close substitute for food at home (the numerator of the budget share).

the benchmark WCPD-vw price index from the market price surveys gives a Gini of just 0.41 (the same as using prices from the local experts). It may not surprise that this pattern also holds in India when Almås *et al.* (2018) use the Engel curve method – they find far more spatial variation in poverty than what official data show. Thus, greater regional inequality in deflators may be a feature (or a flaw) of the Engel curve method.<sup>12</sup>

In Table 2 we report the results when the benchmark spatial deflators use the WCPD-fw approach; this index should not be subject to any income bias if the preferences are non-homothetic. The fixed weight price index is a little higher than the variable weight index, averaging 92.6 outside of the base region compared to 91.2 for the variable weight index. This is to be expected, given that the fixed weight index ignores consumer substitution in response to relative price differences and so will overstate the cost of living outside of the base region. However, apart from this level effect, the regional patterns are very similar to those in Table 1, with a correlation of 0.996 for the values in column (1) of each table. Thus, despite the potential for income bias in a variable weight superlative index, which advocates of the Engel curve method suggest as a reason for not using conventional price indexes, at least for the situation in Vietnam it appears to not matter much.

**Table 2: Spatial Deflators From Expert Informant Prices and from Engel Curves Compared to using Benchmark Price Surveys for the WCPD-fw Method**

	Market Price Survey (1)	Expert Informants (2)	Squared Differences [(1)–(2)] <sup>2</sup>	Engel Curve Deflator (4)	Squared Differences [(1)–(4)] <sup>2</sup>
Urban Red River	100.0	100.0	0.0	100.0	0.0
Urban Mid-Northern Mountains	97.9	96.7	1.4	113.0	227.4
Urban North-Central Coast	90.0	87.8	4.7	80.3	93.8
Urban Central Highlands	91.1	91.3	0.0	65.6	649.1
Urban South East	98.3	98.5	0.0	58.9	1554.7
Urban Mekong Delta	88.7	88.8	0.0	77.1	133.4
Rural Red River	90.5	89.9	0.4	108.7	331.7
Rural Mid-Northern Mountains	96.1	96.4	0.1	137.9	1750.9
Rural North-Central Coast	87.6	85.5	4.2	91.8	17.7
Rural Central Highlands	89.0	90.8	3.0	95.7	44.6
Rural South East	91.4	91.1	0.1	65.0	697.3
Rural Mekong Delta	85.3	84.9	0.1	100.0	216.5
<b>Sum of Squared Differences</b>			<b>14.2</b>		<b>5717.1</b>

*Notes*

The WCPD-fw method is based on equation (3b) in the text, and the Engel curve deflators are based on equation (6). The commune-level means of the expert informant prices are weighted by their inverse variances, based on the triangular distribution.

<sup>12</sup> The Coefficient of Variation (CoV) of the Engel curve deflators in Table 1 is 0.25, while for the benchmark index using the traditional price survey (and for using the expert information prices) the CoV is only 0.05.

With the results in Table 2 close to those in Table 1, it is unsurprising that prices from the expert informants still provide data for price index calculations that closely approximate what the benchmark market price survey shows. The rank correlation between the regional deflators using the price opinions and those using the market surveys is 0.958 with both fixed weight and variable weight indexes, while the product-moment correlations are 0.975 and 0.976. The sum of squared deviations are likewise similar, at 14.2 for the fixed weight indexes and 14.1 for the variable weight indexes. The poor performance of the Engel curve deflators in not replicating the regional patterns in the benchmark price index also carries over into the Table 2 results. For example, the SSD value for the Engel curve deflators is 400 times larger than the SSD using the prices from the expert informants. There is little difference in the SSD values for the Engel curve deflators if the benchmark uses fixed weights or variable weights (SSD=5774 in Table 1 and 5717 in Table 2). Thus, at least in terms of one no-price alternative, of basing spatial deflators on a food Engel curve, using the expert knowledge of local informants better proxies for prices and for the spatial cost of living index that a traditional retail price survey would provide.

## V. Spatial Price Patterns and the Alchian-Allen Effect

The use of food Engel curves to calculate spatial deflators is a recent development but unit values have been used as proxies for prices for longer. For example, as early as 1955 there were warnings that survey groups cover a range of different varieties, each selling for a different price, and if the group mix changes unit values will not reflect prices (Prais and Houthakker, 1955). The particular conditions needed for unit values to be a consistent measure of prices were highlighted by Deaton (1988); the price of each and every food in a survey group – for all kinds of varieties, each of different quality – have to move in fixed proportions over time and space. Fixed relative prices within groups is also known as *Hicksian separability*.<sup>13</sup> This violates a basic pricing feature; the Alchian-Allen effect, where the relative price of quality will vary over space due to transport costs. In this section we consider whether expert informant prices can also show this effect.

If sellers offer a good in a market further from the production point they need to pass on some costs to the buyer, such as for transport (and/or storage). Alchian and Allen (1969) note that such costs will lower the relative price of, and hence should raise the demand for, high-quality goods. In other words, for some transformation cost  $t$ , such as for transport:

$$\frac{p^h}{p^l} > \frac{p^h + t}{p^l + t}$$

---

<sup>13</sup> This is one of the four possible conditions for theoretically consistent micro analysis on more aggregate data. The others are the Leontief composite commodity theorem (quantities for items in the group move in exact proportion), the generalized composite commodity theory (price deviations for each individual food in the group are independent of income and of all group-level price indexes), and homothetic separability of utility.



where  $p^h$  is the price of the high-quality variety, which exceeds that of the low-quality variety,  $p^l$ . Thus, demand for the high-quality variety will be relatively greater, the further from the point of production (and so the quality mix within a group changes). This effect is also known as ‘shipping the good apples out’ from the fact that high quality apples produced in Washington state are relatively cheaper in East Coast markets of the United States, than in the West Coast markets closer to the production point (Borcharding and Silberberg, 1978). The Alchian-Allen effect should occur with any charges for transport, storage or processing, and can account, for example, for why purchases in smaller packages tend to be of higher quality varieties (Gibson and Kim, 2018).

The Alchian-Allen effect has been shown for Vietnam, using traditional price survey data (Gibson and Kim, 2015). A good example is rice, whose production is concentrated in the south, especially in the Mekong Delta (a major export point to the rest of the world, as well as to the rest of Vietnam). High quality rice is relatively more expensive in the south, with a price premium of 47 percent versus 33 percent in the north, because the market surplus flows from south to north and it costs the same to ship high quality rice as low quality rice. This within-group relative price variation lets consumers switch to relatively cheaper items, so the composition of the rice group varies over space. Gibson (2016) shows that south-to-north variation in relative prices of high quality rice tilts the composition of demand to high quality rice in the north, raising the unit value there by six percent, irrespective of actual spatial price differences. This matters to a food price index, or to a food poverty line, because of the big share (over one third) of rice in the food basket. In fact, the Alchian-Allen effect just for rice causes a spurious five percent gap in the head count poverty rate in the north versus the south in 2010 (Gibson 2016).

If price data obtained from local experts are a good proxy for prices collected from market surveys, they should show basic spatial features of the data, such as the Alchian-Allen effect. The results in Table 3 suggest that the prices for different quality rice varieties provided by local expert informants do show the Alchian-Allen effect, much as is seen with the market price survey data. There are six regressions reported in this table, based on provincial averages of the price data, and of quantity data from the VHLSS. We examine how the relative price of high quality to low quality rice (and the quantities demanded) varies with distance from the main city in the Mekong delta, Can Tho, which is taken as the shipping point for excess supply to move throughout Vietnam.

The results in the first column of Table 3 show that the ratio of the average quantity of high quality rice to low quality rice bought in each province rises by two percentage points for every 100 kilometres the province is from Can Tho. In contrast, rice from non-market sources (e.g, from own-production) has no statistically significant Alchian-Allen effect, which is to be expected because self-produced rice is not subject to transport cost. The driving force for this

change in the mix of rice quality that is bought is the change in relative prices; the results in the third column show that, according to the market price survey, the price ratio of high quality to low quality rice falls by 1.20 percentage points for every 100 kilometres from Can Tho. If the prices from the local experts are used, the rate of relative price change with distance is about the same, ranging from  $-1.19$  to  $-1.35$  per 100 kilometres. Thus, in addition to providing data that give similar regional cost of living indexes to what market price surveys show, expert knowledge about prices can also provide a good approximation for a basic spatial feature of prices – the Alchian-Allen effect.

**Table 3: The Alchian-Allen Effect for Rice, Using Traditional Market Survey Price Data and Local Expert Knowledge About Prices**

	Quantity of High Quality Rice		Relative Price of High Quality Rice to Low Quality Rice			
	Relative to Low Quality		Traditional Market Survey	Expert Knowledge Survey		
	Purchases	Non-Market		Minimum	Typical	Maximum
Distance	1.98 (3.36)***	-2.01 (0.81)	-1.20 (4.52)***	-1.35 (5.83)***	-1.32 (5.59)***	-1.19 (5.09)***
Intercept	1.96 (1.13)	30.59 (1.08)	1.50 (48.49)***	1.51 (56.44)***	1.51 (56.02)***	1.51 (56.11)***
R-squared	0.13	0.02	0.22	0.30	0.27	0.23

*Notes*

Results use provincial averages, for  $N=59$  provinces, with the quantity ratios and price ratios in percentage terms. Distance is from Can Tho in the Mekong Delta to the  $i^{\text{th}}$  province (in 100 km units), where Can Tho is for the shipping location for surplus production.  $t$ -statistics in ( ) from heteroscedasticity-robust standard errors; \*  $p<0.1$ ; \*\*  $p<0.05$ ; \*\*\*  $p<0.01$ .

## VI. Comparing Prices from Expert Knowledge with Food Group Unit Values

With the Alchian-Allen effect, average unit values for different regions will not refer to the same quality mix, and so will misrepresent spatial cost of living differences. Prior studies also show that unit values are a poor proxy for prices (for example, Gibson and Rozelle, 2005) but continued use of unit values in applied studies makes further demonstration of this point useful. Therefore, in Table 4 we compare the performance of unit values with that of the prices from the expert informants, in terms of being accurate proxies for the market prices from the benchmark survey. The comparison is restricted to the 30 food groups for which unit values are available (the other VHLSS groups do not have metric quantities available). The first part of the table compares the national means from each of the three types of data for each food group, the second part reports the correlations with the benchmark prices, and the third part shows how many communes have data from unit values and how many have data on local prices provided by the expert informants.

**Table 4: Performance of Expert Opinions and Unit Values as Proxies for Food Group Prices**

VHLSS Food Group	Mean Price	Percent Difference from Mean		Correlation with Prices <sup>a</sup>		Number of Communes with	
		UVs	Opinions	Unit Values	Opinions	UVs	Opinions
Rice (all grades and varieties)	10522	-7.5%	-0.8%	0.53	0.84	748	1043
Sticky rice	16263	-5.0%	3.9%	0.30	0.30	361	914
Bread and wheat flour	24291	-11.0%	-1.6%	0.07	0.88	594	595
Flour noodles and porridge	34310	-34.1%	2.3%	0.07	0.14	942	944
Fresh and dried rice noodles	7313	39.6%	0.2%	0.08	0.90	752	754
Pork of all types	50108	7.7%	0.9%	0.53	0.91	1027	1049
Beef of all types	102732	7.4%	-0.3%	0.26	0.85	647	791
Chicken of all types	75249	-11.5%	-3.7%	0.21	0.64	626	1034
Duck and other poultry	52072	-14.5%	0.8%	0.37	0.85	485	748
Fresh fish and shrimp	84188	-59.7%	-8.4%	0.18	0.47	985	1029
Dried fish and shrimp	58068	18.8%	-2.6%	0.16	0.71	613	629
Eggs (chicken and duck)	2678	-19.9%	-1.0%	0.00	0.82	907	1037
Tofu	11520	5.1%	0.3%	0.39	0.88	956	959
Peas of various types	11062	-16.3%	-1.1%	0.24	0.65	636	700
Morning glory (water spinach)	4486	-4.4%	0.9%	0.46	0.94	870	996
Cabbage	7972	1.7%	0.9%	0.06	0.83	641	653
Tomato	9107	4.2%	0.6%	0.50	0.94	865	875
Orange	20750	-13.1%	2.2%	0.34	0.87	550	578
Banana	6128	18.0%	1.4%	0.36	0.89	709	874
Mango	25146	-30.6%	-1.6%	0.37	0.89	301	344
Cooking sauces	29000	-41.0%	0.4%	0.23	0.77	1034	1035
Salt	5821	-17.6%	4.3%	-0.09	0.55	1010	1017
Sugar and molasses	20425	-11.8%	-1.2%	-0.02	0.74	980	1035
Confectionery	44723	-12.6%	3.1%	0.13	0.62	655	1040
Condensed milk	42559	191.5%	0.5%	0.17	0.69	538	556
Wine and spirits	91393	-84.4%	1.5%	-0.13	0.36	828	1003
Beer	19588	-20.9%	-4.0%	-0.24	0.12	452	795
Water and soft drinks	13807	-11.7%	3.0%	0.03	0.04	397	851
Coffee	98155	-27.6%	0.7%	0.15	0.58	206	586
Tea	92884	-18.8%	-1.1%	-0.04	0.67	868	986
<b>Unweighted mean<sup>b</sup></b>		<b>25.6%</b>	<b>1.8%</b>	<b>0.19</b>	<b>0.68</b>		
<b>Unweighted median<sup>b</sup></b>		<b>15.4%</b>	<b>1.1%</b>	<b>0.17</b>	<b>0.75</b>		
<b>Budget share-weighted mean<sup>b</sup></b>		<b>22.3%</b>	<b>2.1%</b>	<b>0.33</b>	<b>0.73</b>		
<b>Budget share-weighted median<sup>b</sup></b>		<b>7.7%</b>	<b>0.9%</b>	<b>0.37</b>	<b>0.84</b>		

*Notes*

Mean prices are in VND per kilogram (or liter). The items in the table are the only food groups with quantity (or volume) data, which is needed for unit values (UV) to be calculated. The exchange rate at the time of the survey was approximately 19,300 Dong per US dollar.

<sup>a</sup> The correlation with market prices is across all the communes with market prices and unit values (or with market prices and expert opinions) available. This number of observations is reported in the last two columns of the table.

On average across the 30 food groups, the mean unit value differs from the mean price by 25.6%, and for the median food group the discrepancy is 15.4%. These summary statistics treat all food groups equally and so minor groups whose unit values are poor proxies for prices exert a big influence on the results. With budget share-weighted statistics, the mean discrepancy is 22.3%. The commune-level prices opinions from the expert informants have discrepancies from the mean market prices that are an order of magnitude smaller than the discrepancies for the unit values, with means of 1.8% to 2.1% and medians of 0.9% to 1.1%. Likewise, the correlations between the benchmark prices and the expert informant data are far higher than the correlations between unit values and prices; in terms of the budget share-weighted statistics the mean and median correlation is 0.73 and 0.84 for the price opinions, compared to just 0.33 and 0.37 for the unit values. A further difficulty with unit values is that they are missing for many more communes (on average, each food group has no purchasers amongst all VHLSS households in 350 communes). In contrast, there are many more communes with prices from the expert informants, and this greater availability of data also contributes to the better performance of price opinions as proxies for market prices.

Our final comparison, which is reported in Table 5, uses a Laspeyres food price index for the six broad regions (and for the rural and urban sector within each region), which uses either the prices from the traditional survey of markets, the prices from the expert informants, or the unit values. The food groups covered by this index are the 30 groups shown in Table 4. For either product-moment, or ranks, the correlations between the price index using the traditional survey and the price index using the expert informant data are 0.98. In contrast, for the price index based on the unit values, the rank correlation with the benchmark price index is only 0.62 while the product-moment correlation of the unit value index with the benchmark is only 0.70.

Table 5 also shows the squared differences between the benchmark food price index and the other two indexes (columns 3 and 5) and it is apparent that the unit values are an especially poor proxy for prices in the Mid-Northern Mountains region. While the Engel curve deflators in Tables 1 and 2 suggest this is a high cost of living region (exceeding the cost of living for the urban sectors of the two regions with big cities – Red River and South East) the unit values suggest it is a low cost of living region. One explanation is that this is a poor region, and so unit values likely refer to a lower quality mix within food groups than the quality mix in richer areas, while food price differences are not so large (for example, the price indexes suggest this region has four percent lower food prices than in the base region).<sup>14</sup> However, even with cheaper food (on both the price and quality margin), food shares are high because incomes are low, and the Engel curve method mistakes this high food share for a high cost of living. Regardless of the causes, the overall performance (in approximating what price survey

---

<sup>14</sup> Gibson and Kim (2013) show that there is considerable scope for quality downgrading within the food groups covered by the VHLSS, especially for major groups like rice.

data would show) of the food price index based on the unit values is quite poor, with an SSD value that is an order of magnitude larger than when the prices from the expert informants are used (276 versus 27).<sup>15</sup>

**Table 5: Food Price Indexes from Expert Informant Prices and from Unit Values Compared to using Benchmark Price Surveys**

	Market Price Survey (1)	Expert Informants (2)	Squared Differences [(1)-(2)] <sup>2</sup>	Food Group Unit Values (4)	Squared Differences [(1)-(4)] <sup>2</sup>
Urban Red River	100.0	100.0	0.0	100.0	0.0
Urban Mid-Northern Mountains	98.1	96.0	4.2	91.2	46.6
Urban North-Central Coast	87.1	83.6	12.3	92.3	26.5
Urban Central Highlands	92.6	92.5	0.0	89.5	9.5
Urban South East	100.7	100.8	0.0	99.9	0.6
Urban Mekong Delta	86.8	86.9	0.0	87.3	0.2
Rural Red River	89.5	89.6	0.0	86.5	8.6
Rural Mid-Northern Mountains	96.4	96.1	0.1	83.4	168.2
Rural North-Central Coast	84.9	81.8	9.7	85.2	0.1
Rural Central Highlands	89.0	89.7	0.5	86.8	4.7
Rural South East	91.1	90.7	0.2	87.9	10.5
Rural Mekong Delta	82.7	82.7	0.0	83.4	0.6
<b>Sum of Squared Differences</b>			<b>27.0</b>		<b>276.1</b>

*Notes*

The food price index is calculated using the Laspeyres index. The commune-level means of the expert informant prices are weighted by their inverse variances, based on the triangular distribution.

## VII. Conclusions

In this paper we report on a survey experiment that tested an idea that was first discussed almost 40 years ago, which is to gather commodity-wise and spatially disaggregated price data by asking expert informants. In a typical household survey in developing countries, outsiders come into a rural area for just a brief period – perhaps a day or two – in order to survey expenditures. The interview teams have little time (or desire) to go to markets that may be some distance away and may meet only intermittently. Thus, many of these surveys take no record of prices. In contrast, the local residents – and especially informed people such as leaders of women’s groups – have a very good idea about the distribution of local prices. What has been lacking is a straightforward method to elicit the expert local knowledge, in a consistent manner over space (and potentially, over time), and also a demonstration that this method can work. Our approach was to use pictures of the target specifications, so that throughout Vietnam the expert informants were referring to the same items, and to ask about low, typical, and high prices, so that we could use triangular distributions to get the means and variances of local prices.

<sup>15</sup> If we used a WCPD-vw food price index, the SSD when using unit values would be almost 15 times as large as the SSD when using the expert informant prices, so use of the fixed weight Laspeyres index is not the cause of the poor performance of unit values.

Our approach to obtaining price data from expert informants is a hybrid of what has been done previously in Indonesia and PNG. However, the scale of our experiment dwarfs previous surveys and our evaluation of the data provided by expert informants is far more comprehensive. In Vietnam, using expert local informants provided 99.3% of the data that would be obtained from a traditional survey of retail markets. The data from the expert informants allowed spatial price indexes to be calculated that closely approximate the benchmark price indexes calculated using the traditional approach to surveying stores and markets. While there was no time dimension to our experiment, we have no doubt that price data from expert informants would be just as accurate for time-space deflation. We also found that a basic spatial feature of prices – the Alchian-Allen effect – is illustrated with the expert informant prices in about the same way as it is with the prices from the traditional survey approach. This effect matters because it undermines use of unit values as a proxy for prices.

While the idea of using expert informants was suggested long ago it is rarely implemented. Instead, both an older approach – using unit values as a proxy for prices – and a newer approach – using food Engel curves to derive deflators – are often used by analysts in countries that lack price surveys when they calculate poverty lines, or more generally, when doing welfare analysis that needs deflated data. Our results suggest that both of these no-price methods provide very poor approximations to the benchmark deflators that would be provided by price surveys. With regard to using Engel curves, our findings corroborate those of Gibson *et al.* (2017), who find a substantial distortion in estimates of the level, location and change in poverty if Engel curve deflators are used in Vietnam. With regard to unit values, we add to the concerns first noted by Prais and Houthakker (1955) that these are not a valid indicator of price levels if the quality mix within survey groups changes. We also note that a changing quality mix over space (and time) is exactly what the Alchian-Allen effect would predict, and we demonstrate this effect for a key food in Vietnam, rice, using both the market price surveys and the expert informant prices. In light of the weaknesses with no-price methods, and given the feasibility of asking local experts about prices – with guides such as photographs to ensure consistency – we recommend more household surveys in developing countries should attempt to experiment with gathering price data by using expert local knowledge.

## References

- Alchian, A. and W. Allen (1967) *University Economics*, 2nd Edition, Wadsworth Publishing Company: Belmont, California.
- Almås, I., Kjelsrud, A. and Somanathan, R. (2018) A Behaviour-based Approach to the Estimation of Poverty in India. *Scandinavian Journal of Economics* (forthcoming).
- Borcherding, T. and Silberberg, E. (1978) Shipping the good apples out: the Alchian and Allen theorem reconsidered. *The Journal of Political Economy* 86(1): 131-138.
- Breuer, C. and von der Lippe, P. (2011) Problems of Operationalizing the Concept of a Cost-of-Living Index. *MPRA Paper* No. 32902.

- Deaton, A., Friedman, J., and Alatas, V. (2004) Purchasing Power Parity Exchange Rates from Household Survey Data: India and Indonesia. *Princeton Research Program in Development Studies Working Paper*.
- Deaton, A. (1988) Quality, quantity, and spatial variation of price. *American Economic Review* 78(3): 418-430.
- Deaton, A. (1989) Household survey data and pricing policies in developing countries. *The World Bank Economic Review* 3(2): 183-210.
- Deaton, A. and Dupriez, O. (2011) Spatial Price Differences Within Large Countries. *Mimeo*, Princeton University.
- Diewert, E. (2005) Weighted Country Product Dummy Variable Regressions and Index Number Formulae. *Review of Income and Wealth* 51(4): 561-570.
- Dumagan, J. and Mount, T. (1997) Re-examining the Cost-of-Living Index and the Biases of Price Indices. *Department of Commerce Working Paper ESA/OPD*, 97-5.
- Gibson, J. (2016) Poverty measurement: We know less than policy makers realize. *Asia and the Pacific Policy Studies* 3(3): 430-442.
- Gibson, J. and Kim, B. (2013) Quality, quantity, and nutritional impacts of rice price changes in Vietnam. *World Development* 43(1): 329-340.
- Gibson, J. and Kim, B. (2015) Hicksian Separability Does Not Hold Over Space: Implications for the Design of Household Surveys and Price Questionnaires. *Journal of Development Economics* 114(1): 34-40.
- Gibson, J. and Kim, B. (2018) Economies of scale, bulk discounts, and liquidity constraints: Comparing unit value and transaction level evidence in a poor country. *Review of Economics of the Household* 16(1): 21-39.
- Gibson, J., Le, T., and Kim, B. (2017) Prices, Engel curves, and time-space deflation: Impacts on poverty and inequality in Vietnam. *The World Bank Economic Review* 31(2): 504-530.
- Gibson, J. and Rozelle, S. (2005) Prices and unit values in poverty measurement and tax reform analysis. *The World Bank Economic Review* 19(1): 69-97.
- Gong, H. and Meng, X. (2008) Regional Price Differences in Urban China 1986-2001: Estimation and Implication. *Discussion Paper* No. 3621, Institute for the Study of Labor (IZA), Bonn.
- Hamilton, B. (2001) Using Engel's Law to Estimate CPI Bias. *American Economic Review* 91(3): 619-630.
- Prais, S. and Houthakker, H. (1955) *The Analysis of Family Budgets*, New York: Cambridge University Press.
- Rao, P. (2005) On the Equivalence of Weighted Country-Product-Dummy (CPD) Method and the Rao-System for Multilateral Price Comparisons. *Review of Income and Wealth* 51(4): 571-580.
- Saunders, C. and Grootaert, C. (1980) Reflections on the LSMS group meeting. *Living Standards Measurement Study Working Paper* No. 10, The World Bank.
- Selvanathan, E. (1991) Standard Errors for Laspeyres and Paasche Index Numbers. *Economics Letters* 35(1): 35-38.
- Van Veelen, M., and van der Weide, R. (2008) A Note on Different Approaches to Index Number Theory. *American Economic Review* 98(4): 1722-1730.
- Winters, L. A., McCulloch, N., and McKay, A. (2004) Trade liberalization and poverty: The evidence so far. *Journal of Economic Literature* 42(1): 72-115.
- Wood, D., and Knight, J. (1985) The collection of price data for the measurement of living standards. *Living Standards Measurement Study Working Paper* No. 21, The World Bank.
- World Bank (2012) *Well Begun, Not Yet Done: Vietnam's Remarkable Progress on Poverty Reduction and the Emerging Challenges*, World Bank: Hanoi.

**Appendix Table 1: Items in the Price Survey and Expert Informant Survey, and Number of Communes With Price Data**

Description of target specification	Number of communes with prices for target	Number of communes with imputed price	Number of communes with price opinions
White rice #1 (lower quality e.g. IR50404)	935	111	950
White rice #2 (premium variety e.g. Bac Huong)	868	153	893
Sticky rice, Nep Nhung type	850	195	864
White bread - 100g local loaf, unpackaged	993	23	982
Instant noodles - Hao Hao brand, 75g pkt	1027	22	1038
Fresh rice noodles	1028	9	1004
Pork: Rump	1038	0	1022
Pork: Belly	1038	0	1021
Beef	982	7	959
Fresh beef rib	970	14	945
Battery (caged) chicken, 1kg whole	811	43	778
Live free range chicken	987	1	978
Carcase of free range chicken	677	72	647
Whole local duck	834	48	818
Pork- pie	929	3	898
Lard	1022	1	1007
Cooking oil, Neptune brand, 1 liter	930	119	916
Fresh fish, 1 whole Carp ca. 500g	800	176	810
Salt-water tiger shrimp, 7-10cm long	643	208	662
Fresh-water shrimp, 3-5 cm long	801	96	795
Salted fish	811	139	835
Chicken eggs, ca. 65g each, 10 eggs	968	70	983
Tofu	1040	4	1027
Fresh pea	821	87	801
Water morning glory	1044	2	1029
Cabbage	1007	6	981
Tomato	1045	0	1024
Orange	812	141	815
Banana	918	118	931
Mango	609	172	569
Fish sauce, Nam Ngu-Shinsu brand, 500ml bottle	1047	2	1030
Salt, MS brand, 1kg bag	407	638	493
White sugar, Bien Hoa brand, 1kg	678	363	714
Fruit candy, Hai Ha brand, 100g	689	309	734
Condensed milk, Omg Tho, 380ml	662	381	636
Vodka, Ha Noi, 300ml bottle	659	176	624
Bottled beer - Ha Noi, 450 ml bottle	474	332	497
Bottled beer - Sai Gon, 450 ml bottle	566	162	590
Soft drink, Coca Cola, 330ml can	947	65	956
Fruit juice, Twister brand, 330ml can	808	172	836
Bottled water, La Vie brand	744	203	800
Powdered coffee, Trung Nguyen, 250g	672	250	637
Dried tea, Trung Nguyen, 1kg	756	229	770
Cigarettes, Vinataba brand, pack of 20	598	363	641
Cigarettes, Craven brand, pack of 20	633	304	628
Outdoor meals - breakfast	972	45	941
Outdoor meals - lunch/dinner	956	5	914
Kerosene, Petrolimex, 1 liter	953	11	915
Gasoline, 92 octane, Petrolimex, 1 liter	1015	1	999
Laundry powder, Omo brand, 800g pack	1046	3	1030
Shampoo, Rejoice brand, 360ml bottle	911	118	851
Soap, Lux brand, 90g bar	800	231	793
Toothpaste, PS brand, 200g tube	1003	42	995
Toilet paper, An An brand, 10 rolls	559	476	555
Television set, Samsung CRT, 21 inch	669	40	582
Men's haircut	1030	4	1022
Women's haircut	1013	4	987
Tailoring of men's trousers	1016	0	998
Tyre puncture repair	1034	4	1041
Brick, locally made, two hole construction	782	171	722
Medicine, Paracetamol, 500mg, 10 tablets	964	67	965
Medicine, Decolgen, 500mg, 4 tablets	1001	23	999
Educational supplies - student notebook	385	623	421
Educational fee at public secondary school	919	0	915
<b>TOTAL</b>	<b>54606</b>	<b>7857</b>	<b>54213</b>