Can caloric needs explain three food consumption puzzles? Evidence from India

Shari Eli and Nicholas Li*

University of Toronto

October 21, 2012 – Version 1.0
Preliminary

Abstract

We argue that differences in caloric needs can help explain three food consumption puzzles that have been noted in the literature. The first puzzle (Deaton and Dreze (2009)) is the stagnant or declining caloric intake of Indian household despite significant economic growth between 1983-2005. The second puzzle (Deaton and Paxson (1998)) is the negative relationship between household size and caloric intake per capita after controlling for total expenditure per capita, which is inconsistent with basic theories of household scale economies. The third puzzle is the substantial decline in caloric intake for older Indian households which seems to suggest the absence of any consumption-smoothing consistent with the permanent-income life-cycle hypothesis, a decline that is at odds with results for the United States where there is little or no decline. We combine data on household food consumption with a novel measure of caloric needs based on time-use data and anthropometric measures of net nutritional outcomes to shed light on all these puzzles. Our results indicate that caloric needs can explain a substantial part of all three puzzles, and that consequently attempts to use food and caloric intake to measure household welfare – which are very common in the literature using Engel's law or absolute poverty lines based on caloric intake – need to carefully control for substantial variation in caloric needs across households. When caloric needs vary substantially, lower caloric intake can actually correspond to an increase in welfare as it frees up resources for higher food quality and non-food expenditures.

*We would like to thank Pierre-Olivier Gourinchas, Chang-Tai Hsieh, Ronald Lee, and Ted Miguel for their comments and suggestions. We also thank Pranab Bardhan, Ethan Ligon, Elizabeth Sadoulet, as well as participants of the Berkeley Development Lunch and Seminar and University of Toronto SWEAT seminar for helpful comments. Lucas Parker provided outstanding research assistance. Nicholas Li gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada, the UC Berkeley Institute for Business and Economics Research, and the Center for Equitable Growth. Shari Eli gratefully acknowledges the financial support of the NICHD grant T32-HD007275 and NIA grant T32-AG000246. All errors are ours.

1. Introduction

Food consumption is widely used to measure poverty and welfare. A tradition going back to Engel's work on household budgets in England has argued that a lower consumption share of food corresponds to greater welfare as it indicates that the most basic need is satisfied. The relationship between food consumption and incomes – Engel's law – has been applied to analyze household welfare over time (Costa (2001), Hamilton (2001)), across countries (Almas (2012)), and to calculate household economies of scale – equivalence scales – across households of different sizes (Barten (1964), Deaton and Paxson (1998), Lanjouw and Ravallion (1995)). A more direct measure of the satisfaction of basic needs is caloric intake, which is often used to calculate absolute poverty lines as well as to infer whether households are able to smooth consumption over the life-cycle (Aguair and Hurst (2005), Hicks (2010)). Other welfare measures based on food consumption patterns have been proposed, such as the calorie-expenditure elasticity (Logan (2009)) and staple-share of calories (Jensen and Miller (2010)).

Given the strong posited link between food consumption welfare, there are three widely-noted empirical patterns that pose a puzzle for basic economic theory and measurement. The first puzzle, noted in India (Deaton and Dreze (2009)), China (Du et al. (2002)) and England during the Industrial Revolution (Clark et al. (1995)), is that societies undergoing rapid growth and structural transformation have often seen stagnant caloric intake despite rapid growth of incomes. This is puzzling because cross-sectional data from poor countries always reveals a strong positive relationship between household expenditure and caloric intake, but expenditure growth over time is not reflected in higher caloric intake. This general pattern also appears to be exist when comparing urban and rural sectors within a country or rich and poor countries – the richer and more developed the area, the lower is caloric intake for a household holding real expenditures constant.

The second puzzle is that holding per capita expenditure constant, larger households have lower food expenditure and lower caloric intake than smaller households. While larger households are often poorer in developing countries, when total resources per person are held constant we would expect some scale economies for the household that would result in greater consumption of income-elastic private goods like food with no close substitutes. The classic model of Barten (1964) predicts that larger households would be better off and would consequently consume more food per capita, a prediction completely at odds with the data for numerous countries (Deaton and Paxson (1998)), data that seem to imply that household scale economies lead to worse nutrition. Deaton and Paxson (1998) find that the negative relationship between household size and private good expenditure per capita is stronger for food and for poor countries, suggesting that there may be something special about food demand in poor countries. Are larger households really worse off nutritionally because of their lower food expenditures and caloric intake per capita, and how should we think about equivalence scales in this context?

The third puzzle is that food expenditure falls significantly for older households, contradicting the permanent-income life-cycle hypothesis that predicts smooth consumption despite income that varies over the life-cycle. Aguair and Hurst (2005) argue that in the United States this pattern does not reflect a failure of the permanent-income life-cycle hypothesis, which applies to consumption and not expenditure – upon retirement, older US households substitute intensive shopping and home cooking for meals out to maintain the quality and caloric-intake of their diet despite a sharp drop in expenditures. Hicks (2010) finds only a small (2.7%) drop in caloric-intake for Mexican households over age 55. In India, however, caloric-intake falls sharply beginning in the 50s, suggesting that no such smoothing occurs. Does this reflect a breakdown of the permanent-income life-cycle hypothesis, perhaps due to financial frictions and psychological forces (e.g. lack of savings vehicles or inability to anticipate future income or commit to future consumption), or could it be due to changes in caloric needs?

In this paper we argue that a single factor – differences in caloric needs across households – can help explain these three puzzles. We use time-use data from India to create an original measure of caloric needs, and combine this with data on food consumption and anthropometrics (height/weight and BMI) to explore the extent to which caloric needs can help predict differences in food consumption. We show that caloric needs vary considerably with occupation, access to household facilities like electricity, gas, and water, household size and age, and that the differences in caloric needs generated by these factors co-vary strongly with caloric intake and are consistent with observed differences in adult weight. Our results imply that decreasing food expenditure and caloric intake should often be interpreted as raising household welfare by freeing up resources for higher quality food or non-food expenditures; in poor countries, caloric intake is only a monotonic predictor of household welfare *conditional* on caloric needs.

We find that movement out of physically demanding sectors like agriculture and manufacturing combined with access to electricity, gas and water (which provide access to labor-saving appliances and lower the need for firewood and water collection by households) can explain all of the differences in rural versus urban food consumption patterns in India, but only about a third of the difference that takes place over the 1983-2005 period, suggesting that alternative explanations (changes in relative prices and the introduction of new consumer goods) may help explain why some rapidly-growing countries have not experienced a large increase in caloric intake per capita. Larger households are able to lower their caloric requirements through specialization and household scale economies - they lower the amount of market work required to generate a given level of expenditures and they also decrease the amount of home production, leading to significantly higher leisure per member. Older households in India experience a sharp drop in caloric intake but an even sharper drop in caloric needs, driven in equal parts by biological metabolism and a reduction in activity levels; while the decline in metabolism is an obvious explanation for falling caloric intake over the life-cycle, in richer countries the decline in physical activity levels is likely to be less of a factor, and the smoothing behavior observed by Aguair and Hurst (2005) and Hicks (2010) leads to a larger gap between caloric intake and needs and hence greater weight gain.

The paper is organized as follows: Section 2 lays out the three food consumption puzzles in more detail. Section 3 provides a simple theoretical framework. Section 4 describes the data. Section 5 presents our main results, and Section 6 concludes.

2. Three Food Puzzles

2.1. 1."Growth–calorie puzzle" (Deaton and Dreze (2009))

Statement: Caloric intake per capita and total expenditures per capita are positively related in the cross-section in India and other poor countries. This suggests that economic growth will raise caloric intake. However, demand for food and especially calories appears to be stagnant during several periods of rapid economic growth, consistent with a downward shift in the calorie-total expenditure relationship (the "calorie Engel curve"). A similar downward shift in demand for calories

and food has been observed comparing richer and poorer areas in the cross-section, i.e. urban versus rural or rich versus poor country.

Figure 1 provides a graphical depiction of this "puzzle" for India. Within a particular sector (rural or urban) and a particular year (1983 or 2005) we observe an upward-sloping relationship between caloric intake per capita and total expenditures per capita ("calorie Engel curve") for five person households, which is not surprising given the high-levels of poverty and undernutrition observed in India. What is surprising is that despite a growth in real total expenditures per capita of over 30% between 1983 and 2005, caloric intake per capita actually declined in rural areas and was roughly constant in urban areas. Figure 1 reveals that these two facts are reconciled mechanically by the downward shift in the calorie Engel curves for both rural and urban areas. These calorie Engel curves imply that some factor is driving down demand for food and calories over this period across the expenditure distribution (though perhaps more for higher-income households). One potential clue provided by figure 1 comes from comparing rural and urban areas – the calorie-Engel curves for urban areas are shifted down relative to those for rural areas in the same year, such that the urban calorie-Engel curve for 1983 looks similar to the one for rural areas in 2005. This suggests that factors that differ across rural and urban India but also vary over time – differences in the physical intensity of occupations like farming, manufacturing, and services or in access to electricity, gas for cooking/heating and piped water – could be driving the shifts in calorie Engel curves.

While the pattern of stagnant food expenditures and caloric intake despite economic growth revealed by these Indian data could be an artifact of measurement, the result appears to be quite robust to various assumptions and imputations regarding likely sources of measurement error like meals to guests, meals taken outside of the household, and increased consumption of restaurant meals and processed food (see the appendix for details). Deaton and Dreze (2009) note that a similar pattern of caloric decline been observed using an alternative data set that uses a shorter recall window and more direct measurement of food intake. This Indian pattern of economic growth and stagnant caloric intake has also been observed for China in recent years (Du et al. (2002)) and England during the Industrial Revolution (Clark et al. (1995)).

A similar and perhaps related phenomenon has been observed when comparing food Engel curves over time for the United States (Costa (2001), Hamilton (2001))

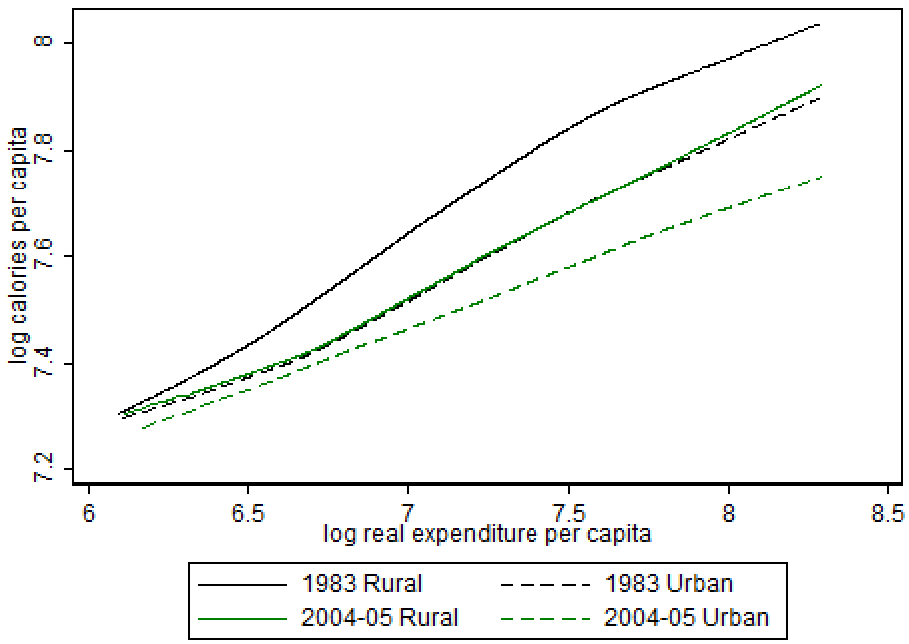

Figure 1: Puzzle 1: Calorie Engel curves by year and sector (five person households)

or across countries (Almas (2012)) – richer countries and later periods seem to have downward shifted food budget shares conditional on real expenditures. These studies interpret the downward shift as evidence of mismeasurement in official price indexes (the CPI or Penn World Tables), which may overstate the rate of price increases over time (or the higher prices in rich countries) for several reasons (e.g. product quality, new varieties, outlet and substitution bias). According to this logic, richer countries and later periods should have lower food budget shares because they are further along the same, overlapping Engel curve that holds for all countries and periods due to higher real expenditures - the downward shift we observe is just an artifact of mismeasurement of prices, once we have controlled for relative price differences that could also shift the food budget share conditional on real expenditure. While mismeasurement of the price level can conceivably explain the patterns in figure 1, the pattern holds up using price indexes calculated directly from the household survey data used to calculate caloric intake or food budget shares, even after correcting for quality-effects on unit values and chaining the price indexes. This suggests that another factor may be at play.

Following the suggestion in Deaton and Dreze (2009) we explore whether differences in caloric needs over time and across rural and urban India can explain the shifts in figure 1, focusing on the role of physical activity levels related to occupation and household facilities like electricity, gas, and piped water. We do not delve deeply into various alternative explanations, but provide some evidence later that shifts in the relative price of food and non-food may have played some role, and that the shifting composition of the overall household budget is consistent with the introduction and availability of new consumer goods, the availability of complementary goods (like electricity), and the perceived higher return of certain "investment" expenditures like education.

2.2. 2. "Household size–calorie puzzle" (Deaton and Paxson (1998))

Statement: Caloric intake per capita is negatively related to household size *conditional on total expenditure per capita and household demographics* in India and other poor countries. This seems to imply worse nutritional outcomes for larger households. However, economic theory predicts that in many cases household scale economies should raise household welfare in part by increasing per capita consumption of private goods. The Barten (1964) model notes that as household size increases, the price of more public/shared goods falls generating a negative substitution effect and positive income effect on the demand for private goods. As food expenditure (and caloric intake) has a relatively large income elasticity and few substitutes in poor countries, we would expect to see this positive relationship between household size and caloric intake per capita (conditional on total expenditure per capita) most clearly in poor countries; instead, we see that the relationship is the strongest and most negative for food (compared to other private goods) and poor countries (compared to rich ones).

Figure 2 provides a graphical depiction of our second puzzle using Indian data from 1983. Holding household size constant, households with higher expenditures per capita consume more calories per capita as we expect. Holding expenditures per capita constant, we see that households with more adults consume less calories per capita – their calorie Engel curves are shifted downward for any level of total expenditure per capita. This is exactly the opposite of what the Barten (1964) model would predict if the expenditure elasticity of calories is very positive (which it clearly is from the figure) and the lower effective price of public/shared goods does not lead to large substitution away from all private goods.

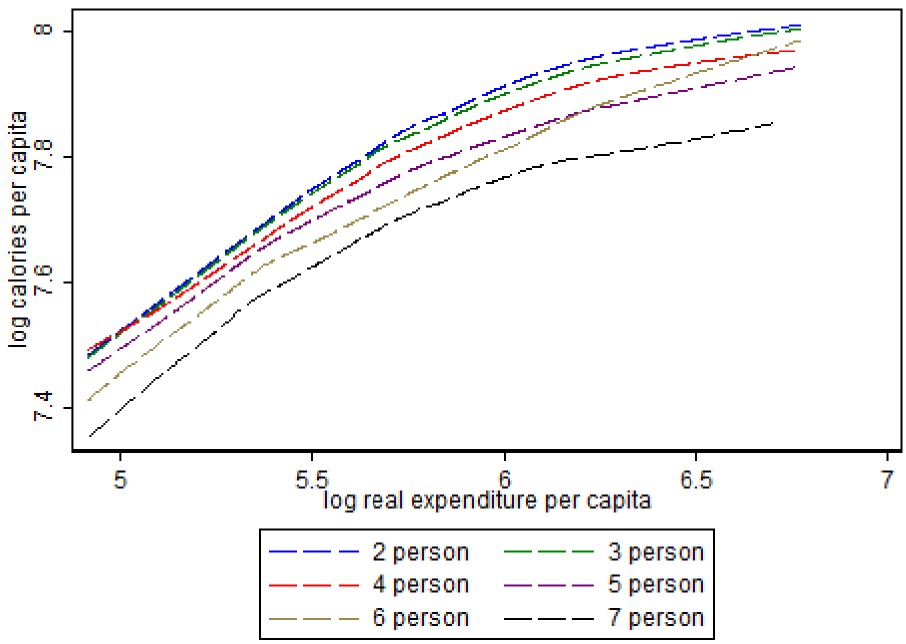

Figure 2: Puzzle 2: Calorie Engel curves for different sized households (all adult), 1983

The most obvious explanations for this figure – that larger households have different demographic structures including more children or different numbers of females – turns out not to be the case, as this result holds controlling for detailed household demographics. Moreover, this result has been documented (for food expenditures per capita rather than calories per capita) in many different developing countries like Thailand, Pakistan, and South Africa by Deaton and Paxson (1998). While it is weaker in richer countries like Taiwan, France and Great Britain, it also appears to hold quite strongly in the 19th century United States (Logan (2008)). Gardes and Starzec (1999) find a similar result using household panel data from Poland. An potentially important finding from Deaton and Paxson (1998) is that the negative relationship between household size and food does not appear to hold for other private goods like clothing, entertainment services, or alcohol and tobacco, suggesting that the empirical failure of the Barten (1964) model may be due to the special nature of food consumption rather than a more fundamental failure of the models predicted income and substitution effects on private good consumption.

Deaton and Paxson (1998) provide a long list of possible explanations for the empirical puzzle they document, including economies of scale in food purchase (bulk-discounting) and in preparation, wastage, collective household models, a large price elasticity for food, measurement error, intra-household inequality and what they call "calorie overheads." Calorie overheads correspond to what we call "caloric needs" and we pursue this as a possible explanation for the puzzle. The idea is that larger households are able to achieve the same level of income per capita (or correspondingly expenditure per capita) and the same level of home production with less physical exertion, either through specialization of household members or through direct time-use/caloric need scale economies (e.g. the caloric need per capita rises by less for a seven member than a two member household when one person in the household goes to the forest to gather and chop firewood for the days cooking needs). However, we do not rule out other possible explanations. Recent studies have explored the contribution of bulk-discounting (Abdulai (2003)), parametric misspecification (Perali (2001)) and systematic measurement error and recall bias (Gibson and Kim (2007)), but neither of these factors seems able to resolve the puzzle on their own – our findings in terms of calories per capita (rather than food expenditures per capita) would seem to rule out bulk-discounting, and while the results on measurement error in Gibson and Kim (2007) are quite convincing, the puzzle persists in their data even using the most detailed, diary-based food expenditure surveys.

2.3. 3. "Life-cycle–calorie puzzle" (Aguair and Hurst (2005))

Statement: Older households experience a sharp decline in caloric intake per capita in India. The permanent-income life-cycle hypothesis predicts that households smooth consumption across periods given their expected (permanent) income. The large decline in food expenditure upon retirement in rich countries has been interpreted by Aguair and Hurst (2005) as consistent with the permanent-income life-cycle hypothesis because the decline in food expenditure by most retired households does not correspond to a decline in food consumption; households substitute time-intensive shopping and home production (cooking) to maintain the quality and (caloric) quantity of food they were consuming. The large decline in caloric intake in India appears to be inconsistent with this hypothesis and requires an alternative explanation to the one offered by Aguair and Hurst (2005) for the United States.

Figure 3 presents the coefficients from a regression of total expenditures, food expenditures, and caloric intake on average adult age for rural 2-adult households in India. We separate life-cycle/age effects from cohort effects by pooling the 1983, 1987-1988, 1993-1994, 1999-2000, and 2004-2005 NSS survey rounds. The data show a steep decline in total expenditures, food expenditures and caloric intake beginning around age 50. Based on the evidence provided by Aguair and Hurst (2005), a similar figure for the United States would show a similar steep decline in total expenditures and food expenditures beginning a bit later, but would show little or no decline in calorie intake. Hicks (2010) finds that caloric intake declines by 2.7% for Mexican households after age 55. In India, the decline in caloric intake is much larger – as high as 10% – and is of a similar magnitude to the decline in food and total expenditures. That Indian households do not use the consumption smoothing mechanisms used by American households is not surprising – consumption of food outside the home is very low to begin with, offering little margin for adjustment upon retirement or when market hours decline in late life, and the returns to intensive shopping behavior may be lower in a relatively undeveloped retail environment – but it raises the question of why consumption smoothing appears to fail for caloric intake.

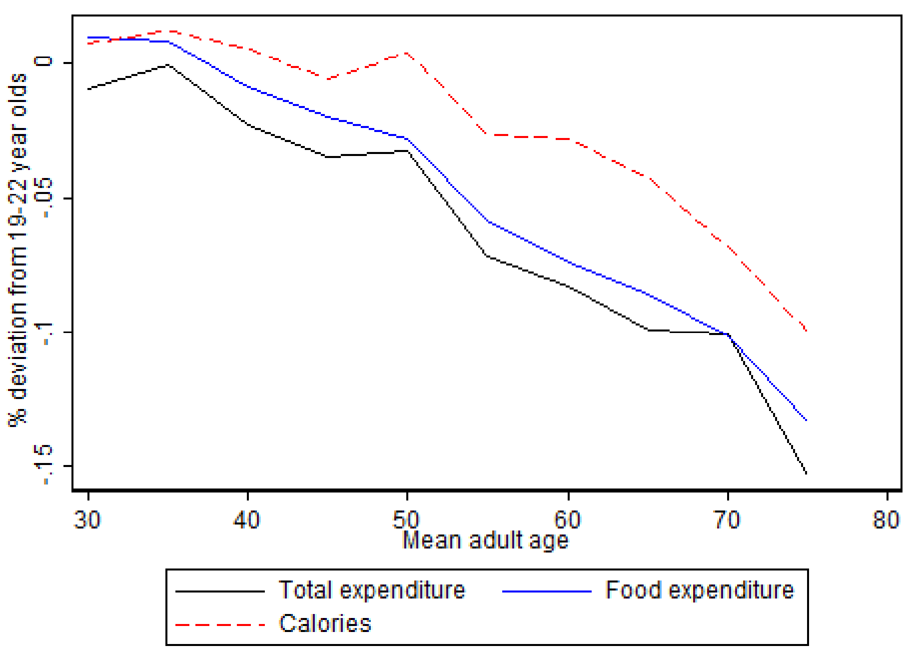

Figure 3: Puzzle 3: % deviations from 19-22 year olds by age group (1983-2005, rural 2-adult households)

We explore the possibility that the pattern observed in figure 3 is driven by a fall in caloric needs in late life. While it is well known that metabolism declines with age resulting in lower caloric needs for basic bodily functions, this factor does not seem to lead to a significant decrease in caloric intake in the United States and Mexico where the metabolic decline is presumably similar. We posit that in addition to the direct effect of age on metabolism and basic caloric needs, changes in physical activity levels also play a larger role for Indian households than for those in richer countries. If market and home production activities are more physically intensive in India, and there is a reduction in these activities in late life (either due to lack of opportunity or declining ability and capacity for physically demanding work), we would expect to see a larger decline in caloric needs over the life-cycle which may in turn lead to a larger decline in caloric intake and food expenditures. While there is little literature on household consumption and expenditure over the life-cycle in developing countries, we do not rule out other possible explanations. The most likely alternative explanations include a lack of savings vehicles and instruments, present-biased preferences or other psychological factors that impede saving behavior, and a larger role for idiosyncratic and unanticipated shocks to permanent income, preferences, and health.

3. Theory

Consider a simple model where households trade-off consumption of three goods – food quality (Q), food quantity (calories, C), and non-food (Q_{nf}). If the consumer's food problem is separable, we can consider the food quantity and quality trade-off as a simplified expression of a more complicated food choice problem with many goods and relative prices but only these two characteristics – the relative price of food quality and quantity here is then a shadow price that depends implicitly on all relative food prices. A simple Stone-Geary version of the food sub-problem would be:

$$\max_{Q,C} \left(C - \bar{C}\right)^{\alpha} (Q)^{1-\alpha} \;\; s.t. \; p_c C + Q \leq X_f \tag{1}$$

where the parameter $\bar{C}$ is the caloric needs of the household that were required to generate the household expenditure, X_f are food expenditures, and p_c is the (shadow) price of calories, with the price of quantity normalized to one.[1]

Solving this problem yields optimal calorie choice $C = \alpha(X_f/p_c)+(1-\alpha)\bar{C}$, which is increasing in food expenditures and caloric needs but decreasing in the price of calories relative to quality. While our hypothetical food quality is not directly observed in the data, the model gives a proxy in the form of calories per unit of food expenditure $(C+\bar{C})/X_f = (\alpha/p_c) + (1-\alpha)(\bar{C}/X_f)$ which is increasing in food expenditures but *decreasing* in caloric needs. The total food (indirect) utility from this problem is given by

$$U_f = \frac{\alpha^\alpha(1-\alpha)^{1-\alpha}}{p_c^\alpha}\left[X_f - \frac{1-\alpha}{\alpha}\bar{C}p_c\right] \tag{2}$$

which is increasing in food expenditures and decreasing in caloric needs.

We model demand between food and non-food with a CES function of their sub-utilities:

$$U = \left(U_f^{\frac{\sigma-1}{\sigma}} + U_{nf}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{3}$$

with budget constraint $X_f + Q_{nf}(U_{nf})p_{nf} \leq Y$. The price of non-food relative to food is p_{nf} and total expenditure is Y. Let $U_{nf} = \gamma_{nf}Q_{nf}$ where γ_{nf} represents some combination of exogenous shifters that affect preference for non-food, including preferences, quality, and variety. Substituting this and equation 2 we have the consumer problem:

$$\max_{X_f, Q_{nf}} \left(\left[\frac{\alpha^\alpha(1-\alpha)^{1-\alpha}}{p_c^\alpha}\left[X_f - \frac{1-\alpha}{\alpha}\bar{C}p_c\right]\right]^{\frac{\sigma-1}{\sigma}} + [\gamma_{nf}Q_{nf}]^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{4}$$

subject to $X_f + Q_{nf}p_{nf} \leq Y$.

Denoting $\gamma_f = \frac{\alpha^\alpha(1-\alpha)^{1-\alpha}}{p_c^\alpha}$ and $\bar{C}^* = p_c(\frac{1-\alpha}{\alpha})\bar{C}$, the solution for food expenditure is given by $X_f = (Y + [(p_{nf}\gamma_f)/(\gamma_{nf})]^{1-\sigma}\bar{C}^*)/(1 + p_{nf}\gamma_f/\gamma_{nf})^{1-\sigma}$ with non-food demand given by $Q_{nf} = \frac{Y-X_f}{p_{nf}}$. If food and non-food are substitutes ($\sigma > 1$) then food

[1] Caloric needs can be endogenized and made dependent on a choice of labor supply and ultimately household expenditures. We do not pursue this here but it has interesting implications for some welfare metrics. Logan (2009) suggests that the slope of calorie Engel curves is an intuitive measure of hunger or welfare and potentially superior to using budget share or total calorie consumption, as it takes into account the marginal propensity to consume on a basic necessity that should be falling in the standard of living. However, if the generation of expenditures is more calorie intensive in some areas and periods, this can result in steeper calorie Engel curve slopes regardless of the actual level of welfare of the population, though it will still be the case that *conditional on expenditures* a locally steeper calorie Engel curve slope implies lower welfare.

expenditure increase in the price of non-food p_{nf} and decrease in the taste shifter for non-food γ_{nf}. Note that our primary object of interest – the calorie-total expenditure Engel curve – can be derived by substituting equation the solution for X_f into the solution for C above, yielding:

$$C = \alpha/p_c[\frac{Y + [(p_{nf}\gamma_f)/(\gamma_{nf})]^{1-\sigma}\bar{C}^*}{(1 + p_{nf}\gamma_f/\gamma_{nf})^{1-\sigma}}] + (1-\alpha)\bar{C} \tag{5}$$

We can also derive a similar expression for food quality, calories per food expenditure ($[C + \bar{C}]/X_f$).

We emphasize several predictions of the model that guide are analysis:

1. Both food quantity and quality are increase in total expenditures.
2. An increase in caloric needs ($\bar{C}$) will increase caloric intake (C) conditional on total expenditures, shifting up the calorie Engel curve.
3. An increase in caloric needs ($\bar{C}$) will decrease calories per food expenditure ($[C + \bar{C}]/X_f$) or food quality conditional on total expenditures.
4. A fall in the relative price of non-food or an increase in the taste for non-food will lead to lower food expenditure (conditional on total expenditures), and this will decrease both food quality and quantity.

Our most important intuition, which goes through in a wide class of models, is that food quality measured by calories per food expenditure ($[C + \bar{C}]/X_f$) is helpful for distinguishing shocks to caloric intake driven by caloric needs or by relative food/non-food prices and tastes. Forces that pull resources out of food lead to reductions on both the quality and quantity margin, while a fall in caloric needs frees up resources that are partly allocated to food quality and partly allocated to non-food items on the margin. This suggests that in addition to analyzing the effect of occupation, household facilities, household size and age on caloric needs and caloric intake, we should also look at the impact of these variables on food quality (calories per rupee of food expenditure) to ensure that the data are consistent a primary role for caloric needs and are not driven by some other mechanism.

For example, when we regress caloric intake on access to electricity, a finding that calories decline with access to electricity may simply be due to an increase in household demand for non-food items – like televisions, fans, and computers – that divert budgets from food to non-food. However, if this is the case then we would also expect to see a decline in food quality, as optimizing households should be reducing on both quality and quantity margins. Similarly, if larger households simply have lower effective prices for non-food goods and the substitution effect is large, we would expect a fall in both caloric intake but also in food quality; conversely if the main factor behind the lower caloric intake is lower caloric needs for large households, we might instead see a rise in food quality for larger households. The intuition that food composition is a good metric for satisfaction of nutritional requirements is consistent with recent work by Jensen and Miller (2010) who argue that the share of calories consumed that come from staples reflects the degree of hunger and undernutrition in the population.

4. Data

4.1. Caloric Intake

To measure caloric intake we use the National Sample Survey (NSS) rounds for 1983, 1987-1988, 1993-1994, 1999-2000, 2004-2005. These "thick" rounds contain household level data for over 100,000 households. Consumption is measured item-by-item at the household level using a 30-day recall period, with both expenditures and quantities recorded (expenditures are imputed for home-produced foods). [2] A major advantage of these data is the sheer size of the sample and the large number of household variables (mostly defined in a common way over time and across other Indian household surveys) that allows for many different cuts of the data.

In order to construct caloric intake from this data, the standard approach (adopted by Deaton and Dreze (2009) among others) is to multiply the quantities by caloric conversion factors from Gopalan et al. (2004) and several other sources. While this works well for many goods, some assumptions and imputations are required. The three main issues are (1) treatment of food with missing or imprecise quantity data (whose caloric conversions per quantity may be certain), (2) composite or processed

[2]The notable exception is the 55th (1999-2000) survey round which used a 30-day and a 7-day recall period. Critics observed that using a shorter additional recall period biases upward consumption measures over the 30-day period, leading to overestimation of the decline in poverty. See Deaton and Dreze (2002) or Deaton and Kozel (2005) for discussion. When pooling multiple rounds our results are not sensitive to excluding the 55th round entirely.

food items with unknown caloric conversion factors (even though quantity may be precise), and (3) meals to/from others that bias the numerator or denominator of a household calories per capita.[3] In the data appendix we discuss these issues in detail and explore the sensitivity of caloric intake to several different assumptions and imputation methods. Our numbers are close to those from Deaton and Dreze (2009) but differ substantially from those of several other authors (unfortunately many studies using these data do not report how their measures are constructed). Our preferred calorie estimates use direct calorie conversion whenever possible and impute calories for items missing quantities or caloric conversion factors using the (expenditure) weighted average of calories per rupee from other foods in the same "group" (e.g. "other vegetables" generate the same calories per rupee as the expenditure-weighted average of all the vegetables for which we observe quantities and caloric conversion factors).

4.2. Caloric needs

Our caloric needs measure is a novel use of time-use diary data; by multiplying detailed activities by an activity intensity factor and basic metabolic requirements for each age and gender, we derive individual and household level caloric requirements. Our data come from the one-time India Time Use Survey (TUS), which was implemented by the same National Sample Survey Organization that carries out the NSS. The data was only collected for one year (between July 1998-June 1999) and six states (Haryana, Madhya Pradesh, Gujarat, Tamil Nadu, Orissa, and Meghalaya) but contains 18,620 households. Every household member over age five was asked their time-use over the previous 24 hours as well as for abnormal and variant days (e.g. weekends, trips into town). Time-use is recorded in 20 minute increments and is classified into 154 different types of activities. A major advantage of this survey is that it records many household variables in the same format as the NSS consumption surveys, including monthly household expenditures, age, gender, education, and occupation, and allows for analysis of home production activities that are typically missing from employment surveys.

To convert activities in the TUS into caloric requirements we use age/gender minimum caloric requirements (corresponding to the Basal Metabolic Rate or BMR) mul-

[3]Calories from alcohol only play a minor role but are included in our main measure of caloric intake.

tiplied by a scaling factor. We used our own intuitive classification based on four activity levels, as well as a more detailed classification based on matching activities using a detailed online calorie per hour calculator. The data appendix contains the details of our imputation procedure. Children under 6 do not have time-use recorded in our data so we use caloric requirements from the India Council for Medical Research (ICMR). For most of our analysis we aggregate to the household level as we do not have individual caloric intake.

Other than the general difficulty and measurement error involved in matching TUS activities to a particular intensity level, there are two major limitations of the caloric needs measure we use. First, we do not observe the heights and weights of individuals – these would affect baseline caloric requirements in a systematic way. Second, we are unable to capture differences in energy intensity *within* specific activities, which is likely to be particularly important for activities where there is lots of substitution of animal and machine power for human energy. This issue arises primarily for transport (the mode of transport is not recorded in the survey) but also for some agricultural activities. We expect this omission to bias up the energy requirements of richer households that use more capital and animal power for their activities. Despite these limitations, we believe our measure of caloric requirements is superior that of the India Council of Medical Research, which only feature three classes of energy intensity for adults - heavy, moderate, and sedentary. The ICMR guidelines do not allow for different activity levels for children and seniors or for the metabolic effects of aging. Our measure generates continuous differences in caloric requirement along both an extensive (length of the work day) and intensive (type of activity) margin and for both market and home production. Most critically, the TUS gives us many household variables – including household size, age, total expenditures, occupation, and education – that are defined identically in the consumption data, allowing us to compare the effects of a given variable on both caloric intake (from the NSS) and caloric needs (the TUS) simultaneously.

4.3. Anthropometrics

Another source of insight into the interaction between caloric needs and caloric intake is provided by anthropometric data. The India National Family Health Survey (NFHS) was collected for the years 1992-1993, 1998-1999, and 2005-2006 by the Indian Ministry of Health and Family Welfare. The nationally representative survey focuses on child and maternal outcomes – only child heights and weights are recorded for all three survey rounds – but in recent years also includes measurement of the heights and weights of women between the ages of 15-49 (beginning in 1998) and men between the ages of 15-54 (beginning in 2005). In addition to measuring the heights of weights of eligible members from the sampled households, the data contain a complete household roster (including information on the age, sex, and education of every household member), and household level variables including principal occupation and household asset ownership. The data do not measure household expenditures, which makes it difficult to compare households on this basis to those NSS and TUS data, but the asset data is much richer than other surveys and can be used to generate a wealth index that captures much of the same variation. We consider both the five point wealth index provided by the survey (based on a principal component analysis of numerous asset and housing related variables) and generate our own measure that is comparable across years, based on the sum of indicator variables for whether the household has: piped water, a flush toilet, a television, a refrigerator, a bicycle, a motorcycle, a car, an electric fan, and a sewing machine.

In addition to the height and weight variables, we focus on the Body Mass Index generated by the data. This index is constructed using the formula BMI=(weight in kg)/(height in $m)^2$. Note that the body mass index essentially gives us a measure of weight and net nutritional intake normalized by height – as heights change across survey rounds and cohorts, comparing adult BMI provides a better indicator of past differences in caloric intake and needs than comparing weight alone. While height arguably provides the best indicator of childhood nutritional outcomes, BMI is likely to be a better measure of adult nutritional outcomes.

5. Results

5.1. Framework

For the results that follow, we regress an outcome variable Y on a variable of interest X and set of controls. The outcome variables we examine are the following:

- Household caloric intake per capita
- Household caloric needs per capita

- Household food quality proxied by calories per rupee of real expenditure
- Household food quality proxied by the share of staple calories in total calories, where staples include all grains, cereals and cereal substitutes
- Individual-level anthropometric outcomes, specifically adult BMI for women, or for women and men.

The variables of interest X that we believe affect caloric needs and through them caloric intake are the following:

- Occupation: NCO(1968) 1-digit classification, consisting of professional, administrative, clerical, sales, service (non-sales), primary (agriculture), secondary (manufacturing) and other
- Education: mean years of schooling for all household members over 18
- Household facilities: household use of electricity, use of common biofuels (firewood, dung) instead of gas/propane for heating/cooking, and access to piped water (anthropometric survey only).
- Household size
- Mean age of household members over 18 (in five-year categories)

We include a full set of control variables except where otherwise noted (or collinear with the X variable), including cubics in log real expenditure per capita (or cubics for household wealth for the anthropometric data), cubics in household size, and demographic controls in the form of the ratio of males and females in 3 year age groups up to age 16-18, the ratio of males and females above age 55. For the individual-level anthropometric outcomes we include age and sex controls to account for typical demographically-driven BMI patterns. We also use geographic dummies at the most disaggregate-level provided by the surveys, which is the village for rural areas and city blocks for urban areas. This ensures that our results are not driven by geographic variation in prices, tastes, and product variety. We focus primarily on the 1999-2000 NSS and 1998-1999 TUS for caloric intake and needs (and restrict to the six common TUS states) but use the other survey rounds where appropriate (particularly for separating cohort and age effects), and we focus on the 2005-2006 NFHS data because it includes adult men and unmarried women.

The Y variable is entered in logs, while the X variable is entered non-parametrically with a series of dummy variables for each outcome. This means that the coefficients we report should be interpreted as percentage deviations from the baseline (or omitted) level of X. As our results typically control for household expenditure and wealth, these coefficients should be interpreted as *shifts* in the caloric intake and caloric need Engel curves, and not movement along the curves driven by expenditure differences. Thus when identifying the effects of occupation on caloric intake and needs, we are identifying the effects of occupation that are orthogonal to total expenditures.

5.2. Calories and economic development

Figure 4 combines the caloric intake and caloric needs data for rural and urban sectors of India in the 1998-2000 period. We see that the downward shift in the calorie Engel curve for urban areas is mirrored by a downward shift of similar magnitude for the caloric need Engel curve. While calorie Engel curves slope upward, implying richer households consume more calories, the caloric need Engel curve is flat or slightly downward-sloping, implying that richer households should be gaining weight over time while poorer households are losing weight.[4] Taken together, these curves imply that the level of real expenditures at which the average household is in caloric balance is lower in urban areas. Since urban areas also have higher real expenditures on average, we would also expect adults in urban areas to be substantially heavier and the difference would be increasing with age.

While the existence of a rural/urban differential in caloric needs may explain the difference in caloric intake, it is not clear which particular factors are playing a role. To unpack the "urban effect" identified in figure 4 we first turn to the household occupational classification, which isolates at a crude level the differences in caloric intake and needs across broad occupation classes like agriculture versus clerical work. Figure 5 presents the coefficients on different occupation dummies from a regression of caloric intake or caloric needs on the occupation dummies and the full set of controls. The omitted category is professionals, so the results indicate that households with the other occupation categories tend to have both higher caloric intake and higher caloric needs conditional on the level of expenditures and demographics. Primary (agricultural) sector households have almost 20% higher caloric needs

[4]As discussed earlier, we might expect the true curve to slope downward even more.

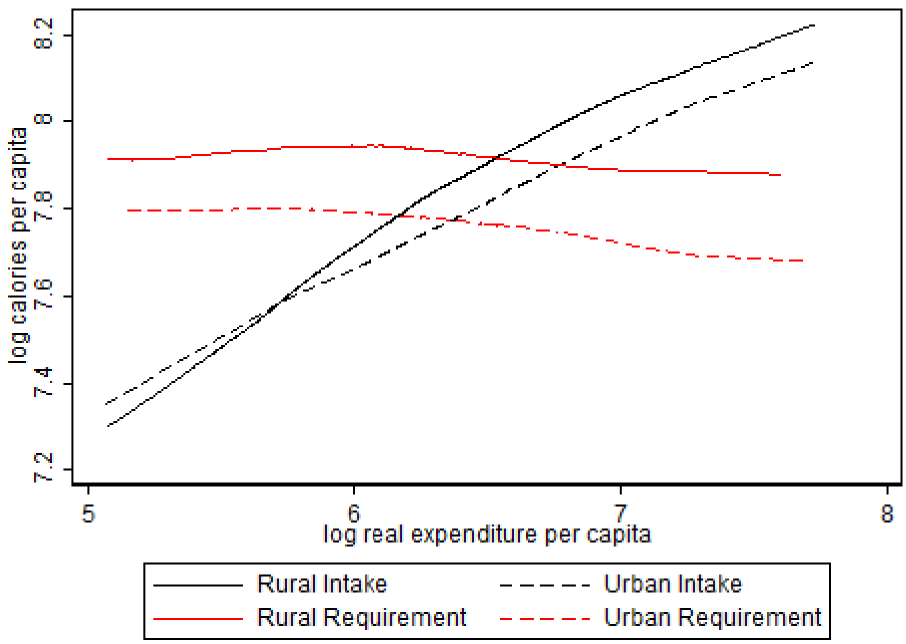

Figure 4: Calorie intake and requirement Engel curves for 2-person households (no children, 1998-2000)

to reach a given level of expenditure than professionals, and this coincides with over 10% higher caloric intake. We stress that these coefficients say nothing about nutritional adequacy per se because the omitted occupation may be in net balance or deficit and because nutritional adequacy depends strongly on expenditures (which vary systematically with occupations) given the upward-sloping calorie Engel curves and generally flatter caloric need Engel curves – these coefficients are only capturing the vertical shifts capture by figure 4.

The magnitudes of the coefficients on caloric intake and needs are generally in line, with primary and secondary sector workers having the highest caloric need and intake coefficients. To see this better figure 6 presents a scatter plot of the occupation coefficients for caloric intake against the occupation coefficients for (A)caloric need, (B)calories per rupee, and (C)the staple share of calories. In all three cases we find a strong upward slope. The fact that the slope in panel A is greater than one is consistent with the impression from 5 that the shifts in caloric need are not passed through entirely into the shifts in caloric intake, which is consistent with the reduction in food quality observed in panels B and C.

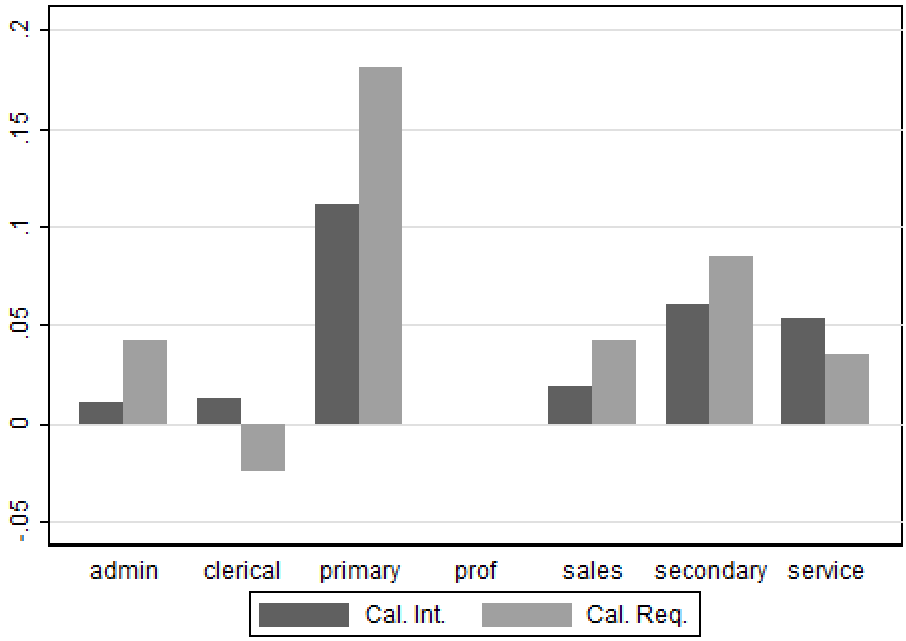

Figure 5: Percent deviation of calories from "professional" occupations (1998-2000)

This last point is also consistent with the differences in adult BMI across household occupations. Figure 7 shows that most occupations have lower BMI than professionals and administrators. This effect is much larger when we do not control for the wealth index, but even conditional on household wealth we find that certain occupations, particularly in the primary and secondary sector, tend to produce lower body masses. This is what we would expect if, conditional on expenditures, these occupations had higher caloric needs, which according to our model will translate into (1)higher caloric intake, (2)lower food quality, and (3)lower excess caloric intake above need (and hence lower body mass).

We next turn to average adult education, which varies more continuously than occupations and better captures the "average" occupation of the household when there are multiple working adults. Figure 8 presents the coefficients on dummies for average years of adult education for both caloric intake and caloric needs. The two series match each other very closely, with high-school graduates having caloric intake and caloric needs that are about 10% lower than households with no education *conditional on expenditure.*

Figure 9 shows that food quality tends to rise with education, as calories per rupee

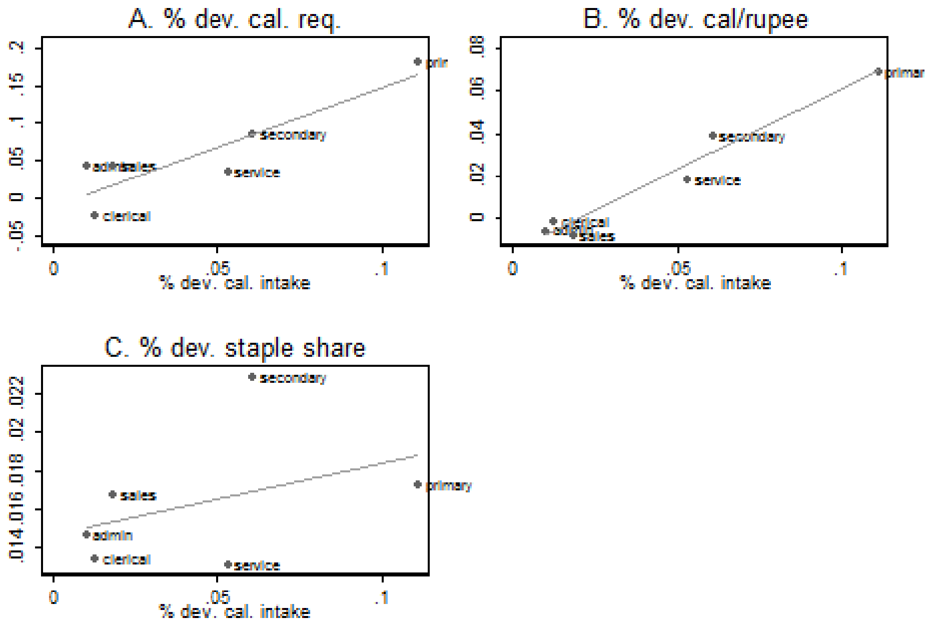

Figure 6: Scatter of deviation in calorie intake against other variables (1998-2000)

and the staple share both drop. This suggests that caloric needs should be falling more rapidly than intake with education, perhaps more evidence that our measure of caloric needs is biased upward for better-off households. At a minimum it confirms that the lower conditional caloric intake for educated households is not a sign of caloric inadequacy.

Turning again to the body mass index for adults, Figure 10 shows that more educated households have significantly higher body mass than less educated households. Part of this result is driven by an income-effect of education, operating through movement up along the calorie Engel curve, but even after conditioning on household wealth we see that more educated households have up to 4% higher body mass (equivalent to 4% higher weight conditional on height).

Finally, we turn to our original question of what particular factors drive the rural-urban difference in caloric intake and caloric needs. We explore this issue by considering whether the residual "urban dummy" from a regression of outcome Y on controls can be reduced by including the specific occupation and education variables we consider above as well as certain household facilities. Thus while the previous results used fine geographic controls, we here simply lump together the rural and ur-

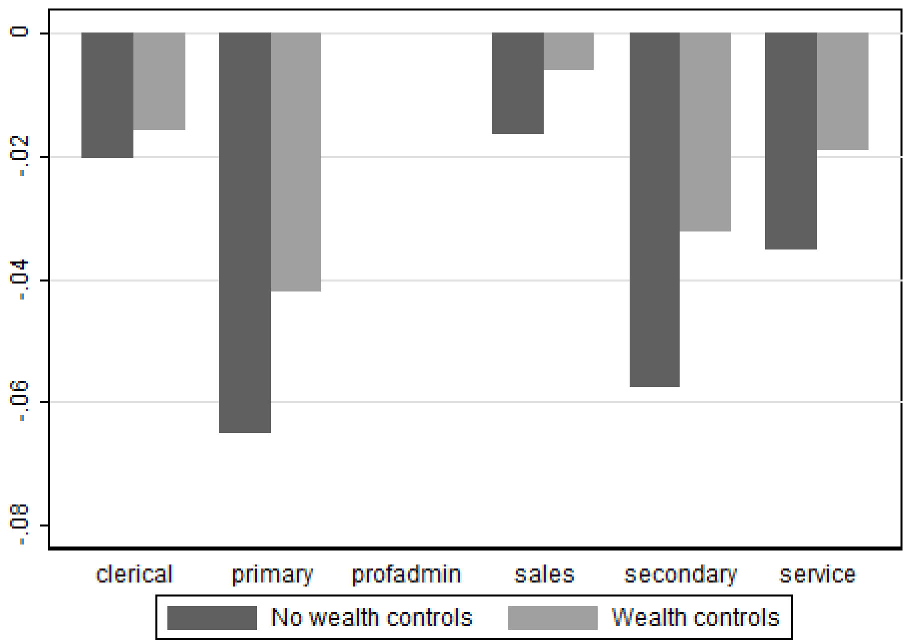

Figure 7: Percent deviation of adult BMI from "other" occupations (2005)

ban areas of six Indian states in the 1998-2000 period and compare them. We examine all of the outcome measures Y which are as always entered in logs.

Table 1 begins by reporting the unconditional urban-rural difference in the first row. We find that urban households have unconditionally higher caloric intake, lower requirements, consume higher quality food on one measure (but lower on the staple share measure) and have significantly higher body mass. In the next row we see that after controlling for demographic variables and more importantly cubics in real household expenditures per capita – that is, after we control for the calorie Engel curve and the fact that urban households have higher real expenditures – we see our initial puzzle, which is that urban households have roughly 13% lower caloric intake and requirements, consume higher quality food, and have modestly higher BMI. These capture the unexplained "shifts" in the Engel curves for the various outcomes.

The subsequent rows all include these same controls and add our orthogonal shifters for occupation or education. We find that occupation generally does a better job than education of capturing these shifts, and that controlling for occupation (and thereby comparing only households with the professional occupation in rural and urban areas) reduces the unexplained urban dummy by about 50%. We also consider

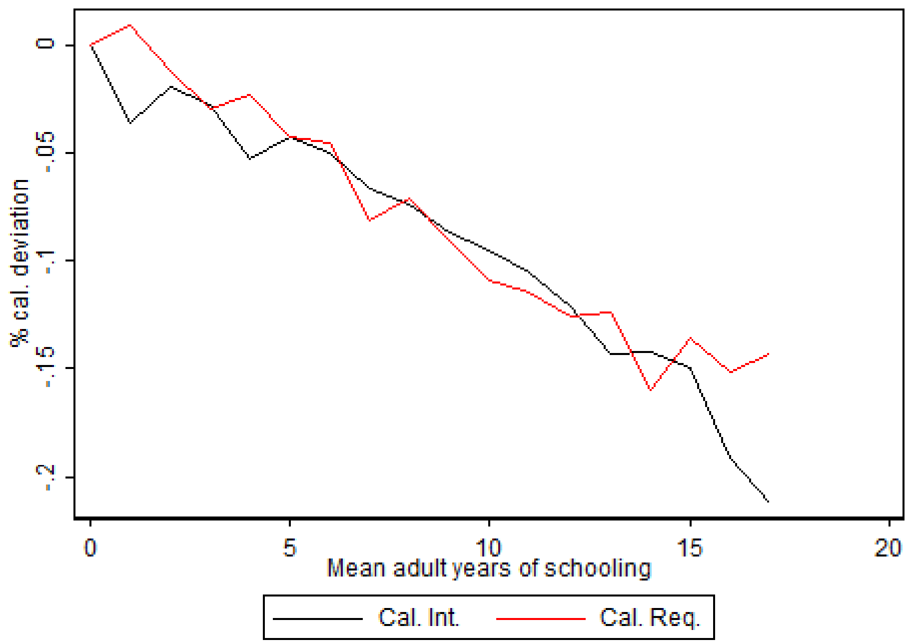

Figure 8: Percent deviation of calories from zero years schooling (1998-2000)

the effect of just three variables that capture household facilities that are likely to impact caloric needs – use of electricity, which varies widely across rural and urban areas and allows households to access a wide range of labor saving appliances, and use of firewood and animal dung, the two main biofuels that are widely-used in rural areas in lieu of liquid propane gas (LPG) and that can generate substantial caloric needs when collected. These three dummies for energy use have a very large impact on our outcome measures. Combining all of our controls variables, the unexplained "urban" dummy is reduced to only 1% for intake, 3$ for requirements, and low levels for food quality and BMI.[5] We thus conclude that the caloric needs explanation for differences in urban and rural caloric intake is highly plausible and quantitatively consistent with the data.

We next turn to the larger question of whether we can explain the decline in caloric intake over time in India that occurred despite substantial growth in real expenditures. If the factors that drive this are similar to the ones that drove the urban-rural

[5]Somewhat surprisingly we find that urban food quality is actually lower after controlling for all of these variables, though this may reflect a small relative price effect including better access to subsidized staples through the public distribution system (PDS).

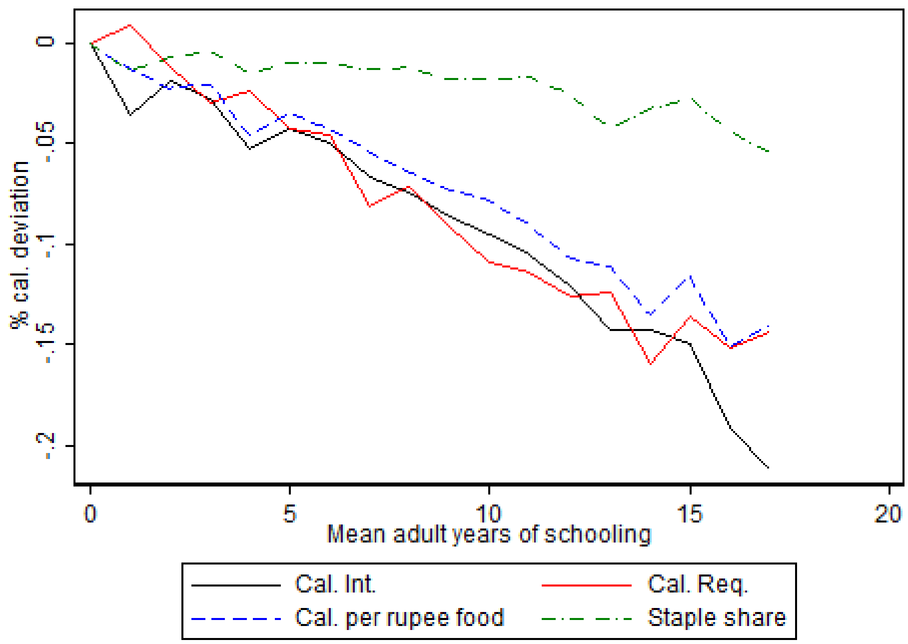

Figure 9: Percent deviation from zero years schooling, other indicators (1998-2000)

difference, we would expect the same variables to explain most of the 2005 dummy effect when regressing pooled 1983 and 2005 data for caloric intake on a set of controls. Of course, some factors that operate over time – such as shifts in the labor intensity within certain occupations, the growth of electricity and LPG access, and perhaps most importantly shifts in relative prices and the availability of new products – may not operate at all across sectors in the 1998-2000 period. Another limitation of our exercise is that we only have a single cross-section for time-use, so we cannot examine changes in caloric needs over time using the methodology we used for the urban versus rural comparison.

The results of this exercise, reported in table 2, indicate that the same variables that explain virtually all of the urban-rural differences in caloric intake, caloric needs, food quality and BMI in 1998-2000 do a significantly poorer job of explaining the decline in calories for both rural and urban areas over the 1983-2005 period. In particular, the unexplained drop is 14.4% in rural areas and 10% in urban areas, and including the full set of controls reduces this unexplained drop to only 11.2% and 8.3% – occupation, education, and access to electricity and propane gas only explains about 20% of the drop in caloric intake over this period. However, these same variables do explain

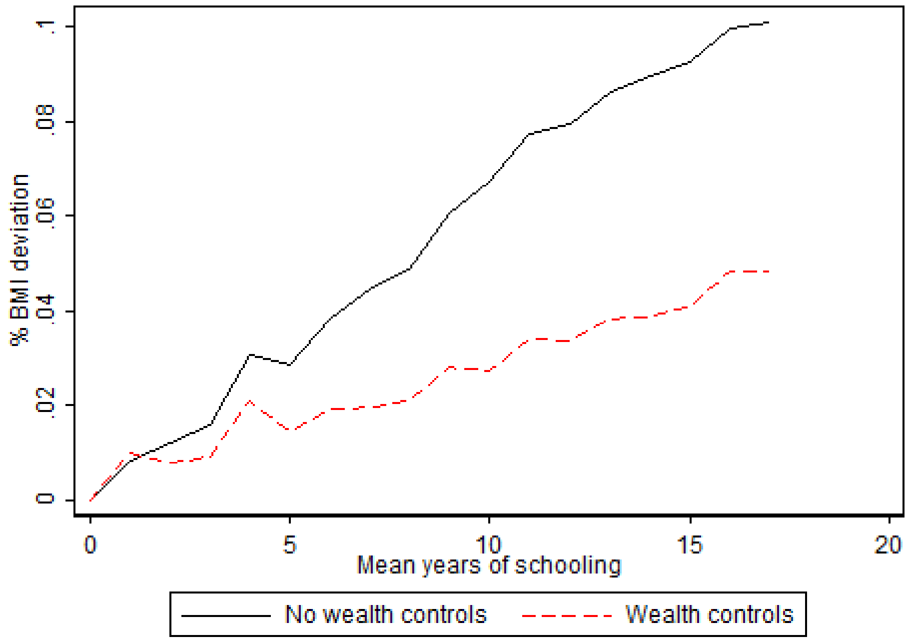

Figure 10: Percent deviation of adult BMI from zero years (2005)

a larger share of the increase in food quality that we observe. Taken together, these results suggest that either our controls for caloric need are inadequate for reasons discussed earlier, that other factors are at play, or most likely both.

As our goal is to assess the role of caloric needs in generating several different food consumption puzzles we do not wish to delve too deeply into the alternative explanations for this particular puzzle, but we have two pieces of evidence that suggest that caloric needs are not enough to explain the downward shift in calorie Engel curves over time in India. The first piece of evidence comes from Figure 11 which reveals that while food and non-food prices have increased at roughly the same rate over the 1983-2005 period (with some fluctuation in certain sub-periods) there is a substantial divergence between certain non-food prices – energy prices, which are heavily regulated by the government, have increased much more slowly than prices for clothing and other miscellaneous goods. This would be consistent with households using energy more intensively over time (particularly in combination with increases in access to electrical and gas networks and LPG delivery services), which would reduce caloric needs but also lead to other changes in consumption patterns such as increased purchase of electricity-using appliances. There is plenty of evi-

Table 1: Can we make urban dummy zero?

Y variable	Cal. intake	Cal. Req.	Cal/rupee	Staple share	BMI male	BMI female
No controls	0.019	-0.108	-0.192	-0.030	0.065	0.086
	(0.004)	(0.005)	(0.004)	(0.001)	(0.001)	(0.001)
Dem./exp./wealth	-0.126	-0.129	-0.049	-0.027	0.007	0.020
	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
Including dem./exp./wealth controls						
Occupation	-0.077	-0.042	-0.013	-0.025	0.003	0.016
	(0.004)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
Education	-0.102	-0.094	-0.032	-0.024	0.007	0.019
	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
Energy	-0.038		0.017	-0.009	0.005	0.017
	(0.004)		(0.003)	(0.001)	(0.001)	(0.001)
All	-0.011	-0.029	0.036	-0.011	0.001	0.014
	(0.004)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
All vs. Dem./exp./wealth						
Share explained	91.32%	77.14%	174.14%	61.04%		

Standard errors in parentheses.
Intake data 1999-2000.
Requirement data for 1998-1999.
BMI data for 2005

dence that ownership of these appliances has increased at a faster rate than would be predicted based on cross-sectional expenditure elasticities alone.

A further piece of evidence comes from analyzing the composition of household budgets directly – if food budget shares have fallen holding real expenditure constant, which categories have been rising? Figure 12 reveals that the downward shift in food has not been accompanied by a uniform upward shift for all other categories (consistent with a pure food/non-food relative price effect) or by upward shifts that are proportional to the Engel curve slope (which would be consistent with mismeasurement of the price-level denominator in the X-axis real expenditures). Rather, certain categories – particularly use of fuel and light ("light") and education ("edu") – have increased substantially, others have increased slightly (transportation, services, durables, and non-durables) and others are mixed (clothing, medical care). The 10 percentage point decline in budget share for food can be almost entirely explained by the upward shift in energy use and education expenditure, which suggests that relative price effects, complementarity with new goods, and investment related expenditures can explain a substantial part of the downward shift in calorie Engel curves over time.

Table 2: Can we make 2005 dummy disappear?

Y var.	Cal. intake	Cal. intake	Staple share	Staple share	BMI female	Child Height
Sector	Rural	Urban	Rural	Urban	Both	Both
Comp. year	1983	1983	1983	1983	1998	1992
No controls	-0.053	-0.042	-0.069	-0.034	0.015	0.018
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dem./exp./wealth	-0.144	-0.100	-0.039	-0.009	-0.023	0.012
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Including dem./exp./wealth controls						
Occupation	-0.132	-0.101	-0.037	-0.010	-0.020	0.012
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Education	-0.133	-0.096	-0.031	-0.006	-0.017	0.010
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Durables	-0.125	-0.083	-0.025	0.001	-0.022	0.012
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
All	-0.112	-0.083	-0.021	0.002	-0.013	0.010
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
All vs. Dem./exp./wealth						
Percent explained	22.19%	16.99%	46.42%	123.69%		

Standard errors in parentheses.

5.3. Calories and household size

We now turn to the evidence on caloric intake and household size. Rather than examining food expenditure per capita as in Deaton and Paxson (1998) we focus on caloric intake per capita as this deals with the issue of direct economies of scale in purchasing (bulk-discounting) directly and in our view is a stronger and more puzzling result. However the results using food expenditure per capita are similar to those for caloric intake per capita. For the analysis of household size, we always hold constant the expenditure per capita of the household. While we could use all household types and include various flexible controls for demographics, we focus on results for households with no children to minimize the influence of the number and parametric specification of household demographics on the results. This also helps us in terms of avoiding certain measurement issues (breast-feeding, the absence of activity levels for children under five in the time-use data) and avoids some (but not all) of the alternative explanations for the puzzle due to intra-household allocation. Our demographic controls thus consist of the ratio of females in the household and controls for the mean adult age.

Figure 13 presents the coefficients on different household sizes relative to the omitted category (two adult households). We see that for both per capita caloric intake and per capita caloric needs, there is a substantial decline for larger households

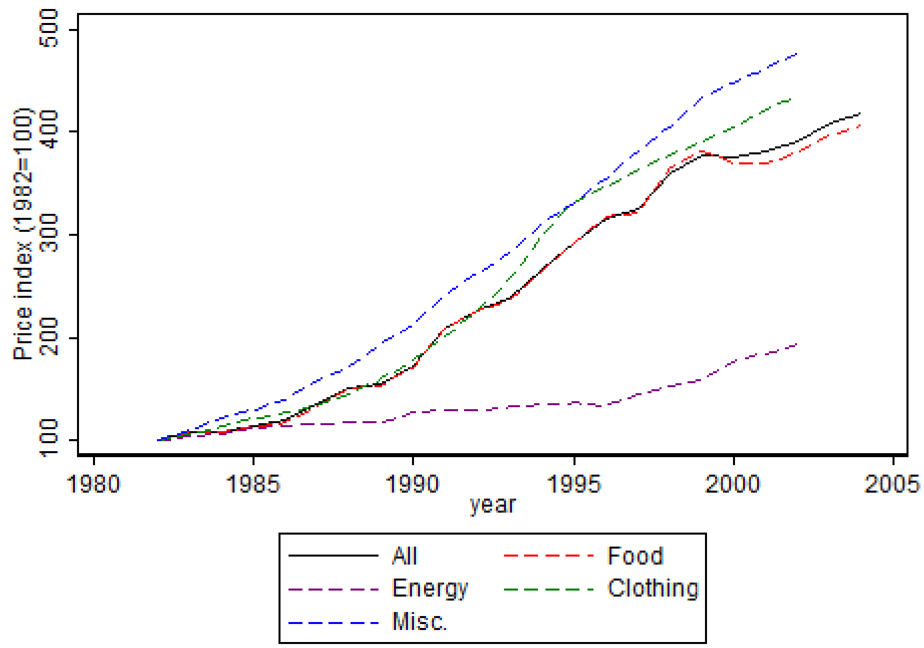

Figure 11: Food and non-food prices

holding constant expenditure per capita. The patterns and magnitudes are somewhat different, however, as caloric intake declines monotonically by 10% going from 2 to 6 adults while caloric requirements increase slightly from 2 to 3 but falls by about 4% for 6 adult households. For larger household sizes, caloric intake continues to decline but caloric requirements appear to level off around 4% lower than two adult households. At first glance it appears that differences in caloric needs can only account for up to 40% of the decline in caloric intake for larger households.

The reason larger households seem to have lower caloric needs is shown in figure 14. Larger households require lower market hours per capita to achieve the same expenditure per capita, and also exhibit lower home-production hours per capita – the different is made up entirely by greater leisure per capita. The result for market hours seems to be primarily driven by agricultural activities and transportation, while the results for home-production hours hold for a wide variety of activities (such as shopping, cooking, gathering firewood and water). There are a range of explanations for these findings, including greater scope for specialization in larger households that yields higher income per hour worked, direct caloric-need economies for some activities (e.g. transport for shopping or gathering firewood and water, cooking) but also

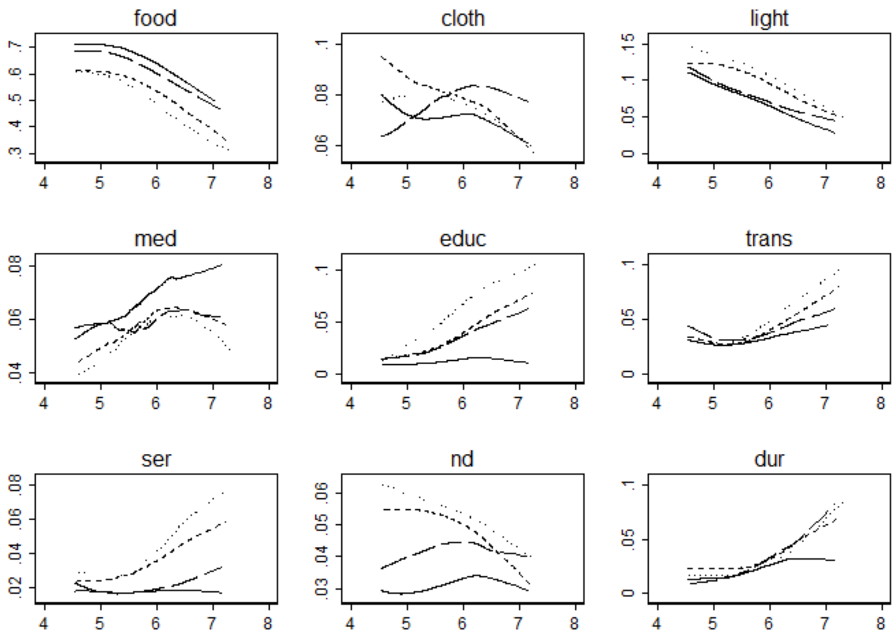

Figure 12: Shifts in budget share Engel curves (Rural 1983(solid), Urban 1983 (long dash), Rural 2005 (short dash), Urban 2005 (dot)

complementarity in leisure across household members. Regardless of the economic mechanism behind this finding, our result indicates that caloric needs provide an important part of the explanation for the Deaton and Paxson (1998) food–household-size puzzle, and help explain why the negative relationship is strongest for food in poor countries – these are the places where caloric needs are likely binding and most strongly related to caloric intake and food expenditures, and where specialization and scale economies are likely have the greatest impact on caloric needs.

Further evidence for this proposition can be found by turning to our other caloric adequacy measures in Figure 15. Here we find that both measures of food quality increase with household size (where food quality is inversely related to calories/rupee and the staple-share of calories). This is inconsistent with greater hunger and undernourishment in larger households, and also suggests that we may be overstating the decline in caloric intake or understating the decline in caloric needs by using the NSS and TUS data. One explanation for this consistent with the literature is that systematic recall biases cause larger households to understate their food consumption (relative to other household expenditures) by more than smaller households. Our

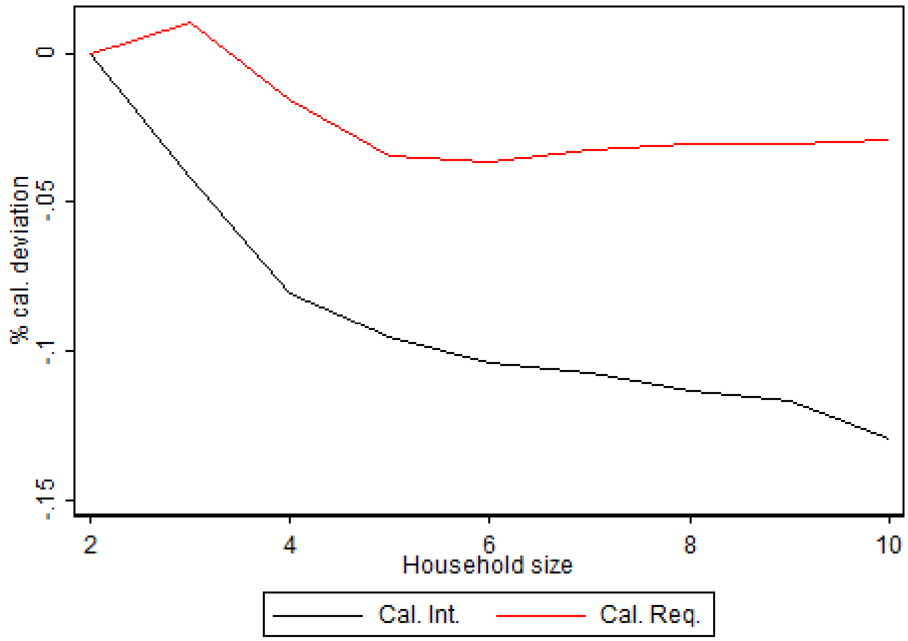

Figure 13: Percent calorie deviation relative to 2-adult household (1998-2000)

caloric needs measure may understate the decline in caloric needs if some activities are performed by multiple household member simultaneously in way that decreases the caloric intensity for each member.

Consistent with these findings, we find some evidence that larger households are not worse of in terms of net caloric balance. Figure 16 shows that adults in larger households tend to have higher BMI than adults in smaller households unconditionally, despite the fact that unconditionally larger households have lower expenditure per capita. The NFHS data do not allow us to control for household expenditure, but do allow for calculation of a wealth index. The problem is that this wealth index is based on household asset ownership is unlikely to be a good predictor of wealth per capita, the preferred corollary of expenditure per capita in the other datasets. When we control for the household wealth index we find some decline in BMI for larger households, but mechanically the larger households have a lower "wealth index per capita." We can control for the "wealth index per capita" by simply dividing the wealth index by household size, which yields a steeper upward sloping relationship between household size and BMI, but this measure has an even less obvious interpretation than the wealth index.

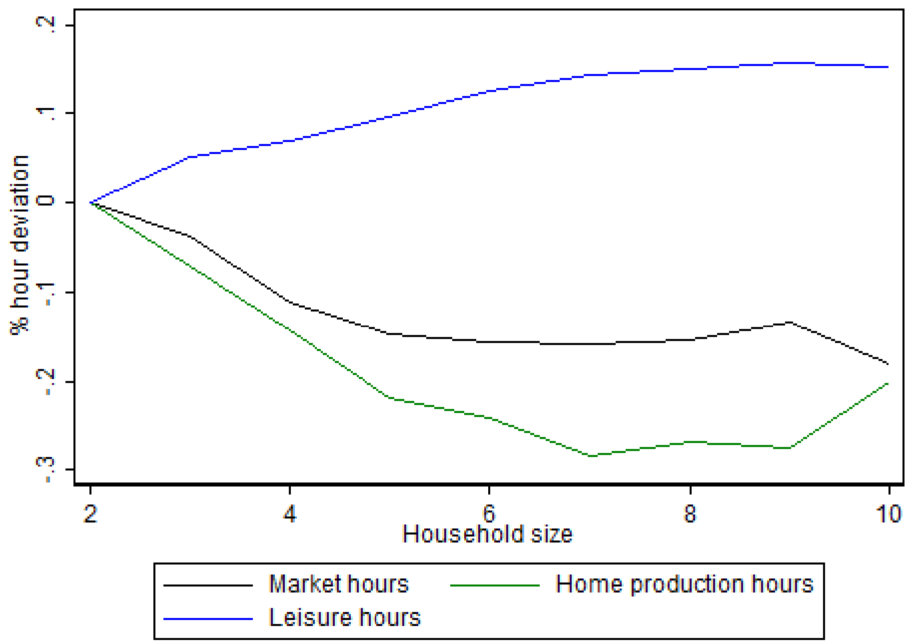

Figure 14: Percent deviation in hours per capita relative 2-adult household (1998-1999)

Altogether we believe these results support a significant role for caloric needs as an explanation for negative relationship between food and caloric intake per capita and household size when conditioning on total expenditures per capita. However, our results leave some room for other explanations that have been advanced in the literature, particularly measurement error in caloric intake that leads to understatement of food intake by larger households.

5.4. Calories over the life-cycle

Our final set of results concern caloric intake over the life-cycle. Here we depart from our previous practice and first focus on results that are not conditional on total expenditures. The reason for this is that the decline in caloric intake in late-life is related to the unconditional decline in total expenditures and food expenditures, so we begin by focusing on the unconditional results. In these regressions the omitted category is households that are 19-22 year old and we group average adult age into five-year categories (23-27,28-32,33-37,38-42,etc. up to 72-77). We also focus on results

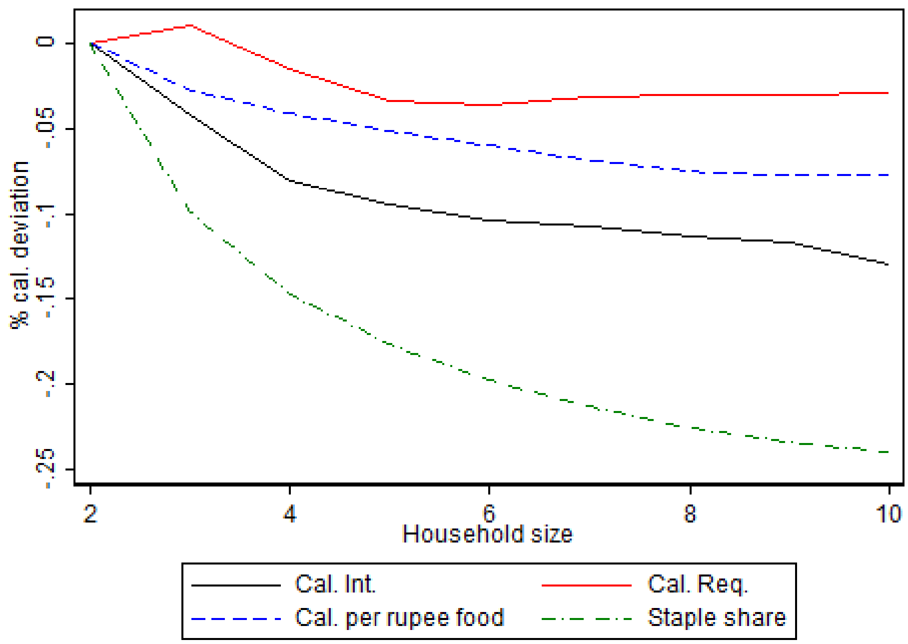

Figure 15: Percent deviation relative to 2-adult household, other indicators (1998-2000)

using the 1998-2000 cross-section for comparability between the NSS and TUS data. While this comes at the cost of not controlling for cohort effects, when we do this for the NSS data by pooling all of our survey rounds from 1983 to 2005 the life-cycle patterns look very similar, suggesting that cohort effects are likely to be small here.[6] These regressions control for a cubic in log household size, the ratio of male and female adults, and the ratio of male or female children in 3 year age groups up to age 18.

Figure 17 presents the basic results from combining the NSS and TUS data to look at caloric intake and requirements over the life-cycle. We observe a substantial decline in caloric intake, as high as 20% when comparing 75 year olds to 20 year olds, but an even larger decline in caloric needs that begins in the early 30s and reaches over 30% by the time households are in their 70s. This is the only one of our three puzzles where the drop in caloric needs appears to be substantially greater than what

[6]Even if cohort effects are large, our comparison of the NSS and TUS data would be informative about the combined effect of cohort and age since the NSS and TUS data are for the same 1998-2000 period.

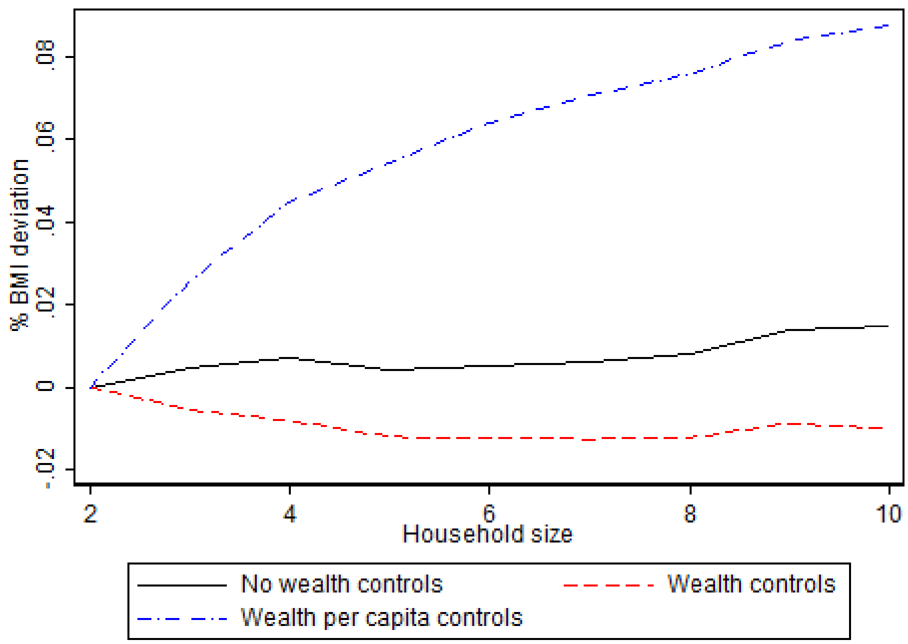

Figure 16: Percent deviation in BMI relative to 2-adult household (2005)

is required to generate the observed drop in caloric intake.

An obvious issue is that our caloric needs measure declines with average adult age by construction, because our caloric needs measure multiplies an activity intensity level by the basal metabolic rate for the age/sex of each household member and this rate declines substantially for older men and women. To see how much this drives our results, as opposed to a decline in physical activity, Figure 18 separates the decline in caloric needs that comes from this age adjustment from the one that is due only to physical activity levels. Thus the series “cal-req. no adjustment” shows the hypothetical caloric needs of older households if they maintained a 19-22 year old metabolism. Roughly half of the decline in caloric requirements is driven by the metabolism effect while the other half is driven by a decline in physical activity levels. The figure provides further insight into the reason for the decline in physical activity levels by decomposing hours into market work, home production and leisure. We see that older households have similar home production levels but substitute market work for leisure.

When we examine our food quality measures in Figure 19, the results are at first puzzling, as we find a decline in food quality for older households even though our

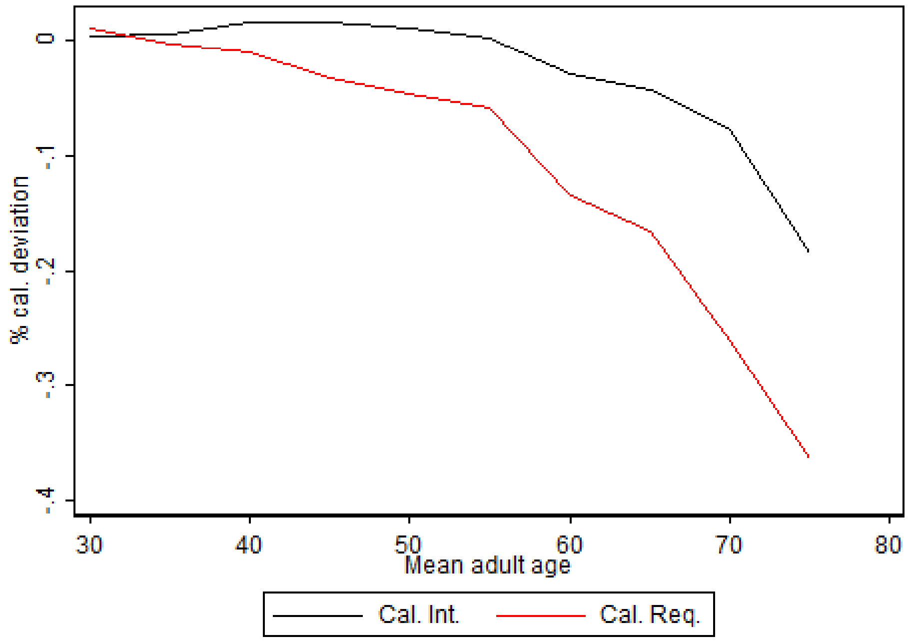

Figure 17: Percent deviation in calories relative to 19-22 year olds (1998-2000)

data say that their caloric needs fall by more than their caloric intake. However, all our results up to this point have been unconditional, and we already observed earlier that both total and food expenditure decline substantially for older households – this decline could push households to reduce food quality even when caloric needs are falling.

When we condition on household expenditures – thereby comparing the atypical older households that have maintained higher expenditures in old age to their younger counterparts – the results change substantially. Figure 20 shows that conditional on expenditure, the decline in caloric intake for older households is about halved in magnitude (from 20% to 10%) and is concentrated in the early-60s to mid-70s, while the decline in caloric requirements is equally large. More to the point, we then find that food quality actually increases in old-age as older households have a lower staple-share of calories and consume less calories per rupee of food expenditure, which is consistent with the findings for caloric intake and needs conditional on expenditures.

If accurate, our unconditional results indicate that Indian households should be increasing in BMI as they age – even though their caloric intake falls relative to 19-22

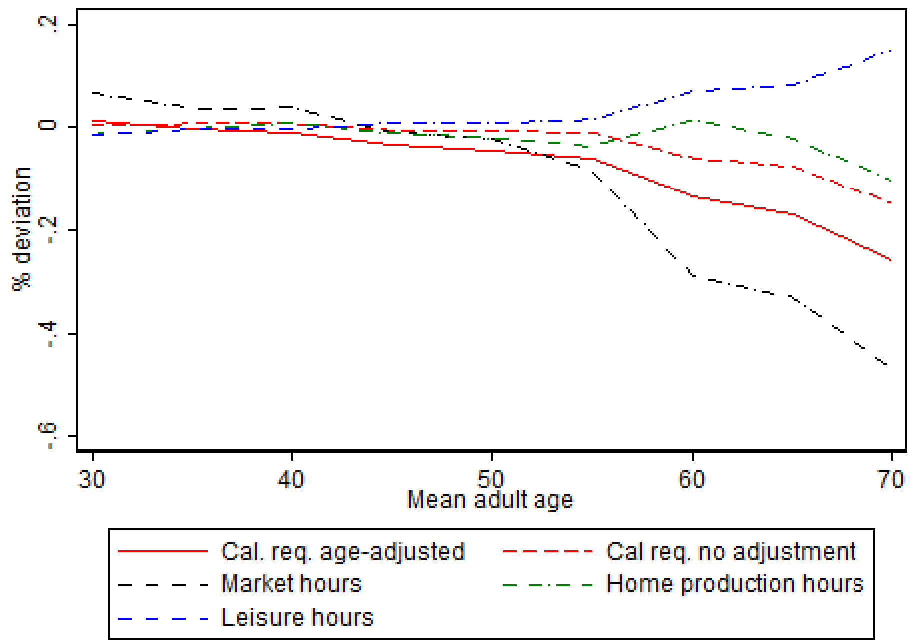

Figure 18: Composition of percent deviation relative to 19-22 year olds (1998-2000)

year olds, caloric needs fall by more. We see this clearly in figure 21, which shows that both men and women have a BMI increasing with age. BMI is also larger for the 2005 cohort than the 1998 cohort, implying some increase improvement in net caloric intake. A striking feature of the data, though one consistent with our earlier results, is that the slope of the BMI-age relationship is much steeper for urban than rural areas. Our results suggest that this occurs because urban households have considerably lower caloric needs – even though their caloric intake and food expenditures are fairly similar to those of rural households on average, they gain weight at a much faster rate as they age because of the difference in caloric needs coming from differences in physical activity levels. A significant limitation of the DFHS data though is that it only covers adults up to age 55 – this makes it difficult to draw conclusions from the anthropometric data about the ages where caloric intake and needs fall the most, the 60s and 70s.

Overall, the conditional results of figure 20 suggest that falling caloric needs over the life-cycle do play a substantial role and may explain some of the unconditional decline in expenditures in late-life, but the substantial decline in overall (including non-food) expenditures and the unconditional decrease in food quality in figure 17

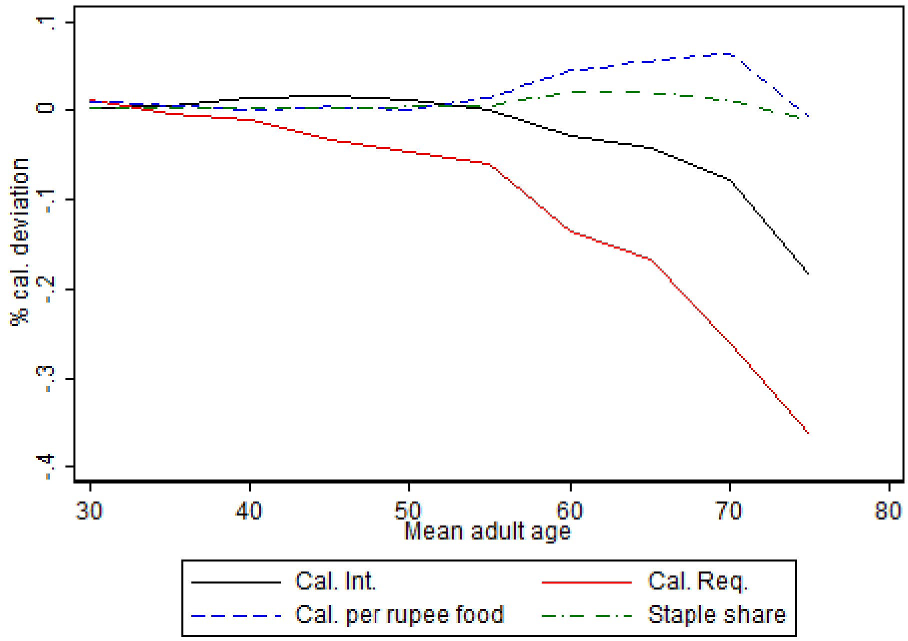

Figure 19: Percent deviation relative to 19-22 year olds, other indicators (1998-2000)

suggest that households may be failing to smooth consumption completely. This topic deserves greater scrutiny as there is little empirical evidence on late-life consumption in developing countries.

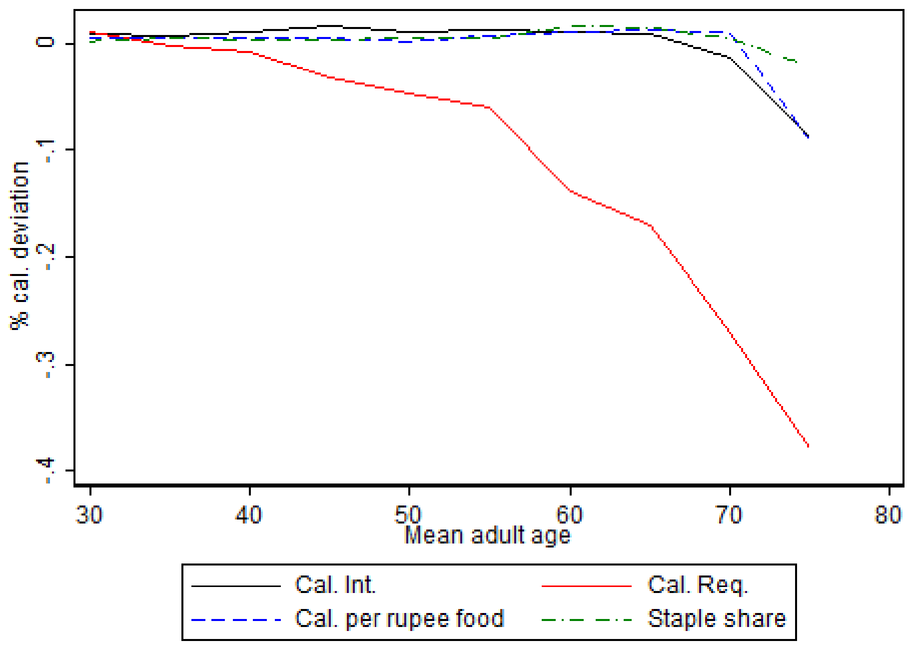

Figure 20: Conditional on expenditure: percent dev. relative to 19-22 year olds, other indicators (1998-2000)

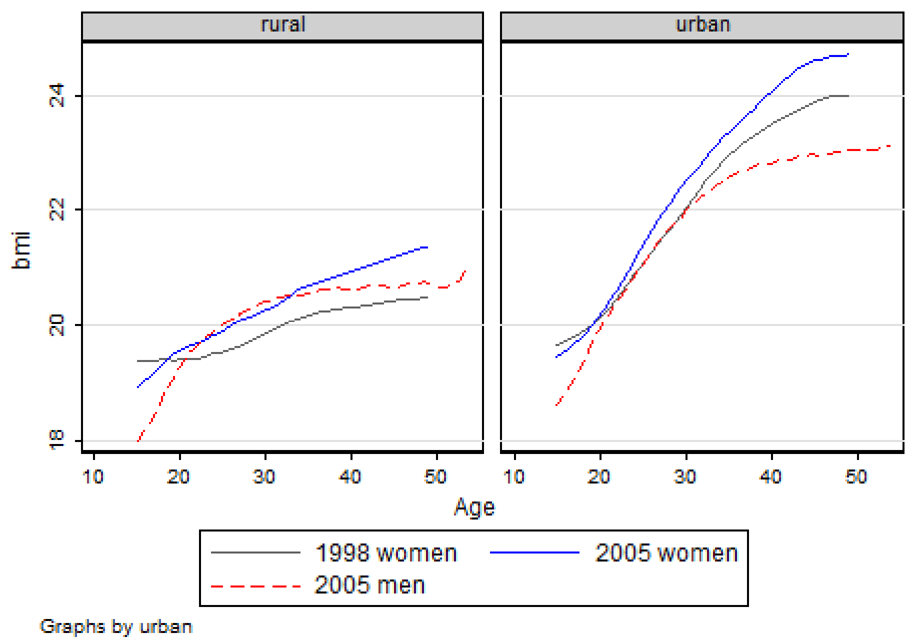

Figure 21: BMI for adults by age and sector (1998,2005)

6. Conclusion

To be written.

References

Abdulai, Awadu, "Economies of Scale and the Demand for Food in Switzerland: Parametric and Non-Parametric Analysis," *Journal of Agricultural Economics*, 2003, *54(2)*, 247–267.

Aguair, Mark and Erik Hurst, "Consumption vs. Expenditure," *Journal of Political Economy*, 2005, *113(5)*, 919–948.

Almas, Ingvild, "International Income Inequality: Measuring PPP bias by estimating Engel curves for food," *American Economic Review*, 2012.

Barten, Anton P., "Family Composition, Prices and Expenditure Patterns," in Gordon Mills Hart, Peter E. and John K. Whitaker, eds., *Econometric Analysis for National Economic Planning*, London: Butterworths, 1964.

Chatterjee, Srikanta, Allan Rae, and Ranjan Ray, "Food Consumption and Calorie Intake in Contemporary India," *Working Paper*, 2007.

Clark, Gregoy, Michale Hubermand, and Peter Lindert, "A British Food Puzzle," *The Economic History Review*, 1995, *48(2)*, 215–237.

Costa, Dora L., "Estimating Real Income in the US from 1888 to 1994: Correcting CPI Bias Using Engel Curves," *Journal of Political Economy*, 2001, *109(6)*, 1288–1310.

Deaton, Angus and Christina Paxson, "Economies of Scale, Household Size, and the Demand for Food," *Journal of Political Economy*, 1998, *106(5)*, 897–930.

_ and Jean Dreze, "Poverty and Inequality in India: A Re-examination," *Economic and Political Weekly*, 2002, *September 7*, 3729–3748.

_ and _ , "Food and Nutrition in India: Facts and Interpretations," *Economic and Political Weekly*, 2009, *44(7)*, 42–65.

_ and Valerie Kozel, "Data and Dogma: The Great Indian Poverty Debate," *World Bank Research Observer*, 2005.

Du, S, B Lu, F Zhai, and B Popkin, "A New Stage of the Nutrition Transition in China," *Public Health Nutrition*, 2002, *5(1A)*, 169–174.

Gardes, Francois and Christophe Starzec, "Economies of scale and food consumption: a reappraisal of the Deaton-Paxson paradox," *Working Paper*, 1999.

Gibson, John and Bonggeun Kim, "Measurement Error in Recall Surveys and the Relationship Between Household Size and Food Demand," *American Journal of Agricultural Economics*, 2007, *89(2)*, 473–489.

Gopalan, C, B V Rama Sastri, and S C Balasubramanian, *The Nutritive Value of Indian Foods*, Indian Council of Medical Research, Hyderabad: National Institute of Nutrition, 2004.

Hamilton, Bruce W., "Using Engel's Law to Estimate CPI bias," *American Economic Review*, 2001, *91(3)*, 619–630.

Hicks, Daniel L., "Consumption Volatility, Marketization, and Expenditure in Emerging Market Economies," *Working Paper*, 2010.

Jensen, Robert and Nolan Miller, "A Revealed Preference Approach to Measuring Hunger and Undernutrition," *Working Paper*, 2010.

Karan, Anup and Ajay Mahal, "Health, nutrition and poverty: Linking nutrition to consumer expenditures," in "Financing and Delivery of Health Care Services in India," National Commission on Macroeconomics and Health, Ministry of Health and Family Welfare, Government of India, 2005.

Kumar, Praduman M and Madan M Dey, "Long-term Changes in Indian Food Basket and Nutrition," *Economic and Political Weekly*, 2007, *September 1st*, 3567–3572.

Lanjouw, Peter and Martin Ravallion, "Poverty and Household Size," *Economic Journal*, 1995, *105*, 1415–1434.

Logan, Trevon, "Economies of Scale in the Household: Puzzles and Patterns from the American Past," *NBER Working Paper 13869*, 2008.

— , "The Transformation of Hunger: The Demand for Calories Past and Present," *The Journal of Economic History*, 2009, *69(2)*, 388–408.

Perali, Federico, "The Second Engel Law and Economies of Scale: an Empirical Puzzle to be Resolved," *Working Paper*, 2001.

A Data Appendix: Imputing caloric intake

A1. Calorie Intake

Deaton and Dreze (2009) have previously documented the decline in caloric intake in India between 1983 and 2005 using this data – see their paper for a discussion of various issues. Table 3 presents their estimates of per capita calorie consumption over this period. We also present independent estimates calculated by other authors with the same data - surprisingly the different studies disagree on both the direction and magnitude of calorie changes.[7] While Deaton and Dreze (2009) find a large decline in rural areas and modest decline in urban areas, Chatterjee et al. (2007) find a decrease in rural areas and an increase in urban areas, while Kumar and Dey (2007) find an increase in both areas. Both Kumar and Dey (2007) and Chatterjee et al. (2007) find that in recent years urban India has higher per capita consumption of calories than rural India. Below we also report calorie intake from the NNMB as calculated by Deaton and Dreze (2009). These data, presented at the bottom of table 3 show a dramatic decline in calories that is over the double the size of the decline for comparable states in the NSS based on the calculations of Deaton and Drèze.

As there is some disagreement about the direction and magnitude of the trend in calorie consumption and as quantitative evaluation is important to us we delve deeper into the construction of calorie intake measures. We cannot address the issue of systematic under-reporting or over-reporting using the NSS data alone, and the 30 day recall period and reliance on a single informant may bias measured food consumption in several ways. Beyond measurement error in the data itself, there are also several important assumptions and imputations that affect the calorie estimates. These can be broadly divided into 3 categories - (1)food items with no quantity data or imprecise quantity units (even though caloric conversion factors may be accurate), (2)composite food items with unknown calorie conversions (even though the quantity measures may be precise), and (3)meals received and given by the household that are not accounted for in total calories or household size (and hence bias estimates of calories per capita). Several items - most notably processed foods, beverages, and cooked meals - suffer from both the first and second problems, and there are some items with inconsistent measurement of quantity and different units across the five survey years we examine. The third problem takes two forms - meals received for free by household members (which are not recorded in the household consumption data but are sometimes recorded on the household roster) and meals given by the household to non-members. A fourth but less important issue is treatment of alcoholic beverages, which are typically not factored into food expenditures or calorie consumption but are potentially an important source of both for some households.

[7] As none of the studies make explicit the details of data-cleaning and calorie imputation we cannot pinpoint the reason for the divergent estimates.

A2. Data issues

To get a sense of the magnitude of these issues, table 6 reports some summary statistics for consumption of the different sets of "problem" goods.[8] The first row reports the share of food expenditures on goods with no quantity data, which has been increasing over time and is higher in urban areas. Many of these goods fall into the processed food and "other" categories. The second row reports the share of expenditures on composite commodities - defined as those commodities with "other" in the description (with three exceptions - "palak/other leafy vegetables" and "other edible oils" are excluded, as their caloric content is likely to be very similar to other products in that category, and cereal substitutes are included since they include varied goods like tapioca, jackfruit suits, and sago). This narrow definition of composite commodities excludes some processed foods that could be considered composite commodities, like biscuits or salted refreshments but includes categories like "other vegetables" and "other animals" and "other dairy products" that contain quantity information. The expenditure share of the composite commodity categories has risen over time and is higher in urban areas. The third row reports spending on all items in the processed foods and beverages categories, which contain several notable composite items, items lacking quantity data, and uncertain caloric conversions - the food expenditure share of this category is much larger for urban households and it has increased by about 3 percentage points for rural and urban households over the sample period, almost doubling for rural households. The fourth row presents expenditures on cooked meals, a subset of the expenditures on processed foods, which is higher in urban areas but has actually decreased over time. Cooked meals include both restaurant meals and transfers in kind from employers so this decline need not imply a decline in restaurant meals - it could also imply increased formalization of employee-employer relations and a shift in wage versus in-kind payment. Note that the expenditure share on cooked meals remains very low compared to what is observed in wealthy countries and middle-income developing countries (CITE). The fifth row shows expenditures on alcohol as a share of food expenditures, and while there has been a 25% increase the level remains low but slightly higher in rural areas.

The sixth and seventh rows of table 6 show the share of expenditures that can be directly converted to calories using either a conservative or a liberal imputation criteria. The conservative criteria only converts calories directly for goods that both include (a)quantity units in weights or volumes (as opposed to units or missing quantities, as is the case for most beverages, processed food, cooked meals and some fruits and other goods) and (b)obvious calorie conversions (which rules out most composite commodities even if they are measured in KG). The liberal criteria attempts to convert virtually all goods directly and only excludes goods with no quantity measures. Goods with discrete units are converted to masses, and the published caloric conversion tables (from Gopalan et al. (2004) or Karan and Mahal (2005)) are supplemented with data from the IndiaMD website and other sources. The conservative criteria only covers 80% of food expenditures in urban areas and about 90% in rural areas, and the share covered declined by 2-4% over the sample period. The liberal

[8] Unless otherwise noted, all summary statistics reported are weighted using the multiplier factors provided by the surveys. We use the combined central and state samples and use data from the 17 biggest states, urban Delhi, and Meghalaya.

criteria covers over 95% and 90% of urban and rural food expenditures respectively, with a 1.4%-2.3% decrease in expenditure share. There is thus an intrinsic trade-off between measurement error induced by attempting to broaden the coverage for direct calorie conversion and the error induced by imputing the caloric content of the unconverted part of food expenditures.

We next turn to measurement of unrecorded meals to the household and meals provided to others. The expenditure data records all expenditures by the household on food and this includes food that is given to guests, as part of ceremonies, or to employees - provided they do not live with the household and therefore do not qualify as household members. An accurate measure of per capita calorie consumption by the household requires a downward adjustment to calorie consumption due to these meals to others. Conversely, each household receives free meals as guests of other households, through school or other public programs, or from employers. The NSS instructions require that these free meals not be recorded under household consumption (with their value imputed at market rates), unless their is some payment. Thus subsidized meal purchases would be recorded but free meals from school or employers would not. There is some ambiguity as meals from employers would constitute transfers in kind and should technically be recorded in the consumption data but due to uncertain valuation this is often not the case. Since some meals are received from institutional employers or schools it is not necessary that these free meals given to others and those received balance out on average.

Table 7 provides summary statistics on the share of households giving or receiving free meals, the mean number of meals given and received in the last month, and the median number of meals given or received conditional on giving or receiving meals. There is a clear pattern with rural households providing more meals to others than urban households and a reverse pattern for free meals received (until the last survey round). The pattern over time is less clear and a bit inconsistent, with some implausibly large jumps. As expected on average meals given exceed meals received, since all of the meals given would typically be recorded for both the giving and receiving household, while meals given by non-household employers, schools, government programs would not be recorded. While the distribution of meals given and received is quite skewed - with a few households hosting large ceremonies and a few households heavily dependent on free food received - the average effect is not quite large and is unlikely to significantly bias estimates of calorie consumption per capita. Table 7 also includes the quantity of purchased cooked meals consumed, with the main lesson being that cooked meals are much more important to urban than rural households and their consumption has declined, particularly in urban areas. Thus the decline in expenditure share from table 6 is not simply due to the availability of cheaper cooked meals.

A3. Calorie estimates

In light of these issues we construct several different measures of calorie consumption using different imputation schemes, which helps to clarify which basic facts are quite robust and which depend on contestable assumptions. Table 8 presents calories per capita per day using several different imputation schemes. There are three steps to the imputation procedure. We begin with either the conservative or liberal direct conversion of calories. For goods that normally have quantities reported but are sometimes missing quantities we use the median unit value (expenditure/quantity) to impute quantity, and we also censor quantities so that no household purchases a good for a unit value more than 20 times more or less than median unit value. These two steps ensure that the calorie measurements for categories with relative few quantity observations - especially processed foods - are not biased by the presence of outliers. Next we impute the non-converted part of food expenditures using either (a)calorie/rupee for directly converted goods by household, (b)the average calorie/rupee for directly converted goods across all households, or (c)the group average calorie/rupee averaged across all households. Imputation (a) allows the calorie per rupee of expenditure to vary across households, with richer households typically having lower calories per rupee of directly converted expenditure and hence less imputed calories per rupee of non-converted expenditure. Imputations (b) and (c) remove this idiosyncrasy by averaging across all households, by sector and survey round to control for differences in prices. Measure (c) allows differences in average calorie/rupee conversion rates across different food groups, which is important given the large range in calories/rupee documented later. When performing this imputation we can also consider an adjustment factor - for example, to take account of the fact that most of the unmeasured calories come from goods with generally high cost per calorie (e.g. processed foods, beverages, other meats, ice cream) we might apply a factor of 0.5 to the calories/rupee measure.[9] Finally, having imputed the calories of the missing food, we also need to consider outliers in the data, so we calculate both the uncensored mean, the median, or the trimmed mean which drops households in the top and bottom 1% of food expenditures and direct calories imputed.

The first row of table 8 presents the uncensored mean calories per capita per day using the liberal direct conversion and imputing the rest of the calories by multiplying the rest of expenditures by half of the calorie per rupee of expenditures directly converted for each household. This captures the fact that most of the imputed calories come from foods with a generally higher cost per calorie than the average directly converted basket, and allows the cost per calorie to rise with household budgets. The next five row each change one parameter at a time. The second row does uses the conservative direct conversion, meaning that a greater share of expenditures are imputed. The third row uses a one to one adjustment factor instead of a a one half factor, thereby assuming that the non-converted foods have a similar price per calorie as the directly converted expenditure. The fourth row imputes the non-converted expenditure using the sectoral annual average rather than the household-specific calorie/rupee factor. The fifth and sixth rows report the median and censored mean, which trims the 1% tails of the food expenditure and converted calorie distributions.

The seventh row of table 8 imputes the unmeasured calories using group-specific conversion factors equal to the average calorie per rupee for each group, averaged across all households. Direct imputation is done using the liberal conversion criteria (which ensures that there are at least 4 goods in each group with direct calorie conversion). Since imputation is now done by each group there is less concern about imputing the low cost per calorie of grains or pulses to goods like 'cooked meals,'

[9] Deaton and Dreze (2009) do this explicitly for cooked meals, implying that a cooked meal is equivalent to the aggregate food consumption basket with a markup of 100%.

'other processed food' and 'other beverages' so we do not multiply by one half. For comparison the eighth row assumes that the imputed goods have a calorie/rupee rate half as high as the rest of the goods in the group - this might be more reasonable for some categories, such as ice-cream (which could have twice the cost per calorie as milk), other fruit (given that coconut, singara, and dried fruits and nuts are directly converted and have high calories per rupee), and cooked meals (compared to pickles, sauces, jam/jelly, and cakes). The ninth row presents the group results of row seven but trimming the 1% tails of expenditure and calories.

Altogether, the estimates presented in table 8 strongly suggest that there has been a large decline in calories per capita for rural households and that rural households in 1983 consumed significantly more calories than urban households on average. However, there is some uncertainty about whether urban calories per capita have risen or declined and whether calories per capita in urban areas exceed those in rural areas in 2004-05. These results are sensitive to the imputation method. Using medians we sometimes find a modest increase in calories per capita in urban areas, though the range in table 8 is quite small at -74 to 18. Using group-specific, average or higher calorie/rupee adjustment factors also tends to shift the rural-urban gap in 2004-05 in favor of urban households.

The bottom two rows of table 8 present our two preferred specifications, corresponding to row (6) and row (9), but adding in calories from alcohol and the effect of a 'household adjustment factor.' This factor accounts for free meals and meals to others by assuming that they have the same calories per capita of other meals consumed by the household. The precise formula used is

$$hh.\ adj.\ factor = \frac{pay\ meals\ at\ home\ + pay\ meals\ outside\ + free\ meals}{pay\ meals\ at\ home\ + pay\ meals\ outside\ + meals\ to\ others} \tag{6}$$

Note that the 55th survey round (1999-2000) did not record meals to others so it is excluded from this calculation, even though one can include a positive inflation factor accounting for free meals consumed. Comparing rows (10) and (11) to (6) and (9) we see that these last two adjustments have a minimal effect. The adjustments tend to increase calories per capita in urban areas but by a greater amount in the early period. In rural areas the pattern is reversed, with a slightly negative adjustment in the early period and positive in the later period. The net effect is thus to decrease the fall in calories in rural areas and increase in the fall in calories in urban areas, and a modest reduction in the rural-urban gap. The magnitude of the effect overall is at most 20%. Throughout the rest of the paper we use the estimate of row 10 as our baseline measure of caloric intake and check it against the other measures, noting the differences only if they are economically significant.

A final issue that we cannot address with our data is that the nutritional content of particular foods may vary over time and space. Many foods lose some of their nutritional content with transportation over longer distances and storage, the composition of the 'other' goods may vary systematically over different areas and periods, and the caloric content of processed foods may also vary. To the extent that transportation lowers caloric content for goods that we measure this would tend to decrease urban relative to rural calories and might also lower caloric intake further over time. For goods with unknown caloric content, our imputation procedure may capture some of these effects, as areas and periods with higher calories per rupee for directly converted goods might also have higher calories per rupee for imputed goods, but we cannot be certain.

B Data Appendix: Imputing caloric requirements

To record time-use information, the surveyors attempted to interview each member of the household over age 5 about their time-use over the preceding 24 hour period. Busy, reluctant or incapable members had their time-use recalled by another household member. Time-use was captured for up to three separate types of days - normal, abnormal, and variant - to capture variations in the weekly schedule including market days, weekend activities, etc. The measure we use is based on a weighted average of these three days based on how many days the household reported of each type in the preceding week. The interview team included both a male and female interviewer as the goal of the survey was to measure and validate the contribution of women to economic life in India.

The survey also records a number of other variables that are recorded in the same format as in the NSS consumption surveys - monthly expenditures, land ownership, religion, and scheduled caste/tribe at the household level and age, gender, education, and occupation for each household member. Unfortunately, the Time Use Survey was not carried out simultaneously with the NSS consumption survey, which means that comparable consumption data is only available for the July 1997-June 1998 period or the July 1999-June 2000 period. The closest geographical match is at the district level as individual villages and cities are not recorded or geocoded.

To go from time-use to caloric requirements, we use two methodologies. The first is based on classifying the 154 different types of activities into four different levels energy requirement, using an on-line calorie per activity calculator to guide our personal judgment. The intensity level of each activity is relative to a complete state of rest (activity level 0), and we get the intensity factors from the calories per hour website at http://www.caloriesperhour.com/. The intensity ratios (relative to complete rest) are 1.7 for level 1 (playing cards), 3 for level 2 (cooking/housework/walking 2mph), and 5 for level 3 (chopping wood/push-mowing). To give some concrete examples from our data, activity level 3 includes ploughing, preparing land, cleaning of land, wood cutting, chopping and stocking of firewood, and building and construction of dwellings. Activity level 2 includes cooking, sweeping, and assembling machines, equipment and other products. Activity level 1 corresponds to sedentary labor such as service in government, professional work, reading, and watching tv. Activity level 0 corresponds to sleeping or 'doing nothing, rest and relaxation.' As an alternative we also hired a research assistant to match each activity in our data to an activity from the website at the most detailed level possible. The results are quite similar – the most notable difference is that we classified transport activity as level 2 but our RA matched transport activities to something closer to level 3. Our classification of activities into different levels of intensity and matching scheme are available upon request. Assuming that households sleep 8 hours a day, spend 8 hours awake at intensity level 1 and then another 8 hours working at intensity levels 1/2/3 for heavy, moderate, or sedentary market work a 26 year old man weighing 70 kilograms would require 3952/2928/2272 calories. This lines up roughly with the ICMR recommendations of 3800/2875/2424 calories.

We take as the baseline caloric requirements those corresponding to a 70 KG 26 year old man. We then convert this energy requirement by a multiplicative factor corresponding to the average relative Basal Metabolic Rate (BMR) for a given age and gender. The BMR captures the energy consumed by the body at a complete state of rest for a given age and gender, and it multiplicatively scales the energy requirement of different activities that consume more energy than resting. BMR typically rises and falls in age, starting out higher for women but peaking earlier. Our baseline female is 62 KG. For children under age 6 we use the daily energy requirements from the India Council for Medical Research (ICMR). For infants aged 0-6 months and 7-12 months, for which the ICMR gives energy requirements by weight, we use energy requirements for 1-3 year olds. These should be a reasonable approximation of calorie requirements based on an average child growth chart plus an extra energy requirement for lactating mothers. The NSS data do not report pregnancy status so we are likely to underestimate the calorie requirements for pregnant women by about 300 calories per day according to the ICMR.

The ICMR provides daily energy requirements for adult men and women as well children of different ages, but adult caloric requirements are only divided into three activity cells - heavy, moderate, sedentary. They also do not take account of activity levels by children, an important omission given that they have separate age/gender cells for boys/girls aged 13-15 and 16-18, age ranges where child labor inside and outside the household is likely to be quite important in some areas. The ICMR theoretically provides us with an alternative set of energy requirements for analysis but we prefer our measure for several reasons – it allows us to account for household age and life-cycle effects for adults and labor by children and adolescents, we can match energy requirements to a variety of household characteristics rather than industry or occupation (which would be the only way of imputing household calorie requirements in the NSS using the ICMR recommendations), and we have a much more fine-grained measure of energy requirements that has both an extensive margin – number of hours working on different activities – and an intensive margin – requirements for activities of different intensity.

The most important limitation of the TUS data is that we do not have a measure of the intensity of individual activities. While many agricultural tasks are likely to be highly labor intensive some may have assistance from mechanical and animal energy sources. This issue also occurs for all transportation related activities - since the TUS does not record mode of transport, we assume an activity level of 2 which would tend to overstate energy requirements for motorized vehicular transport but understate energy requirements for walking and cycling. Other limitations include the lack of data on height and weight for individuals or systematic biases in activity recall.

C Biomechanics Appendix

While we have shown that data on caloric needs and anthropometrics can qualitatively resolve our three food consumption "puzzles" in this section we delve into the biomechanics to explore feedback from height and weight to caloric needs, and from excess caloric intake over caloric needs on weight. This allows us to assess the quantitative fit between our three data sources – caloric intake, caloric needs, and anthropometrics – and provide some insight into mismeasurement and within-household distribution issues.

C1. Feedback from height/weight to caloric needs

Our caloric need estimates are based on basal metabolic rates for a given age and gender for a given height and weight (70KG for men, 62KG) for women. These are high relative to the weights used by the Indian Council for Medical Research caloric needs calculations (60KG and 50KG respectively) and the anthropomorphic data (56KG and 49KG). While this is largely irrelevant for our qualitative exercise since it scales all individuals equally (conditional on gender composition) it matters for measuring the exact caloric balance of individuals because the basal metabolic rates (BMR) that determine caloric requirements at rest are reasonably sensitive to weight. It also matters if there are systematic differences in height and weight across rural/urban areas, cohorts, household sizes and ages. In general we would expect these effects to partially offset any excess caloric intake over caloric needs since those households will on average be taller and heavier, increasing their caloric needs.

To assess the quantitative importance of this channel, we rely on the anthromoporphic data combined with the widely-used revised Harris-Benedict BMR formula which is based on regression analysis of clinically measured BMR.[10] The formula for women is

$$(13.397 * weight\ in\ KG) + (6.25 * height\ in\ cm) - (5.677 * age) + 5 \qquad (7)$$

and the formula for men is

$$(9.247 * weight\ in\ KG) + (3.098 * height\ in\ cm) - (4.33 * age) + 447.593 \qquad (8)$$

Height is important because the height to weight ratio is typically correlated with "fat-free mass" which is one of the primary determinants of the basal metabolic rate – evidence from the scientific literature suggests that about 75% of the variation in BMR across individuals can be explained by height, weight, age and gender leaving about 25% to idiosyncratic metabolic differences.[11] While fat-free mass is the ideal predictor, the Harris-Benedict formula using height, weight and age is the best we can do with the data available.

Figure 22 presents the differences in caloric needs due only to differences in height, wage, and age across several dimensions. We omit young children from the analysis and pool men and women where relevant, using the same regression controls as we did earlier with the BMI regressions. Note that several of these figures look similar to the body-mass index graphs presented earlier owing to the high correlation between BMI and BMR (0.74 for adults). The top left panel indicates that urban households

[10]We use the formula from A.M. Roza and H.M. Shizgal "The Harris Benedict equation reevaluated." American Journal of Clinical Nutrition. Vol. 40, No. 1 (July 1984): 168-182.

[11]Johnstone et a. (2005) "Factors influencing variation in basal metabolic rate include fat-free mass, fat mass, age, and circulating thyroxine but not sex, circulating leptin, or triiodothyronine" American Journal of Clinical Nutrition, vol.82(5):941-948.

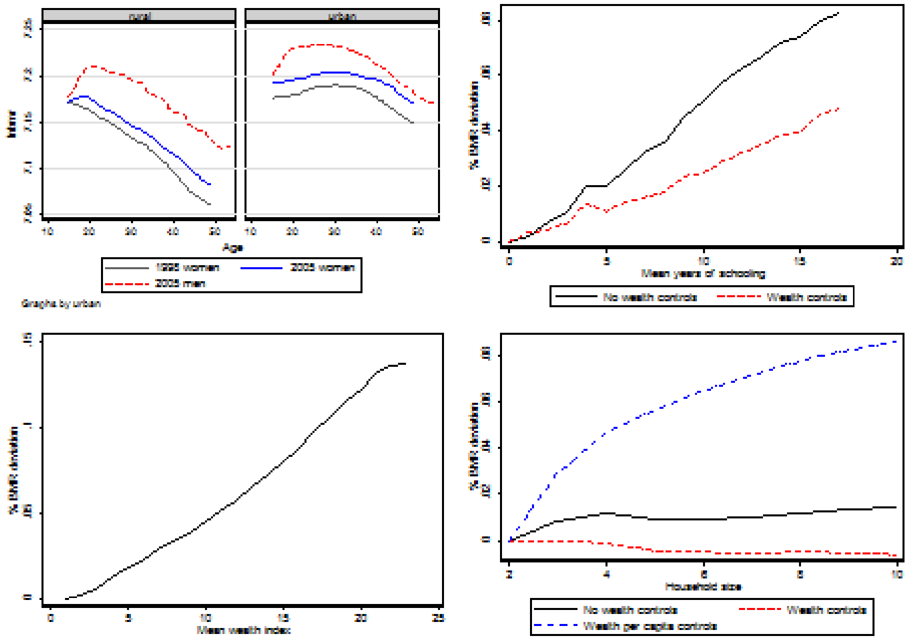

Figure 22: Variation in adult BMR due to age, height and weight

would have BMR about 4% higher than rural households, which offsets about one third of the 12% difference in caloric needs that we measure assuming similar BMR across rural and urban households. This would be smaller after controlling for the fact that urban households are wealthier – the wealthiest households in the data have up to 15% higher BMR coming from their greater height and weight. The increase over time between the 1998 and 2005 cohorts is quite small at about 2%. The decline in BMR from its peak around age 20 to age 50 is slightly smaller for rural households than the average age conversion factors we used earlier though this figure has no controls for cohort effects; urban households, which the data indicate gain more weight over the life-cycle, have a much smaller decline in BMR with age.

Higher BMR for more educated households (conditional on household wealth) generates about 4% higher caloric needs, compared to the 15% difference we measure based on activity levels assuming constant BMR. Finally, the results for household size are difficult to interpret given the impossibility of controlling for per capita expenditure with the anthropometric data – conditional on the total wealth index larger households have lower BMR but conditional on the wealth index per capita, BMR rises significantly in larger households to a degree (8%) that would more than offset the decline in caloric requirements due to lower activity levels (about 4%). As we noted earlier the "wealth per capita index" is difficult to interpret since our wealth index is a count of the number of household durables and amenities.

C2. Weight gain and excess calories

A commonly used but naive formula for weight loss states that 7700 excess or defficient calories will generate a kilogram of weight gain or loss.[12] In fact the effects of caloric surplus or defficiency are highly non-linear both because caloric surplus and defficiency that lead to weight gain and loss affect caloric needs through the feedback effect described above, and because the basal metabolic rate can respond by adjusting energy routed to tasks like bone maintenance and internal organ functioning. Aggregate evidence from weight loss studies indicates that people on diets (defined as a caloric deficit of 500 or more) only lose about 5 KG per year, when the standard formula would predict a 23.7KG weight loss – this suggests that for weight loss, the naive formula is off by about a factor of 5.[13] While we would typically expect weight gain to be asymmetric since body has a harder time increasing the BMR than decreasing it, people who were previously calorie constrained may see a substantial rise in their BMR in response to additional calories consumed.

Figure 23 presents evidence that the average urban women between ages 14 and 42 gained 4 to 5 kilograms of weight between 1998-2005, while rural women gains 2 to 2.5 kilograms. While these averages may be affected by attrition it is unlikely to be a major factor over this age range. Based on the naive formula, this weight gain implies excess daily calories of 12 to 15 in urban areas and 6 to 7.5 in rural areas. Using the evidence from weight loss studies and assuming symmetry, this level of weight gain would imply 60-75 daily excess calories for urban women and 30-37.5 daily excess calories for rural women.

The average daily excess calories (per capita) for households with adult females between the ages of 14 and 45, using 1998 time-use data and 1999-2000 consumption data, is 25. This is well within reason given the medical literature. However, these numbers are -75 and 228 for rural and urban – this strongly suggests that we are underestimating caloric surplus for rural households and overestimating caloric surplus for urban households, although we cannot say whether this is coming from the caloric intake side (including unequal distribution within the household) or the caloric need side.

[12] E.g. the US government website nutrition.gov – http://www.nutrition.gov/weight-management/strategies-success/interested-losing-weight.

[13] See Marion J. Franz et al. (2007). "Weight-Loss Outcomes: A Systematic Review and Meta-Analysis of Weight-Loss Clinical Trials with a Minimum 1-Year Follow-Up," Journal of the American Dietetic Association, Volume 107(10):1755-1767.

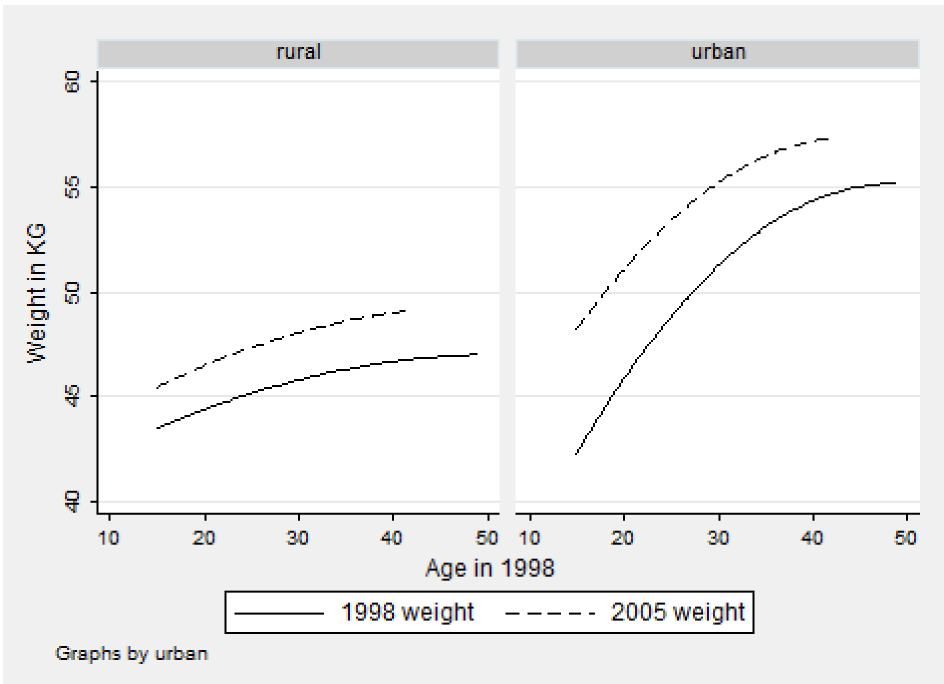

Figure 23: Weight gain 1998-2005 for women, by age

Table 3: Estimates of mean per capita calorie consumption in India

Authors	Sector	1983	1987-88	1993-94	1999-00	2004-05	Δ 1983-2005
Deaton and Drèze	Rural	2240	2233	2153	2148	2047	-193
(2009)	Urban	2070	2095	2073	2155	2021	-49
Chatterjee, Rae	Rural		2135	2100	2097		
and Ray (2007)	Urban		2073	2091	2169		
Kumar and Dey	Rural	2205			2332		
(2007)	Urban	1972			2440		
Meenakshi and	Mean	2219			2132		
Vishwanathan (2003)	Median	2076			2024		
NSS for NNMB states		2131	2139	2076	2020	1960	-171
(Deaton and Drèze)	Year	1975-79	1988-90	1996-97	2000-01	2004-05	
NNMB		2340	2283	2108	1954	1907	-405
			Our estimates				
Group imputation	Rural	2313	2285	2234		2140	-172
	Urban	2230	2234	2214		2136	-94
All food imp.	Rural	2320	2293	2244		2154	-166
	Urban	2178	2180	2192		2121	-58

Meenakshi and Vishwanathan (2003) report data by state for both sectors combined.

NNMB are the independent estimates from the National Nutritional Monitoring Bureau reported in Deaton and Dreze (2009), which cover a subset of states. Above are NSS estimates from Deaton and Dreze (2009) for the same set of states in nearby years.

Table 4: Comparison of Time Use Survey and NSS

	Sector	TUS Mean	TUS Median	NSS Consumption Mean	NSS Consumption Median
		Requirements		Intake	
Per capita calories	Rural	2363	2323	2236	2088
	Urban	2091	2122	2327	2180
Per capita calories alt.	Rural	2491	2473	2232	2095
	Urban	2200	2274	2277	2163
MPCE	Rural	459	400	505	429
	Urban	804	694	947	734
Hhsize	Rural	4.07	4	4.82	5
	Urban	4.10	4	4.40	4
Age of head	Rural	43.20	40	44.40	42
	Urban	42.32	40	43.43	42
Male head	Rural	0.90		0.90	
	Urban	0.91		0.90	
Adult males	Rural	1.10		1.10	
	Urban	1.19		1.10	
Adult females	Rural	1.10		1.19	
	Urban	1.10		1.16	
Years schooling	Rural	3.58	3	3.20	2
	Urban	8.33	8.8	7.65	7.5

Table 5: Minutes per day on various activities, by sector and gender

Activity	Household		Male adult		Female adult	
Sector	Rural	Urban	Rural	Urban	Rural	Urban
Primary	649.75	73.94	314.13	38.14	153.28	18.13
Free collection	73.33	17.13	12.22	2.39	38.10	10.25
Secondary	98.22	192.74	52.46	107.57	15.27	20.10
Tertiary	113.87	485.63	69.81	305.96	12.62	41.83
Total Market	935.16	769.44	448.62	454.06	219.27	90.32
Cook	229.04	233.30	5.40	5.99	161.51	171.63
Other hh maint.	230.15	241.47	23.06	19.56	137.30	157.02
Care for others	65.84	71.04	10.07	10.62	47.27	55.23
Total Nonmarket	525.02	545.81	38.54	36.18	346.09	383.88
Learning	248.41	317.16	7.83	18.46	2.31	12.12
Social	262.69	515.60	56.81	118.41	34.55	113.95
Sleep	1841.55	1817.84	528.54	503.76	515.28	511.11
Television	104.39	313.43	27.27	74.14	23.51	91.94
Other	1024.33	747.90	332.41	235.48	298.99	236.69
Total Leisure	3481.37	3711.93	952.86	950.24	874.64	965.80

Note: children under 6 are excluded from the household measure.

Table 6: Problem foods for calorie imputation

		1983	1987-88	1993-94	1999-00	2004-05
Share of food expenditure by problem category						
		38	43	50	55	61
No quantity	Rural	0.011	0.016	0.021	0.022	0.036
	Urban	0.028	0.035	0.042	0.041	0.057
"Other"	Rural	0.028	0.028	0.028	0.028	0.035
	Urban	0.025	0.026	0.031	0.033	0.045
Proc. food and bev.	Rural	0.052	0.064	0.066	0.073	0.085
	Urban	0.144	0.150	0.157	0.146	0.159
Cooked meals	Rural	0.016	0.019	0.012	0.014	0.013
	Urban	0.059	0.060	0.056	0.046	0.046
Alcohol	Rural	0.012	0.013	0.013	0.014	0.015
	Urban	0.009	0.011	0.011	0.012	0.013
Share of food exp. with calorie conversions						
Conservative	Rural	0.906	0.891	0.889	0.880	0.865
	Urban	0.808	0.797	0.788	0.799	0.787
Liberal	Rural	0.972	0.965	0.967	0.963	0.949
	Urban	0.909	0.904	0.900	0.912	0.895

Table 7: Cooked meals, meals to other households and free meals (per 30 days)

		1983	1987-88	1993-94	1999-00	2004-05
		Cooked meals				
Mean number	Rural	1.985	2.512	1.561	1.606	1.537
	Urban	6.574	6.557	5.577	4.580	4.468
Share consuming	Rural	0.074	0.092	0.063	0.049	0.058
	Urban	0.154	0.171	0.144	0.122	0.127
Cond. Median	Rural	12	12	12	16	12
	Urban	30	28	30	27	20
		Meals to guests, employees, ceremonies				
Mean number	Rural	14.650	10.311	10.429	.	7.862
	Urban	10.483	12.178	6.208	.	6.039
Share consuming	Rural	0.407	0.377	0.141	.	0.447
	Urban	0.354	0.350	0.104	.	0.382
Cond. Median	Rural	10	10	12		8
	Urban	10	10	12		7
		Free meals				
Mean number	Rural	7.717	6.572	6.005	6.220	11.473
	Urban	8.152	6.993	7.040	6.342	7.836
Share consuming	Rural	0.261	0.228	0.193	0.179	0.329
	Urban	0.235	0.218	0.196	0.177	0.233
Cond. Median	Rural	14	12	16	20	24
	Urban	18	16	20	20	22

Table 8: Daily calories per person: different imputations

Direct conv.	Cal./rupee + adj. fact.	Stat.	Sect.	1983	1987-88	1993-94	1999-00	2004-05	Δ 1983 to 2005
Lib.	Ind x0.5	Mean	Rural	2350	2302	2226	2217	2121	-229
			Urban	2156	2165	2128	2201	2085	-70
Cons.	Ind x0.5	Mean	Rural	2305	2295	2213	2201	2105	-200
			Urban	2124	2150	2107	2170	2057	-67
Lib.	Ind x1	Mean	Rural	2377	2337	2261	2255	2171	-206
			Urban	2227	2254	2223	2287	2189	-39
Lib.	Avg. x0.5	Mean	Rural	2358	2312	2233	2223	2130	-229
			Urban	2214	2235	2189	2254	2159	-55
Lib.	Ind x0.5	Median	Rural	2158	2150	2107	2099	2027	-131
			Urban	2007	2046	2045	2114	2025	18
Lib.	Ind x0.5	Mean	Rural	2328	2297	2229	2217	2124	-205
		1% trim	Urban	2141	2154	2147	2203	2092	-49
Lib.	Gr.avg. x1	Mean	Rural	2341	2293	2215	2209	2104	-237
			Urban	2159	2208	2141	2243	2095	-64
Lib.	Gr. avg. x0.5	Mean	Rural	2327	2274	2200	2188	2081	-246
			Urban	2106	2135	2078	2169	2032	-74
Lib.	Gr. Avg. x1	Mean	Rural	2321	2289	2219	2212	2110	-211
		1% trim	Urban	2168	2202	2169	2250	2106	-62
Including calories from alcohol and hh. adj. factor									
Lib.	Ind x0.5	Mean	Rural	2320	2293	2244		2154	-166
		1% trim	Urban	2178	2180	2192		2121	-58
Lib.	Gr.avg. x1	Mean	Rural	2313	2285	2234		2140	-172
		1% trim	Urban	2230	2234	2214		2136	-94