

Spatial price differences within large countries

Angus Deaton

Olivier Dupriez

Princeton University and World Bank

July 2011

We are grateful to Ingvild Almås, Bettina Aten, Yuri Dikhanov, Alan Heston and Robert Hill for comments on a previous version of this paper.

1. Introduction

National and international statistical systems are strangely reticent on differences in price levels within countries. Nations as diverse as India and the United States publish inflation rates for different areas, but provide nothing that allows comparisons of price levels across places at a moment of time. The International Comparison Project, which at each round collects prices and calculates price indexes for most of the countries of the world, publishes nothing on within country differences, and in some important cases including China, Brazil, and India, rural prices are either not collected or are underrepresented, which potentially distorts the comparison between large and small countries, Deaton and Heston (2010). Yet spatial price indexes are required if we are to compare real incomes across areas, and both national and global poverty estimates require intra- as well as international price indexes. There are also good grounds for suspecting that price levels differ across regions. Several large Indian states are larger than most countries, and consumption patterns are sharply different across the subcontinent. And for the same reasons that we expect price levels to be lower in poorer countries—the Balassa-Samuelson theorem—we would expect prices to be lower in poorer areas within countries, at least if people are not completely mobile across space.

In this paper, we use unit values from household survey data to document spatial differences in food prices for two large countries, Brazil and India; comprehensive regional price indexes for the United States have recently been calculated by Aten, Figueroa and Martin (2011). Although our estimates are only a

first step towards a more complete accounting, they document substantial differences in food prices between urban and rural areas, and between different regions. These patterns vary by levels of development; in rural India, a significant fraction of foods are produced at home, and these provide a price advantage to rural consumers. Food subsidies are also much more effective in rural than in urban India, and in some Indian states than in others. In richer countries like Brazil, processed foods are relatively more important, so that food prices are sometimes higher and sometimes lower in cities, depending on the food. Our evidence is also consistent with a general diminution of the rural advantage in food prices as countries become richer.

A second concern of our paper is specific to India. Official poverty estimates in India long assumed that urban prices were 15 percent higher than rural prices. Following a 1993 report, Government of India (1993), a more elaborate set of price differentials was used, with differences across states, as well as between the rural and urban areas of each state. Several of these price differentials are implausible, and have long been believed to have come from a calculation error. Most recently, a new set of poverty lines have been proposed by a committee chaired by Tendulkar, Government of India (2009). These differentials were calculated using an extension of the methods in Deaton and Tarozzi (2005) which, like the methods of this paper, use unit values from household surveys. One immediate problem with unit values is that they reflect not only prices, but quality effects, so that price indexes that use unit values without adjustment are likely to be artificially

higher in better-off places. However, these effects were typically thought to be small, see Deaton (1988, and 1997, Chapter 5) for evidence from India and elsewhere, and none of the earlier Indian calculations made any adjustment for quality. However, a recent paper by Dikhanov (2010) finds that, once unit values are adjusted for income effects, the differences in rural and urban prices in India are small enough to ignore. We revisit the Indian results, and reconcile Dikhanov's findings with the previous literature; that reconciliation shows that there are indeed important price differences within India, although they are substantially smaller than those in the new poverty lines. We also take the opportunity to calculate truly multilateral price indexes for India to replace the bilateral indexes in Deaton and Tarozzi (2005) and Government of India (2009).

Our data and methods do not allow us to go (much) beyond food prices. Yet it is clear that many other prices differ from one place to another, particularly anything whose price depends on land prices. Aten, Figueroa, and Martin (2011) find that rents and fuel prices differ sharply from place to place in the US, and that those prices affect the prices of many other goods. Absent prices for non-foods, we follow a different strategy and in a final section examine differences in self-reported well-being by area as evidence about differences in real income. Of course, even if self-reported well-being is a good indicator of well-being, it will be a function not only of real income but also of area amenities, and we see evidence of this in our results.

It might be thought that the failure of the statistical community to produce spatial price indexes simply reflects the fact that such indexes serve no policy purpose. But this is certainly not the case for poverty measurement. In the United States, the National Academy Panel on Poverty and Family Assistance, Citro and Michael (1995), recommended that local price levels be approximated by using indexes of house prices across space. This proposal generated a substantial subsequent research effort within federal statistical agencies, see for example Short, Garner, Johnson, and Doyle (1999), Renwick (2011), Short (2011) and Aten, Figueroa and Martin (2011). In India, the Tendulkar report was a consequence of similar concerns. For global poverty, the World Bank's calculations adjust the PPPs from the International Comparison Project for urban to rural price indexes in a number of large countries, including India and Brazil, see Chen and Ravallion (2010), but the calculations rely on the ratios of local urban to rural poverty lines, themselves sometimes of dubious provenance. Instead, the lack of these indexes more likely reflects the difficulty and cost of producing them, General Accounting Office (1995). Beyond that, because the multilateral price indexes that are required for spatial work are typically not consistent with the inflation rates in local consumer price indexes, multilateral indexes need to be calculated regularly, not just once and updated by the local CPIs.

We note finally that the use of local prices to calculate spatial real incomes does not necessarily embody a commitment to using them to calculate transfer payments. Within countries, where people are much more mobile than across

countries, people will tend to move to the places where they are best off, so that in equilibrium, high price cities, like Delhi, Rio de Janeiro, or New York must have amenities that offset their high housing and transportation costs, see Roback (1982). In such a case, real income poverty may not be enough to justify transfers, see Glaeser (1998) for an analysis that derives practical formulas depending on the sources of the price variation, and the degree of spatial mobility. Such considerations may also have contributed to the lack of interest in calculating spatial price indexes within countries.

The paper is laid out as follows. Section 2 presents our results for India and Section 3 is concerned with Brazil. Section 4 raises some general considerations about the estimates, presents the results on self-reported well-being, and concludes.

2. Unit values and prices for food in India

The Indian data are taken from the *Socio-Economic Survey, sixty-first round*, conducted by the National Sample Survey Organization (NSSO) from July 2004 to June 2005 on a sample of 124,644 households. Data on food consumption were collected using a 30-day recall period. In the Consumer Expenditure Schedule, household consumption data are broken down into 340 goods and services, including 146 food items and non-alcoholic beverages. We focus on the food items here, calculating unit values by dividing expenditures by quantities, and then examining spatial

differences with and without correction for quality effects. We begin with differences between urban and rural households.

2.1 The major foods

Table 1 lists the calculations for the eleven most important food commodities in India, as ranked by their average budget shares over the whole country. The budget shares and unit values in this table include all consumption items, whether purchased for cash, home-grown, or received in some other way, for example as a gift or in barter exchange. We also include purchases through the public distribution system at subsidized prices in “fair price” shops; the important items here are rice, wheat or wheat flour, and sugar. The first two columns list the budget shares for each sector, rural and urban. Median and mean log unit values are shown, together with the difference between them. Note that the median of the log is identical to the log of the median, so that when working with medians, it is of no consequence whether we work with prices or with their logarithms.

One issue with medians which will become apparent later is that they are sometimes identical across sectors, even when one distribution is clearly to the right of the other. This happens because many quantities and expenditures are reported as whole numbers, so that there is sometimes a single unit value that attracts a substantial fraction of the sample. Suppose that 20 percent of the sample reports x , 40 percent reports $x-1$, and 40 percent reports $x+1$, so that the median is x . If the distribution becomes 35 percent below x , 20 percent at x , and

45 percent above x , a clear upward shift, the median remains at x . For this reason, we focus on the means of logarithms of prices, which differs from the procedures reported in Deaton and Tarozzi (2005), for example. Their original concern was with outliers in the unit values, but working with logarithms appears to be sufficient to deal with that here.

For these eleven goods, urban prices are 17 percent higher than rural prices using the medians, and 19 percent higher using the means of logs; these averages come from the differences for each commodity weighted by their budget shares. These estimates are in line with standard Indian numbers, and as we show below, they do not change very much when we include all foods. Both estimates are two percentage points lower if the publicly distributed items are excluded. As we shall see below, foods obtained through fair price shops are much more important to rural than to urban consumers. It is worth also noting that the urban rural difference is not the same for all goods. For example, for the two major cereals, rice and wheat, both of which are produced in the rural sector, the difference in the mean log unit values is 28 and 26 percent. The differences for milk and for fish are 19 and 18 percent. Processed foods like sugar and mustard oil show little or no difference. Tea leaf costs essentially the same in urban and rural, but tea cups, which embody more labor costs, are cheaper in the poorer sector.

These calculations make no allowance for income driven quality effects in the unit values. These are investigated in Table 2 which shows the results of running a household-level regression of the log of each unit value on the logarithm of

household per capita expenditure and an urban dummy. The coefficient on the latter is an estimate of the difference in log unit values between the urban and rural sectors once income effects have been removed and can be compared with the unadjusted difference in logs in Table 1, reproduced here as the third column in Table 2. Note that the ability to run these household-level regressions is another advantage of working with the logs of unit values, rather than with medians, which are obviously not available at the household level.

The income elasticities of the unit values are in some cases substantial, more than 26 percent for rice, 18 percent for cups of tea, 17 percent for wheat/atta, and 11 percent for fish, all categories where there is much quality variation within the category and scope for better-off people to buy more expensive varieties. These estimates are generally higher than those estimated in Deaton (1997, Chapter 5) perhaps because of an expansion of varieties as India has become richer compared with Maharashtra in 1983.) Because of the income effects, the adjusted urban/rural differences in column 2 are smaller than the unadjusted differences in column 3, by almost a half in the case of rice, and by around a third over all these goods. The overall price difference, weighted by an average of the rural and urban budget shares, falls from 19.2 percent with no adjustment to 11.5 percent with the adjustment.

2.2 Unit values by source

The analysis in Section 2 used all unit values for food consumption, irrespective of how the item was obtained. Here we look more closely at unit values by source, with particular attention to the difference between cash purchases and other sources, mostly home production, and between free-market prices and the prices of goods purchased in fair price shops through the Public Distribution System. Table 3 shows the nine cases where there is a non-trivial budget share for non-cash or for subsidized purchases. Sugar, wheat or wheat flour, and rice are provided through the Public Distribution System. There are also significant amounts of non-cash consumption (nearly all of which is of home-produced items) of rice, wheat, and milk (where the non-cash share is larger than the cash share in the rural sector), and much smaller amounts of mustard oil, sugar, and fish. There are substantial differences in mean (and also median, not shown here) unit values. Sugar is about 30 percent cheaper from the PDS, and rice and wheat are less than half the price, with the percentage discount larger in urban areas; that said, the budget shares of the PDS commodities are very small among urban households. Even for rural households, the share of expenditure from the PDS is less than 14 percent of the share of cash purchases for rice, 10 percent for wheat, and 9 percent for sugar. As is to be expected, home-produced goods are also recorded at lower prices, though the differences are relatively modest, 14 percent for rice, 23 percent for wheat, and 9 percent for milk.

We can also calculate the urban price differentials that would be found if we were to use only the unit values from cash purchases. While this is not a desirable calculation in itself—the prices of subsidized and own-produced goods should be included in the price indexes—it is comparable to the differentials that would be observed if only “shop” prices were collected, as is the case in the ICP itself. Calculations of urban/rural price differentials have been carried out by Hill and Syed (2010) using ICP 2005 data for a group of Asian countries for all goods and services (not just food) and they find only small differentials, between two and three percent averaged over five countries, estimates that they regard as implausibly low.

Table 4 reproduces the Indian results in Table 2, but uses only the cash prices, and omitting purchases from the Public Distribution System. As anticipated, this reduces both the raw and adjusted differences in the means of the logarithms of the unit values. For example, for wheat and *atta*, where the unadjusted and adjusted differences in Table 2 are 26 and 16 percent respectively, the corresponding numbers for cash only non-PDS purchases are only 20 and 12 percent. (In this case, it is the exclusion of the PDS that makes the largest difference.) For rice, the numbers fall from 28 and 14 percent to 22 and 9 percent. These two groups of goods are the ones that are most affected, but together they comprise a substantial fraction of the food budget, so the exclusion of the PDS and the move to cash prices has a correspondingly substantial effect on the overall index of urban to rural prices. At the end of the last subsection, we reported the overall

unadjusted and adjusted indexes as 19.2 and 11.5 percent. Using cash prices only, these estimates fall to 13.5 and 6.8 percent.

2.3. Extension to all foods

We began with selected important food items so that the calculations could be transparent. However, the NSS survey provides unit values for 136 foods which together account for 60.4 percent of the rural budget, and for 50.2 percent of the urban budget, both figures calculated from the household budget shares and averaged over households. The survey collects unit values on a few other items, such as tobacco, some fuels, and some clothing, but the total of these items is small, and some of the unit values have proved problematic in the past. For much of the non-food budget, no unit values are collected. So for the moment, we confine ourselves to the foods. Even among those, there are likely some where the unit values are unreliable, but each of the remaining budget shares is small, and so the effect will be limited.

For all 136 foods, the average of the unadjusted urban/rural price difference, with averaged budget shares as weights (the Törnqvist index) is 17.8 percent, only a little less than the average for the eleven goods. With adjustment for income effects, following the same procedures as in Table 2, the difference is 10.3 percent, so that without correction for income effects, the price differential is overstated by about 70 percent. Allowing for income is important, but does not remove the need to make an adjustment, particularly for a country such as India

where a small change in the poverty line has dramatic consequences for the poverty counts.

As in the previous subsection, we have recalculated the food price for urban using only the cash purchases and treating the subsidized PDS goods as separate commodities, rather than combining their unit values with the unit values for the corresponding market goods. As before, this treatment reduces the size of the urban differential, to 13.8 percent unadjusted and 6.9 percent when adjusted for income driven quality effects. This last figure is closer to but remains larger than Dikhanov's estimate of 3.2 percent using cash-only prices (we shall return to the resolution of the discrepancy in Section 7 below). Our cash estimate is close to the 6.9 percent that Hill and Syed (2010) calculate for all prices in Malaysia, but substantially higher than their estimates for Vietnam, Indonesia, Philippines, and Sri Lanka. (It is hard to know what to expect from adding in non-foods; manufactured goods will often be somewhat more expensive in rural areas, while services—like cups of tea—will often be substantially cheaper in the countryside. Hill and Syed use a conceptually distinct quality adjustment procedure, but the effects of adjusting for quality are also consistent with what we find here.) Even so, the main point of these alternative calculations is to show that the omission of non-cash and subsidized purchases can bias downwards the estimate of urban prices relative to a rural base. In the current Indian case, a cash only treatment reduces the measured urban price differential by a third. Of course, there is no reason to suppose that such a result will hold everywhere. While own-produced

foods will always be cheaper, and always be more common in the countryside, the Indian food subsidy system—which favors rural consumers—will not carry through to other countries.

2.4 Alternative calculations using decile groups and median unit values

An alternative, and in some ways more flexible procedure, is used by Dikhanov (2010) who groups the combined rural and urban population into deciles of per capita total expenditure so that households in different sectors but in the same decile have roughly the same per capita expenditure, and then calculates median or mean unit values for each decile group in each sector. The comparison then yields an urban/rural price difference for each decile group, and these can be averaged to give an estimate for the country as a whole. While this method does not allow analysis at the household level, it has other potential advantages, including the less parametric treatment of income effects—allowing them to be different at different points of the distribution—and possibly also the use of medians, which are potentially more robust than working with logs, though logs will also downweight some extreme values.

The results are shown in Table 5. The urban rural difference in log medians does indeed vary by decile groups, and tends to be particularly large among the best off households, perhaps because it is this group that most takes advantage of the richer set of varieties that are available in urban areas. This also raises the possibility that there are further income effects *within* this group, and suggests

that there might be payoffs to a more flexible specification of income effects in the previous analysis. Note also the previously noted problem with medians highlighted by the exact zeroes in the table. In some of those cases, there are positive differences in means of logs within the decile group, so that it is possible that the medians are not sufficiently sensitive to differences in the distribution; this can happen even when the differences are not zero, because the calculation can focus on the difference between the two most common unit values in each sector. Even so, the overall number in the bottom right of the table is 8.6 percent, not much less than the 10.3 percent estimated from the earlier analysis using the means of logs and regression to control income effects.

We have also replicated Table 5 with differences of log means, rather than differences in log medians (results not shown.) The results differ in detail, but little overall. As in Table 5, the largest difference is in the top decile, where there is a 15.1 percent difference in the medians and a 16.3 percent difference in the means. The overall urban/rural price difference is now estimated to be 9.0 percent, as opposed to 8.6 percent in Table 5 using the medians. In spite of the non-parametric advantages of the deciles, our current preference is for the parametric approach using the logarithm of per capita expenditure to control the income effects, and the associated urban/rural price difference for all foods of 10.3 percent.

2.5 Differences in food prices across states

In a large country such as India, geographical differences in prices may be as important as differences between urban and rural within states. This issue is addressed for India's largest states in Table 6.

The left hand panel shows unadjusted results. These were calculated by averaging the budget shares and the logs of the unit values by commodity by state by sector, and then using the averages to calculate multilateral (Törnqvist GEKS) indexes. Each sector of each state is considered a separate geographical unit ("country") in the calculations which therefore automatically yield a set of urban rural differentials by state. (We excluded rural Delhi from the calculations, so there is no urban rural comparison for Delhi.) Rural Jammu and Kashmir is taken as the numeraire with a price level of 1.00.

In the rural sector, the highest food unit values are in Assam, Gujarat, Jammu and Kashmir, while the lowest unit values are recorded in Karnataka, Orissa, UP, and MP. The food price level in Orissa is 83 percent of the price level in Jammu and Kashmir. In the urban sector, Delhi has the highest food prices, 1.19 times the rural price level of Jammu and Kashmir. The lowest, 0.95 and 0.96, are in Orissa and Bihar, so that the gap between lowest and highest is about the same in both sectors. These estimates are unadjusted for income-related quality differences.

To correct for quality, we ran a series of household level regressions, one for each commodity, of the logarithm of the unit value on the logarithm of per capita

total household expenditure, and on a set of dummies, one for each sector of each state. We then used the estimates to calculate an adjusted average log price for each good in each state and sector by calculating what the average log unit value would be for that state and sector if everyone in the state and sector had the Indian average of log per capita total expenditure; this is just the regression prediction evaluated at the mean of log per capita total expenditure. We then inserted these adjusted prices into the multilateral calculation in the same way as for the unadjusted prices.

The results are shown in the right-hand panel. Apart from some shrinkage of the variance, with the poorer states price levels somewhat elevated, the adjusted price levels are quite similar to the unadjusted price levels. The range from top to bottom is now 20 percentage points in the rural sector (Karmataka to Assam), compared with 20 percentage points in the unadjusted numbers, and 18 points (Orissa to Assam) compared with 24 in the urban sector. As with the urban rural price differential, the state differences are surely large enough to be concerned about, at least in some contexts.

These estimates also show that the urban prices, although always higher than rural prices, even after the adjustment, are differentially so in different states. In Kerala, where urban and rural sectors are not sharply defined, the difference is only 4.6 percent. In Maharashtra, by contrast, which covers Mumbai on the one hand, and deep rural areas on the other, urban prices are 16.8 percent higher than

rural prices. Even so, many of the states have about a 10 percent urban rural price differential after adjustment for income effects.

Given that there has been a concern within the ICP about the lack of consistency of definition of rural versus urban in different countries, these estimates are of independent interest because the methodology does not require a consistent definition of urban and rural sectors. By treating each sector of each state as a separate "country," it simply works with whatever are the local definitions and delivers a price differential that matches those definitions, which is what would be required for policy calculations, for example on levels of poverty. In the Indian context, the urban rural distinction is very different in Kerala from Maharashtra, and that is reflected in the estimated price differentials.

2. Spatial food price differences in Brazil

We have replicated the Indian analysis with data from an expenditure survey from Brazil. Data for Brazil are from the *Pesquisa de Orçamentos Familiares 2002-2003*, conducted by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística, IBGE) from July 2002 to June 2003 on a sample of 48,470 households. Data on food consumption were collected using the diary method. In the household questionnaire and in the household and individual diaries, consumption data are broken down into 6,927 products and services, including 3,131 food items and non-alcoholic beverages. Like the Indian survey, the Brazilian survey has nationwide coverage.

Like India, Brazil is large enough so that the neglect of the urban-rural distinction might have important consequences for poverty measurement as well as for international comparisons in the ICP. Its per capita GDP is around three times that of India, so that food is a less important share of the budget. Higher incomes also imply that processed foods are relatively more important than in India. The lesser importance of food affects the relevance of the unit-value based methods used here, which are effectively available only for foods. At the same time, the increasing importance of processed food is likely to move urban and rural prices closer together, or perhaps even to lead to prices being higher in the countryside if they are transported in rather than transported out.

Table 7 presents, in abbreviated form, the same information that was earlier shown for India. The table lists the eleven most important foods, ranked by their budget shares averaged over urban and rural. These are the most important of the 38 foods that we consider, which together comprise 16 percent of the rural, and 9 percent of the urban total consumption budget. The first and second columns list the average budget shares, while the third lists the difference between the means of the logs of the unit values for each of the food groups. For four of the eleven goods—soybean oil, crystal sugar, ground coffee, and frozen chicken—unit values are *lower* in the urban sector. For the seven other categories—all of which are or can be locally produced—rural prices are lower, sometimes very substantially so. Fresh milk unit values are 23 percent lower, and cassava flour unit values 19 percent lower in the countryside. If we take all 38 foods together, and weight by

the averages of the budget shares to give a Törnqvist index, these unadjusted unit values are 4.9 percent higher in the cities, the number listed at the foot of the third column. We followed our earlier procedures to adjust these premia, running household level regressions of the each food's unit values on the logarithm of per capita total consumption and on an urban dummy, and we report the coefficient on the latter in column 4. As in India, this quality correction accounts for some of the urban premia, and the overall Törnqvist index of urban to rural prices falls to 2.89, shown at the foot of the fourth column.

The fifth and sixth columns repeat these calculations using only those cases where the good was obtained for cash, rather than home-grown or received as a gift. As was the case in India, the difference between urban and rural unit values is smaller when home-grown produce is excluded. In general, column 5 shows the same pattern of premia over goods, but is attenuated relative to column 3. The urban to rural Törnqvist using the cash unit values is only 2.08 percent, which falls to -0.03 percent once the unit values are adjusted for quality effects. In Brazil, even more so than in India, the difference between urban and rural food prices largely boils down to the importance of home-grown food. The relative unimportance of quality effects in Brazil is probably also attributable in part to differences in survey design; the Brazilian survey distinguishes 3,131 products compared with "only" 136 in India.

Table 8 presents the final set of results for Brazil, and looks at spatial price differences for five broad regions of the country. As in Table 6 for India, these are

calculated as multilateral (GEKS Törnqvist) price indexes for a five by two disaggregation of provinces and sectors. The unadjusted unit values show very little dispersion across the regions of rural Brazil, except for the South, where food unit values are seven percent lower than elsewhere. There is some variation across regions in the urban premia around the average of 4.9 percent in Table 7. In the right hand panel, adjustment for income effects actually increases dispersion, because there is a negative correlation across regions between unit values and incomes. The adjusted urban premia are small, consistent with the 3.15 overall adjusted premium in Table 7. Clearly, there is limited scope for price variation across these very large regional groupings.

4. Conclusions and an alternative approach

The preceding sections have shown that unit values contain useful information about prices, particularly for foods. However, unit values are not prices, and in particular they are contaminated by quality effects, and the fact that better-off households buy more expensive varieties. And because urban households are on average better-off, a comparison of unit-values without adjustment will overstate the urban/rural price difference. According to the calculations here, a suitable correction for India removes about a third of the difference, reducing a crude estimate for the all foods index from 17.8 percent to 10.3 percent. In Brazil, the unadjusted price difference is smaller, 4.9 percent, and is reduced after adjustment to 2.9 percent.

In both India and Brazil, home produced goods—which are cheaper than purchased goods—make an important contribution to the differential, and the differentials are smaller when we work only with unit values of items bought for cash. In the case of Brazil, which has an extraordinary level of detail in its consumption questionnaire, the difference in cash prices is negligible. In India, the subsidized prices in fair price shops are also important, especially in rural areas. For ICP purposes, the prices of home-grown (and subsidized) items ought to be included. That they are lower through the absence of transport and distribution margins confers a genuine benefit to people who grow and consume these goods, and that benefit is not available to people who must purchase those goods in shops and markets.

We suspect that these issues are less important for rich countries, and indeed the comparison of Brazil and India suggests that Brazil lies somewhere between the US and India in this respect. But there are a number of large, poor countries, such as India, China, Pakistan, Bangladesh, Indonesia, and Nigeria, where rural urban price differentials are likely to be important for the ICP. Some of these countries have data that would permit a replication of the methods of this paper although it should be noted that not all consumption surveys collect quantity as well as expenditure information. There are also a number of surveys, especially in Africa, where quantity information exists, but where it is difficult or impossible to use because the units are not standardized across all observations.

The general methods of this paper also have a number of important limitations. It is clearly correct that households that have higher incomes—or higher total expenditures—usually have higher unit values, a fact that can reasonably be attributed to their buying higher quality items within a heterogeneous group of foods, or buying similar items in more upscale outlets. What is less clear is whether the regressions that we use for the adjustment, which pool urban and rural households to regress the log unit value on an urban dummy and on the log of per capita total household expenditure, provide the right estimates of urban to rural price differentials. The key issue is that per capita expenditure may not mean the same thing in urban and rural sectors, in part because of the very fact that prices are different, but more fundamentally because the patterns of urban life are different from rural life.

Suppose that the unit value of good i is the sum of a log price component and a log quality component, and that, in the simplest possible case, log quality is a linear function of real income, which is money income deflated by the price index for the sector in which the household lives. If so, we can write

$$\ln v_{ih} = \ln p_{is} + \alpha_i + \beta_i(\ln y_h - \ln P_s) \quad (1)$$

where h is a household, i is the good, s is the sector for urban and rural, v is unit value, y is income, and P_s is the sector price index for all goods, which is taken to be 1 in the rural sector. The sector price for good i , p_{is} , is also taken to be 1 in the rural sector, and is the commodity-specific urban price premium. Equation (1)

can be rearranged to correspond to and to interpret the regressions that we have run in this paper. In particular, we have

$$\ln v_{ih} = (\ln p_{is} - \beta_i \ln P_s) + \alpha_i + \beta_i \ln y_h \quad (2)$$

Equation (2) implies that when we run the unit value regressions on an urban dummy and on the logarithm of income, the estimated dummy is *not* the log of the urban price premium for that good $\ln p_{is}$, but the log of the urban price premium minus a term that is the product of the income elasticity of the unit value and the log of the overall price index of urban prices. Hence

$$\ln \tilde{p}_{is} = \ln p_{is} - \beta_i \ln P_s \quad (3)$$

where the superimposed tilde indicates an estimated amount. In the typical case, where the income elasticity of the unit value is positive, and prices are higher in cities, this will mean that the procedures in this paper underestimate the urban price premium for each good, at least if (1) is the true model. Without knowing what the overall price index is, it is difficult to guess the size of the bias. One example is rice in Table 2, where the estimated price index is 0.14, and the income elasticity of the unit value is 0.26, so that if the true all-good index were twice as high in urban as in rural India, for example, the correct price differential for rice would be 0.14 plus 0.26 log 2, or 0.32, and our estimate is less than a half of the truth.

The overall price index is in part determined by food prices, about which we do know something. In particular, the overall price index can be written

$$\ln P_s = \sum w_i^f \ln p_{is} + (1 - w^f) \ln \pi_s^n \quad (4)$$

where w_i^f is the share of the food i in the total budget, w^f is the share of total food, and π_s^n is the urban price index for non-food. Combining (3) and (4) gives an expression for the overall urban price index

$$\ln P_s = \frac{w^f \ln P_s^T + (1 - w^f) \ln \pi_{sn}}{1 - \sum w_i^f \beta_i} \quad (5)$$

where $\ln P_s^T$ is the log Törnqvist index for food that we have been estimating here. Equation (5) shows that the transition from the calculated food index to the overall index depends not only on the price of the omitted non-foods, but also on the size of the quality corrections. However, if the denominator in (5) is small, which will typically be the case, the formula reduces to the standard one of the weighted average of the estimated food index and the unobserved non-food index.

Differential housing costs are clearly an important part of the overall price index for urban relative to rural. Housing is one of the “comparison resistant” items in the ICP, Deaton and Heston (2010), and standard household income and expenditure surveys are typically not helpful. For example, in India, most rural households report no rent of any kind, and while it is sometimes possible to use survey data to impute user costs, the calculations typically depend on hedonic regressions estimated using a small number of possibly unrepresentative households. In India, the NSS statisticians provide some measure of overall rent for urban households. We have not included this in household total expenditure in the

calculations in this paper, but if we were to do so, as did Dikhanov (2010), the price differentials for the foods would have been estimated to be substantially lower. This is exactly the same effect as shown in equations (2) and (3), but with the sign changed, because the inclusion of urban rents scales up urban incomes.

The general problem here is that we have no way of knowing when an urban income is equivalent to a rural income, in part because we do not have an estimate of the overall cost of living in urban relative to rural, but also because we have no way of allowing for other costs of living in a city, such as the costs of transportation, or even the lower costs of food associated with less physical work undertaken in the cities compared with farmers and agricultural laborers.

One solution is to find an observable indicator, on theoretical grounds, that correctly indicates welfare across the sectors. An imaginative procedure along these lines has been used for by Coondoo, Majumdar, and Chattopadhyay (2011) for India, and by Almås and Johnsen (2010) for China. These papers go back to work by Costa (2001) and by Hamilton (2001) that measures price indexes by assuming that the budget share of food indicates welfare, so that rural households and urban households with the same food shares have the same level of welfare; comparison of their incomes thus estimates the price index. As Coondoo et al argue, this procedure allows the calculation of price indexes without any information on prices, though whether or not this is an advantage or disadvantage is arguable. But there is a long literature that argues that the food share is *not* a good indicator of well-being. For example, if children consume relatively more

food than adults, fully compensated households with more children will have larger food shares than households with fewer children, so that it is not correct to compare food shares across rural and urban sectors if the former have more children. More seriously still, rural households are more likely to do heavy manual labor in agriculture, which requires a large calorie intake, and correspondingly high food shares. Once again, we can imagine two equally well-off households, one rural and one urban, with the former working hard in agriculture, and the latter employed in some unpleasant but relatively sedentary urban occupation, or perhaps just constrained to spend many hours commuting to work. By assumption, they are equally well-off, but the former will have a much larger food share. It would take a very large urban price index to bring the urban worker's food share up to the rural worker's food share, so that the method will overestimate urban prices. The procedure has also been criticized on similar grounds in the poverty line literature by Ravallion and Bidani (1994). Note too that Engel's Law postulates that the food share declines with real income. The proposition that households with identical food shares are equally well-off was also asserted by Engel, and is sometimes referred to as Engel's second law to distinguish it from Engel's first law, from which it does not follow, Deaton and Paxson (1998). But Engel's assertion is false, so that methods based on it do not yield credible results.

It is much easier to see why Engel's assertion is wrong than to propose a good alternative. One possibility is to turn to the direct measures of well-being that have been increasingly explored in both economics and psychology. Gallup's World

Poll has collected self-reported well-being (SWB) data for most countries of the world since 2006 using an identical questionnaire in all countries. I focus here on one specific measure, Cantril's "self-anchoring" ladder question, in which the respondent is asked to imagine a ladder with rungs marked 0 to 10. The bottom rung of the ladder corresponds to "the worst possible life for you" and the top rung to "the best possible life for you," and the respondent is asked to report a value from 0 to 10. The national averages of these ladder values are closely related to the logarithm of per capita GDP in purchasing power terms, Deaton (2008), and are thus a plausible candidate for exploring differences in well-being across space within countries.

Table 9 shows the mean ladder scores for selected large countries, including India and Brazil, but also for the world as a whole. These are unadjusted for any other factors that are correlated with self-reported well-being. Except for Pakistan, people who live in rural areas give lower life evaluations than people who live in any other place. The inhabitants of large cities often show the highest life evaluations, although Russians have higher SWB in the suburbs, and Brazilians and Americans do best in villages or small towns. Averaging over all countries together, those who live in the suburbs report the highest SWB, though there is little difference between them and those who live in large cities.

If people are free to choose where they live, and the costs of moving are negligible, we might expect well-being to be the same across space. But we would not expect this in Table 9 because people who live in different places are different

in other ways too so that, for example, people who live in cities tend to be more educated and to have higher incomes, both of which exert an independent positive effect on SWB; in some countries, the ratio of males to females varies by the degree of urbanization. In order to allow for this, we estimate a regression of SWB on education, the logarithm of income, sex, age groups, and dummies for each of the three urbanization categories, treating rural as the base category. The results are shown in Table 10. If we were to assume that the ladder measure of SWB is an indicator of real income, we could calculate the price level in cities relative to rural areas by dividing minus the coefficient on large city by the coefficient on log income. According to this procedure, city prices are 6 percent higher in Brazil, 17 percent higher in India, 20 percent higher in Russia, and nearly 50 percent higher in the US. But for China, Indonesia, Nigeria and South Africa we would be forced to conclude that urban prices are actually lower, in South Africa by an absurdly large factor.

The problem with these calculations is that self-reported well-being—if the concept is accepted at all—refers to something much broader than real income, and in particular includes all of the amenities that come with cities, or more broadly, with living near many other people. From such a perspective, the results in Table 10 make more sense, with relatively few significant effects—which we would expect if people are mobile and we have adequately controlled for individual human capital—but it takes us no further in isolating the real income component of well-being. South Africa is an exception here, with large difference

across areas, and perhaps not surprisingly so given that there remains a great deal of racial segregation over space, even after the formal rules enforcing it have been removed.

The self-reported well-being calculations, like calculations based on food shares, take us no further forward in the search for urban to rural price levels, at least beyond the price indexes for food presented earlier in this paper. Perhaps the differences in urban and rural life in the structure of demand, and in what commodities and services are used for—especially transportation, commutation, and rents—make the search for a non-food price index quixotic. And even if we could measure real income adequately, and following the arguments of Glaeser (1998), it would remain unclear whether it would be appropriate to use those measures of real income as the basis for compensatory transfers or other aspects of anti-poverty policy.

9. Citations

Almås, Ingvild, and Ashild Auglænd Johnsen, 2010, "China's growth miracle—still awaiting the great leap forward," available at <http://www.iariw.org>

Aten, Bettina, Eric Figueroa, and Troy Martin, 2011, "Notes on estimating the multi-year regional price parities by 16 expenditure categories: 2005-2009," Bureau of Economic Analysis, April.

Citro, Constance E., and Robert T. Michael, eds., 1995, *Measuring Poverty: a new approach*, Washington, DC. National Academies Press.

Coondoo, Dipankor, Amita Majumder and Somnath Chattopadhyay, 2011, "Estimating spatial consumer price indices through Engel curve analysis," *Review of Income and Wealth*, 57(1), 138-55.

- Costa, Dora, 2001, "Estimating real income in the United States from 1888 to 1994: correcting CPI bias using Engel curves," *Journal of Political Economy*, 109(6), 1288-1310.
- Deaton, Angus, 1988, "Quality, quantity, and spatial variation in price," *American Economic Review*, 78 (3): 418-30.
- Deaton, Angus, 1997, *The analysis of household surveys*, Johns Hopkins University Press for the World Bank.
- Deaton, Angus, 2008, "Income, health, and well-being around the world: evidence from the Gallup World Poll," *Journal of Economic Perspectives*, 22:2, 53-72.
- Deaton, Angus and Alan Heston, 2010, "Understanding PPPs and PPP-based national accounts," *American Economic Journal Macroeconomics*, 2(4), 1-35.
- Deaton, Angus, and Christina Paxson, 1998, "Economies of scale, household size, and the demand for food," *Journal of Political Economy*, 106(5), 897-930.
- Deaton, Angus and Alessandro Tarozzi, 2005, "Prices and poverty in India," Chapter 16 in Angus Deaton and Valerie Kozel, eds., *The great Indian poverty debate*, New Delhi, Macmillan. 381-411.
- Dikhanov, Yuri, 2010, "Income-effect and urban-rural price differentials from the household survey perspective," ICP Global Office, February.
- General Accounting Office, 1997, *Poverty measurement: issues in revising and updating the official definition*, Report to the ranking minority member, committee on finance, U.S. Senate.
- Glaeser, Edward L., 1998, "Should transfer payments be indexed to local price levels?" *Regional Science and Urban Economics*, 28:1, 1-20.
- Government of India, 1993, *Report of the expert group on the on the estimation of proportion and number of poor*, Delhi, Planning Commission.
- Government of India, 2009, Report of the expert group to review the methodology for the estimation of poverty, Planning Commission, November.
- Hamilton, Bruce W., 2001, "Using Engel's law to estimate CPI bias," *American Economic Review*, 91(3), 619-30.
- Hill, Robert J., and Iqbal Syed, 2010, "Improving international comparisons of real output: the ICP benchmark and its implications for China," University of New

South Wales, Australian School of Business, Research Paper No. 2010 ECON 25,
<http://ssrn.com/abstract=1728428>

Ravallion, Martin and Benu Bidani, 1994, "How robust is a poverty profile?" *World Bank Economic Review*, 8(1), 75-102.

Renwick, Trudi, 2011, "Geographical adjustments of supplementary poverty measure thresholds: using the American Community Survey five year data on housing costs," U.S. Census Bureau, working paper.

Roback, Jennifer, 1982, "Wages, rents, and the quality of life," *Journal of Political Economy*, 90:6, 1257-78.

Short, Kathleen, 2010, "Who is poor? A new look with the supplemental poverty measures," Washington, DC. U.S. Census Bureau, SEHSD Working Paper 2010-15.

Short, Kathleen, Thesia Garner, David Johnson, and Patricia Doyle, 1999, *Experimental poverty measures: 1990-1997*, Washington, DC. U.S. Census Bureau.

Table 1

Budget shares, median and mean log unit values for urban and rural India

	Budget Shares in total expenditure		Median log unit value			Mean log unit values		
	Rural	Urban	Rural	Urban	Difference	Rural	Urban	Difference
Rice	12.99	7.32	2.28	2.48	0.21	2.23	2.51	0.28
Milk	7.40	7.44	2.48	2.64	0.15	2.45	2.64	0.19
Wheat/atta	6.38	5.28	1.97	2.30	0.33	2.01	2.27	0.26
Mustard oil	2.63	1.46	4.01	4.01	0	4.00	4.00	-0.00
Sugar	2.24	1.84	2.89	2.89	0	2.89	2.89	0.00
Potatoes	1.73	1.01	1.95	2.08	0.13	1.87	1.99	0.11
Other oil	1.51	1.55	3.91	3.95	0.04	3.91	3.97	0.06
Fish	1.37	1.00	3.69	3.91	0.22	3.68	3.86	0.18
Arhar	1.24	1.12	3.40	3.42	0.02	3.37	3.42	0.06
Tea (leaf)	1.05	1.04	-2.04	-1.97	0.07	-2.04	-1.96	0.08
Tea (cups)	0.81	1.21	0.41	0.69	0.29	0.45	0.71	0.26
Total/Mean	39.36	30.28			0.17			0.19

Notes: The budget shares are shares in total expenditure, averaged over all households in each sector. The sums in the last row are for these eleven goods only. The means at the foot of the two difference columns are the averages of the differences weighted by the averages of the rural and urban budget shares in columns 1 and 2. Purchases made in fair price shops are included in the rice, wheat/atta, and sugar categories.

Table 2

Income effects on unit values and adjusted urban rural differences: selected goods

	Regression coefficient on lnpce	Regression coefficient on urban dummy	Unadjusted difference from Table 1
Rice	0.262 (154.2)	0.138 (60.7)	0.284 (124.9)
Milk	0.082 (44.0)	0.153 (64.9)	0.194 (88.7)
Wheat/atta	0.174 (73.9)	0.161 (52.1)	0.256 (88.6)
Mustard oil	-0.008 (-6.7)	0.002 (1.2)	-0.002 (-1.5)
Sugar	-0.003 (-3.0)	0.003 (2.4)	0.002 (1.3)
Potatoes	0.084 (40.4)	0.068 (24.5)	0.115 (45.0)
Other oil	0.072 (52.7)	0.016 (8.4)	0.062 (35.3)
Fish	0.106 (29.9)	0.124 (25.0)	0.184 (40.4)
Arhar	0.037 (37.9)	0.035 (28.1)	0.055 (48.5)
Tea (leaf)	0.069 (35.9)	0.046 (18.5)	0.081 (35.4)
Tea (cups)	0.179 (57.7)	0.165 (41.8)	0.262 (71.3)

Notes: Each row comes from a regression of the logarithm of unit value on the logarithm of household per capita expenditure and on an urban dummy. Absolute t -values are shown in parentheses. The regressions are run at the household level, with one observation for each household recording a purchase of the good. The third column is an estimate of the extent to which urban prices are higher than rural prices, and can be compared with the unadjusted difference in means in the final column, taken from the last column of Table 1.

Table 3: Budget shares (percent) and mean log unit values by source and sector

	RURAL				URBAN			
	Cash purchase		Other		Cash purchase		Other	
	Share	Ln uv	Share	Ln uv	Share	Ln uv	Share	Ln uv
Rice from PDS	1.03	1.55	--	--	0.48	1.56	--	--
Rice market	7.60	2.36	4.36	2.22	6.50	2.58	0.35	2.34
Milk	3.00	2.49	4.40	2.40	6.81	2.65	0.63	2.50
Wheat PDS	0.34	1.38	--	--	0.16	1.61	--	--
Wheat market	3.69	2.17	2.36	1.91	4.84	2.31	0.28	1.97
Mustard oil	2.44	4.01	0.19	3.87	1.44	4.00	0.02	3.90
Sugar PDS	0.18	2.62	--	--	0.11	2.62	--	--
Sugar market	2.06	2.92	0.01	2.92	1.73	2.91	0.01	2.98
Fish	1.12	3.69	0.25	3.59	0.96	3.87	0.04	3.53

Table 4

Income effects on unit values and adjusted urban rural differences: selected goods
Cash purchases only and excluding PDS

	Regression coefficient on Lnpce	Regression coefficient on urban dummy	Unadjusted difference
Rice	0.225 (150.6)	0.088 (45.2)	0.216 (110.2)
Milk	0.089 (43.5)	0.141 (55.2)	0.187 (79.5)
Wheat/atta	0.137 (61.3)	0.121 (42.1)	0.198 (74.3)
Mustard oil	-0.007 (-5.3)	0.003 (1.9)	-0.001 (-0.5)
Sugar	-0.009 (-9.4)	0.001 (0.5)	-0.004 (-4.0)
Potatoes	0.095 (42.6)	0.028 (9.8)	0.083 (31.5)
Other oil	0.074 (53.5)	0.015 (7.9)	0.062 (35.4)
Fish	0.088 (21.5)	0.179 (32.0)	0.231 (45.8)
Arhar	0.036 (36.1)	0.031 (25.0)	0.051 (45.1)
Tea (leaf)	0.075 (35.1)	0.042 (15.8)	0.082 (33.3)
Tea (cups)	0.186 (55.4)	0.129 (30.4)	0.233 (59.0)

Notes: Each row comes from a regression of the logarithm of unit value on the logarithm of household per capita expenditure and on an urban dummy. Absolute t -values are shown in parentheses. The regressions are run at the household level, with one observation for each household recording a purchase of the good. The third column is an estimate of the extent to which urban prices are higher than rural prices, and can be compared with the unadjusted difference in means in the final column, taken from the last column of Table 1. Rice, wheat/atta, and sugar exclude purchases from the PDS.

Table 5

Differences in log median unit values (urban minus rural) by decile groups

Decile	rice	milk	wheat	mustard oil	sugar	All food
1	0.061	0.182	0	-0.036	0	0.044
2	0.099	0.182	0.134	-0.018	0	0.077
3	0.105	0.167	0.134	-0.036	0	0.083
4	0.052	0	0.194	-0.036	0	0.062
5	0.025	0.080	0.251	-0.007	0	0.070
6	0.095	0.154	0.182	-0.018	0	0.082
7	0.182	0.154	0.118	0	0	0.093
8	0.175	0.154	0.223	0	0	0.102
8	0.080	0.154	0.223	0.018	0	0.094
10	0.223	0.223	0.258	0.036	0	0.151
Mean	0.110	0.145	0.172	-0.010	0	0.086

Notes: Deciles are defined using the sampling weights, so that each decile contains an equal share of the All India population. The numbers shown are the differences in median log unit values between urban and rural within each decile. The averages in the bottom row are simple averages over the deciles. The final column uses all 134 foods, with the difference in median unit values weighted by the average of the urban and rural budget shares.

Table 6

Törnqvist GEKS multilateral food price indexes by state and sector

	Without correction			With correction		
	Rural	Urban	Ratio	Rural	Urban	Ratio
Jammu & Kashmir	1.00	1.09	1.09	1.00	1.05	1.05
Himachal Pradesh	0.99	1.14	1.15	0.99	1.08	1.10
Punjab	0.96	1.08	1.12	0.95	1.03	1.09
Uttaranchal	0.93	1.05	1.13	0.94	1.03	1.09
Haryana	0.98	1.08	1.09	0.99	1.04	1.06
Delhi	..	1.19	1.12	..
Rajasthan	0.94	1.04	1.11	0.95	1.01	1.07
Uttar Pradesh	0.84	0.99	1.17	0.87	0.98	1.12
Bihar	0.86	0.96	1.12	0.91	0.97	1.06
Assam	1.02	1.18	1.16	1.05	1.14	1.08
West Bengal	0.90	1.08	1.20	0.93	1.04	1.11
Jharkhand	0.87	1.04	1.19	0.93	1.01	1.08
Orissa	0.83	0.95	1.14	0.90	0.95	1.06
Chhatisgarh	0.85	0.99	1.16	0.91	0.97	1.06
Madya Pradesh	0.83	0.98	1.18	0.87	0.97	1.12
Gujarat	1.00	1.17	1.17	1.02	1.13	1.11
Maharashtra	0.95	1.18	1.24	0.97	1.14	1.17
Andhra Pradesh	0.91	1.03	1.13	0.93	1.00	1.08
Karnataka	0.82	1.02	1.24	0.85	1.00	1.18
Kerala	0.99	1.06	1.06	0.97	1.02	1.05
Tamil Nadu	0.92	1.08	1.17	0.94	1.04	1.10

Notes: Small states and rural Delhi excluded. The first panel, without correction, shows multilateral indexes calculated from the means of the logs of reported unit values, averaged by state and sector. The second panel is based on a set of regressions in which the logarithms of the unit values for each good are regressed on the logarithm of per capita total expenditure and a fully interacted set of state and sectoral dummies. These regressions are then used to calculate hypothetical mean log unit values for each sector and state assuming, counterfactually, that each has the same mean log per capita total expenditure. This correction, in the second panel, tends to reduce both the sectoral and spatial differentials. Within each panel the column labeled ratio shows the urban to rural differential for each state from the PPPs in the first and second columns.

Table 7: Budget shares and log unit values by sector, Brazil 2002-03

	Rural share	Urban share	Urban premium	Adj. Premium	Cash premium	Adj. cash premium
Polished rice	1.72	0.97	2.83	-0.03	0.29	-2.40
Soybean oil	1.25	0.58	-6.79	-5.90	-6.66	-5.77
French bread	0.52	1.25	6.20	0.37	5.96	0.02
Crystal sugar	1.28	0.44	-5.25	-4.04	-5.23	-3.84
Fresh milk	1.40	0.26	23.11	22.90	14.09	12.63
Cassava flour	1.13	0.22	18.55	12.34	12.18	7.02
Ground coffee	0.86	0.42	-4.85	-6.46	-6.36	-8.11
Choice beef	0.83	0.38	4.65	2.91	3.20	0.96
Carioca beans	0.81	0.33	7.42	4.57	3.28	0.25
Frozen chicken	0.61	0.51	-5.52	-5.38	-5.87	-5.76
Prime beef	0.61	0.35	9.01	5.94	5.17	1.15
Index over 38 largest share foods	15.96	9.31	4.90	2.89	2.08	-0.03

Notes: All figures are percentages. The first two columns show the average shares of each item in total household expenditure. The third column shows the differences in the means of the logs of unit values, while the fourth column shows the same numbers adjusted for household per capita expenditure and household size. The final two columns replicate these calculations using unit values only for cash transactions. The bottom row of the table shows the Törnqvist indexes for 38 of the most important foods in the survey.

**Table 8: Törnqvist multilateral GEKS food price indexes by region and sector
Brazil, 2002-03**

	Unadjusted			Adjusted		
	Rural	Urban	Ratio	Rural	Urban	Ratio
North	1.000	1.050	1.050	1.000	1.029	1.029
North East	1.001	1.024	1.023	1.007	1.008	1.001
South East	0.999	1.042	1.042	0.979	0.998	1.019
South	0.926	0.996	1.075	0.898	0.954	1.062
Centre West	1.001	1.028	1.027	0.975	0.996	1.021

Notes: The first panel, without correction, shows multilateral indexes calculated from the means of the logs of reported unit values, averaged by region and sector. The second panel is based on a set of regressions in which the logarithms of the unit values for each good are regressed on the logarithm of per capita total expenditure and a fully interacted set of state and sectoral dummies. These regressions are then used to calculate hypothetical mean log unit values for each sector and region assuming, counterfactually, that each has the same mean log per capita total expenditure. This correction, in the second panel, tends to reduce the sectoral differentials. Within each panel the column labeled ratio shows the urban to rural differential for each state from the PPPs in the first and second columns. Calculated using the 38 most important foods in the survey.

Table 9: Mean ladder scores by place of residence, selected large countries and world

	Rural	Village or small town	Suburb of large city	Large city
Brazil	6.59	6.90	6.71	6.65
China	4.51	4.68	4.97	5.22
India	4.67	5.06	5.14	5.20
Indonesia	4.97	5.16	5.17	5.44
Nigeria	4.74	4.95	4.97	5.08
Pakistan	5.34	5.22	5.11	5.56
Russia	5.10	5.18	5.71	5.54
South Africa	4.47	5.41	5.34	6.17
USA	7.29	7.32	7.30	7.10
World	4.94	5.47	5.87	5.82

Notes: The Gallup data are from 2006 to 2010, and all years are pooled for the calculations. Not all countries are sampled in all years, so that the world total will contain possible country selection effects. The means are calculated using within country sampling weights. The world total is computed taking each country as a unit, and is not weighted by population.

Table 10: Coefficients of urbanization and log income on self-reported well-being

	Village or small town	Suburb of large city	Large city	Log income
Brazil	0.036	0.275	-0.024	0.395**
China	-0.217"	-0.007	0.011	0.472**
India	0.112**	0.013	-0.119*	0.734**
Indonesia	-0.009	-0.169*	0.011	0.395**
Nigeria	0.049	0.150	0.257*	0.058
Pakistan	-0.060	-0.546**	-0.140	0.722**
Russia	-0.207**	0.341	-0.057	0.608**
South Africa	0.609**	0.395**	1.372**	0.377**
USA	0.017	-0.146	-0.234*	0.471**
World	-0.007	0.001	0.068**	0.512**

Notes: The Gallup data are from 2006 to 2010, and all years are pooled for the calculations. Not all countries are sampled in all years, so that the world total will contain possible country selection effects. The regressions are run using within country sampling weights. The world results come from a regression that includes country dummies. All regressions include three education categories, age group, and gender controls.