

Hazarika, Gautam; Paul, Sourabh Bikas

**Working Paper**

## India's Calorie Consumption Puzzle: Insights From the Stochastic Cost Frontier Analysis of Calorie Purchases

IZA Discussion Papers, No. 13722

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Hazarika, Gautam; Paul, Sourabh Bikas (2020) : India's Calorie Consumption Puzzle: Insights From the Stochastic Cost Frontier Analysis of Calorie Purchases, IZA Discussion Papers, No. 13722, Institute of Labor Economics (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/227249>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

DISCUSSION PAPER SERIES

IZA DP No. 13722

**India's Calorie Consumption Puzzle:  
Insights From the Stochastic Cost Frontier  
Analysis of Calorie Purchases**

Gautam Hazarika  
Sourabh Bikas Paul

SEPTEMBER 2020

## DISCUSSION PAPER SERIES

IZA DP No. 13722

# India's Calorie Consumption Puzzle: Insights From the Stochastic Cost Frontier Analysis of Calorie Purchases

**Gautam Hazarika**

*University of Texas Rio Grande Valley and IZA*

**Sourabh Bikas Paul**

*Indian Institute of Technology Delhi*

SEPTEMBER 2020

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9  
53113 Bonn, Germany

Phone: +49-228-3894-0  
Email: [publications@iza.org](mailto:publications@iza.org)

[www.iza.org](http://www.iza.org)

## ABSTRACT

---

# India's Calorie Consumption Puzzle: Insights From the Stochastic Cost Frontier Analysis of Calorie Purchases\*

Between the early 1970s and very nearly the present, Indians' per capita calorie consumption declined. This decline, perplexing in the face of rising per capita income when malnutrition is rampant, has been termed India's Calorie Consumption Puzzle. It has been partially attributed to a squeeze in the household food budget. This study employs Stochastic Cost Frontier Analysis to evaluate this explanation, upon the logic that such a squeeze shall likely result in the rising cost-efficiency of calorie purchases, that is, the more economical purchase of calories. Analysis of household expenditure data from India's National Sample Survey reveals that Indian households' purchase of calories did become more cost-efficient at every level of income, suggesting that there was indeed a squeeze in the household food budget, making this a viable explanation of the Calorie Consumption Puzzle. Besides thus investigating India's Calorie Consumption Puzzle, this study demonstrates a novel application of Stochastic Cost Frontier Analysis, to consumption instead of the more common production, in that the method has not previously been applied to the consumption of multiple items treated as inputs yielding an output. Stochastic Cost Frontier Analysis applied to calorie acquisition may be a new way of gauging changes over time in food security, with a rise in cost-efficiency indicating a squeeze in the food budget or declining food security.

**JEL Classification:** I32, O1

**Keywords:** calorie consumption puzzle, Stochastic Cost Frontier Analysis

**Corresponding author:**

Gautam Hazarika  
Department of Economics and Finance  
The University of Texas Rio Grande Valley  
Brownsville, Texas 78520  
USA

E-mail: gautam.hazarika@utrgv.edu

---

\* This paper is forthcoming in Empirical Economics.

## 1. Introduction

This study uses the method of Stochastic Cost Frontier Analysis to examine changes over time in the cost-efficiency of Indian households' purchase of calories. It hopes thereby to evaluate the thesis that per capita calorie consumption in India declined until recently due partly to a squeeze in the household food budget, upon the logic that this ratcheted pressure is diagnosable by the rising cost-efficiency of households' purchase of calories. Further, whereas the cost-efficiency analysis of production has been undertaken often, to our knowledge this study is the first application of Stochastic Cost Frontier Analysis to the consumption of multiple items treated as inputs yielding an output<sup>1</sup>. That typical 'output' in consumption, consumer utility, is unobserved has prevented this manner of cost-efficiency analysis. However, output from the consumption of food in the form of calories, carbohydrate, protein, fat, or micronutrients is computable, making the cost-efficiency analysis of, for example, calorie acquisition, feasible. In sum, the contribution of this paper consists in its novel application of Stochastic Cost Frontier Analysis to calorie acquisition, towards diagnosing a squeeze in the household food budget or decline in food security.

Between the early 1970s and about 2010, Indians' average consumption of calories per capita fell<sup>2</sup> (NSSO, 2014), this despite unambiguous economic growth, which has been rapid since the liberalization of the economy in 1991. The phenomenon has been termed the Calorie Consumption Puzzle (Chandrashekhar and Ghosh, 2003). It is a puzzle for two reasons. First, Indians' calorie Engel curve is positively sloped, that is, household per capita calorie consumption increases in household per capital income (proxied by expenditure) cross-sectionally. It is, thus, puzzling that this positive relationship doesn't hold longitudinally at the level of the nation. Second, anthropometrics indicate that a very large proportion of Indians, children as well as adults, is malnourished. By

---

<sup>1</sup> While Stochastic Frontier Analysis *has* previously been applied to consumption, in that stochastic demand frontiers (for example, Filippini and Hunt, 2012) and hedonic price frontiers (for example, Lee, Park, Oh, and Kim, 2008) have been estimated, the Stochastic Cost Frontier Analysis of the consumption of multiple items, treated as inputs yielding an output, has not been undertaken.

<sup>2</sup> Whereas average daily calorie consumption rose slightly after 2010, it remained lower in 2011-12, the latest period for which data is publicly available, than in 1972-73.

UNICEF, 21 percent of Indian children under the age of 5 were wasted, 38.4 percent were stunted, and 35.7 percent were underweight in 2015-16. These are among the highest rates of child malnutrition in the world. According to the UN's FAO, about half of the adult population is underweight. If this pervasive anthropometrically diagnosed malnutrition bespeaks widespread hunger, it is puzzling that rapid economic growth and rising incomes did not lead to greater calorie consumption.

In a widely cited article, Deaton and Dreze (2009) closely document the decline in per capita calorie consumption between 1983 and 2004-05, establishing that Indians' calorie Engel curve shifted downward over this period, that is, per capita calorie consumption declined at every level of real per capita income (proxied by expenditure). Of their proffered explanations, two seem to have been especially thought provoking in having been taken up by other researchers. They are: (a) Indians' caloric needs have fallen, and (b) there has been a squeeze in the household food budget. This paper takes a novel approach, based on scrutiny of households' costs of acquiring calories, to assessing a squeeze in the food budget. Whereas Deaton and Dreze (2009) observe that Indians have been consuming more expensive calories even as their calorie consumption has declined, and interpret this as casting "some doubt on the hypothesis of a squeeze in the food budget", a more sophisticated, structural approach to examining changes over time in the cost of calories is warranted. We take just such an approach, based on novel use of Stochastic Cost Frontier Analysis. We ask the question: has the cost-efficiency of calorie purchases changed over time? A rise in cost-efficiency, we argue, would be consistent with a squeeze in the household food budget.

Researchers conceive of two reasons for this squeeze. Some believe there has been widespread impoverishment, especially in rural India, from the fraying of the social safety net following liberalization of the economy in 1991 (for example, Mehta and Venkatraman, 2000, Patnaik, 2004, 2007). However, Deaton and Dreze (2009) argue that this is contradicted by rising real per capita income across all percentiles of the distribution of per capita income (proxied by expenditure). Besides, widespread impoverishment would not explain the fall in per capita calorie

consumption at every level of real per capita income (proxied by expenditure), nor the fact that the fall in calorie consumption was more pronounced among those at the upper end of the distribution of per capita income, whose real incomes have unequivocally risen. Hence, others (for example, Sen, 2005) have proposed that the squeeze in the household food budget originates in rising non-food expenditures, since with economic development arrived a proliferation of competing budgetary items: services and non-food goods. In other words, Sen (2005) suggests that any squeeze in the food budget is self-exerted. This naturally complicates interpretations of food security. By Sen's (2005) argument, a household may become less food-secure even as its welfare presumably improves<sup>3</sup>.

Studies that seek to explain the calorie consumption puzzle on the basis of reduction in caloric needs investigate two factors in this reduction. First, India's pernicious disease environment has improved (Duh and Spears, 2017, Siddiqui, Donato, and Jumrani, 2019). Disease tends to raise caloric needs by, for example, impeding the absorption of calories and triggering immunological responses that are energy-intensive. Second, Indians' level of physical activity has declined due to, for example, the changing nature of work (Eli and Lee, 2015).

If the Stochastic Cost Frontier Analysis of calorie purchases revealed rising cost-efficiency, this may be taken to be consistent with a squeeze in the household food budget, making it a viable explanation of the Calorie Consumption Puzzle. If the analysis revealed falling cost-efficiency, this may be regarded consistent with the loosening of food budgets produced by reduced caloric needs. Naturally, a rise in cost-efficiency wouldn't rule out reduced caloric needs as a factor in the Calorie Consumption Puzzle since the rise may be a net rise, that is, the squeeze in the food budget a net squeeze. This is the crux of this study's empirical strategy. Analysis of expenditure data from India's National Sample Survey demonstrates that Indian households' purchase of calories did become more cost-efficient at every level of income, suggesting that there was indeed a squeeze in the household

---

<sup>3</sup> On the other hand, the presumption that a voluntary take-up or purchase must be welfare improving, has been questioned by such advocates of randomized controlled trials as Esther Duflo, who has called it "the moronic revealed preference argument" (Parker, 2010), upon the grounds that poor choices are common enough. Given mounting evidence of the ill effects of poverty on cognition (for example, Mani, Mullainathan, Shafir, and Zhao, 2013), the impoverished may be particularly prone to poor choices.

food budget capable of lowering calorie consumption at every level of income, that is, of causing downward shifts of the calorie Engel curve. Declining food security is, then, a plausible explanation of the Calorie Consumption Puzzle.

The remainder of this study is organized as follows. Section 2 presents the theoretical model used to motivate a caloric cost frontier. Section 3 describes the particular variant of Stochastic Cost Frontier Analysis employed, and the utilized data. Section 4 presents the empirical results, and Section 5 a brief conclusion.

## **2. Motivating a Caloric Cost Frontier**

Assume that calories,  $C$ , are derived from the purchase of two foods, whose quantities are designated  $F_1$  and  $F_2$ . The continuous production function of calories is  $C(F_1, F_2)$ . Even though physical quantities of foods are customarily converted into calories by the means of fixed, food-specific, energy conversion factors, function  $C$  might be quasi-concave instead of linear were calories considered *metabolizable* calories. Calories from the purchase of a food item may not be perfect substitutes for those from the purchase of another for at least two reasons. First, the palatability of a food may decline at the margin, leading to rise in marginal wastage. Second, its digestibility too may decrease at the margin. Hence, the marginal rate of technical substitution between the two foods along an iso-calorie curve may be taken to be decreasing, in that set reductions in a good may need to be offset by progressively larger quantities of the other.

Even the developing world's poor don't view food solely as a source of calories. Banerjee and Duflo (2007) note that "Even for the extremely poor, for every 1 percent increase in the food expenditure, about half goes into purchasing more calories, and half goes into purchasing more expensive (and presumably better tasting) calories." It would appear that food for these poor has recreational as well as caloric value. The authors also note that the extremely poor commonly purchase fewer calories than they can afford, choosing to consume non-food items instead, of which many are recreational. These include alcohol, tobacco, consumer durables, and festivals. In sum,



recreation, designated  $R$ , appears to play a prominent role in the consumer utility of the poor, and derives from the consumption of both food and non-food items. Hence, assume that the continuous and twice continuously differentiable quasi-concave production function of recreation is  $R(F_1, F_2, N)$ , where  $N$  denotes quantities of a composite non-food item.

Assume that consumers derive utility from calories and recreation according to the continuous and twice continuously differentiable quasi-concave utility function  $U(C, R)$ . A consumer may be taken to maximize  $U(C, R) = U[C(F_1, F_2), R(F_1, F_2, N)]$  subject to her budget constraint  $I = P_1F_1 + P_2F_2 + P_3N$ , where  $I, P_1, P_2$ , and  $P_3$  denote, respectively, her income, the price of food item 1, that of food item 2, and that of the non-food item. The following are the first-order conditions for optimality,  $\lambda$  denoting the Lagrange multiplier<sup>4</sup>:

$$U_1C_1 + U_2R_1 - \lambda P_1 = 0, \quad (1)$$

$$U_1C_2 + U_2R_2 - \lambda P_2 = 0, \quad (2)$$

$$U_2R_3 - \lambda P_3 = 0, \quad (3)$$

$$I - P_1F_1 - P_2F_2 - P_3N = 0. \quad (4)$$

If there were no recreational value to food, (1) and (2) above would be modified as, respectively,  $U_1C_1 - \lambda P_1 = 0$  and  $U_1C_2 - \lambda P_2 = 0$ , so that dividing the one by the other yields

$$C_1/C_2 = P_1/P_2, \quad (5)$$

the necessary condition for the minimization of the cost of calories. Food purchases that satisfy (5) may be considered making up the deterministic Caloric Cost Frontier.

On the other hand, dividing (1) by (2) yields,  $(U_1C_1 + U_2R_1) / (U_1C_2 + U_2R_2) = P_1/P_2$ , or

$$[C_1 + (U_2/U_1)R_1] / [C_2 + (U_2/U_1)R_2] = P_1/P_2. \quad (6)$$

Observe that the left-hand side of (6) approaches  $C_1/C_2$ , the left-hand side of (5), that is, calories are purchased more cost-efficiently, as  $(U_2/U_1)R_1$  and  $(U_2/U_1)R_2$  decrease. Hence, calories are purchased more cost-efficiently as: (a)  $R_1$  and  $R_2$  decrease, that is, the marginal productivity of food in the

---

<sup>4</sup> The expression  $f_j$ ,  $f = C, F, R$ , and  $U$ , denotes the first derivative of the function  $f$  with respect to its  $j$ th argument.

production of recreation decreases, (b)  $U_2$ , the marginal utility from recreation, decreases, or (c)  $U_1$ , the marginal utility from calories, increases.

Given diminishing marginal utility from calories, a fall in spending on food, caused by, for example, a decrease in income, shall increase the marginal utility from calories,  $U_1$ . Hence, factors that increase the marginal utility from calories may be described as ‘squeezing’ food budgets in having the same effect as a fall in spending on food, whereas those that lower this marginal utility may be considered loosening food budgets. It follows that a squeeze in the food budget, by increasing  $U_1$ , potentially decreases  $(U_2/U_1) R_1$  and  $(U_2/U_1) R_2$ , promoting the more cost-efficient purchase of calories.

Since the marginal utility from calories is likely positively related to caloric need (one is “as hungry as a hunter”, for example), there may be more pressure on the food budgets of workers in more calorie-intensive strenuous occupations. This may be true as well of the habitants of more diseases-prone areas since illness can impede the absorption of calories and immunological responses are calorie-intensive. Hence, the calorie purchases of those in more strenuous occupations and the habitants of more diseases-prone areas may be more cost-efficient.

The model also supports the argument of a squeeze in the food budget arising from widening of the menu of services and non-food goods available to consumers. With economic development has come an increase in consumer access to services and non-food goods, particularly in urban India. This may be considered to have raised the marginal product of the model’s composite non-food good in the production of recreation, that is,  $R_3$  in (3). For example, greater access to consumer electronics may have raised the recreational value of non-food spending. A rise in  $R_3$  will lead to the substitution of the non-food good for the two food items. The resulting decrease in calorie consumption and consequent increase in the marginal utility from calories,  $U_1$ , may lead to the left-hand side of (6) approaching that of (5), that is, the more cost-efficient purchase of calories.

### 3. Empirical Model and Data

Consider the stochastic frontier cost function rendered in logs as

$$\ln E_i = f(\ln C_i, \ln P_{i1}, \dots, \ln P_{in}) + v_i + u_i \quad (7)$$

where  $E_i$  denotes the expenditure incurred by household  $i$  in the purchase of metabolizable calories  $C_i$ ,  $P_{ij}$ ,  $j = 1, \dots, n$ , signifies the price of food item  $j$  confronting the household, the random errors  $v_i$ , independent of  $u_i$ , are iid  $N(0, \sigma_v^2)$ , and the  $u_i$ , non-negative random variables associated with cost-inefficiency, are obtained by truncation at zero of the distribution  $N(\mathbf{z}_i\boldsymbol{\delta}, \sigma_u^2)$ ,  $\mathbf{z}_i$  being a vector of explanatory variables, termed environmental variables, related to household  $i$ 's cost-efficiency at purchasing calories. This is a version of Battese and Coelli's (1995) normal – truncated normal model. Mean cost-*inefficiency* decreases, that is, mean cost-*efficiency* increases, in the elements of  $\mathbf{z}_i$  whose coefficients are negative. The deterministic portion of (7),  $f(\ln C_i, \ln P_{i1}, \dots, \ln P_{in})$ , represents the household's minimum cost of purchasing calories, interpretable as that to which condition (5) pertains. The error term  $v_i$  may be considered to arise from mismeasurement of the dependent variable or other statistical noise. The error  $u_i$ , measuring the extent by which the household's cost of acquiring calories exceeds the minimum cost of doing so, must be non-negative since it is definitionally impossible to purchase calories at less than minimum cost.

Note that whereas cost-inefficiency in production is caused by such factors as mismanagement of the firm, cost-inefficiency in calorie purchases arises from the facts that (a) consumers derive utility from recreation and that (b) there is recreational, besides caloric, value to food. If consumers didn't value recreation, or there was no recreational value to food, they would simply minimize the cost of calories. As discussed, factors that raise the marginal utility from calories promote cost-efficiency, in bringing the consumer closer to satisfying (5), the condition for cost-minimization. Such factors include poverty, since there is diminishing marginal utility from calories, strenuous work, since it is calorie-intensive, and a diseases-ridden environment, since illness impedes the absorption of calories and immunological responses are calorie-intensive.

Common forms of (7) are the Cobb-Douglas,

$$\ln E_i = \alpha_0 + \alpha_C \ln C_i + \sum_{j=1}^n \alpha_j \ln P_{ij} + v_i + u_i, \quad (8)$$

and the translog,

$$\begin{aligned} \ln E_i = & \beta_0 + \beta_C \ln C_i + \sum_{j=1}^n \beta_j \ln P_{ij} + \beta_{CC} (\ln C_i)^2 + \sum_{j=1}^n \beta_{Cj} \ln C_i \ln P_{ij} + \\ & \sum_{j=1}^n \sum_{k=1}^n \beta_{jk} \ln P_{ij} \ln P_{ik} + v_i + u_i, \end{aligned} \quad (9)$$

stochastic frontier cost functions. The coefficients  $\alpha$  in (8) or  $\beta$  in (9), together with the coefficients  $\delta$  of the environmental variables, may be estimated by maximum likelihood<sup>5</sup>. Since the Cobb-Douglas cost function implies certain rigidities, such as constant demand elasticities and factor cost-shares, the translog function has been preferred in stochastic frontier analysis. However, since, as discussed below, the Cobb-Douglas stochastic frontier cost function is well suited to addressing a critical shortcoming of the data, this study employs, in turn, both the restrictive Cobb-Douglas and the flexible translog functions.

The study's view of calories as *metabolizable* calories, adopted to motivate a quasi-concave calories production function, must now confront the practical difficulty that metabolizable calories are unobserved. This compels their substitution in the empirical investigation by purchased calories. The result is mismeasurement of the regressor  $\ln C_i$  in (8) and (9). If  $C_i$  denotes household  $i$ 's unobservable metabolizable calories and  $C_{pi}$  its observed purchased calories, it must be that  $C_i = C_{pi} \times K_i$ ,  $0 < K_i \leq 1$ , since metabolizable calories are obtained from purchased calories. Hence,

$$\ln C_i = \ln C_{pi} + \ln K_i, \quad (10)$$

where the random variable  $\ln K_i$ , henceforth termed  $k_i$ , assumed independent of  $u_i$ , has a one-sided distribution since it is non-positive. Substituting (10) into, say, (8) yields

$$\ln E_i = \alpha_0 + \alpha_C \ln C_{pi} + \sum_{j=1}^n \alpha_j \ln P_{ij} + v_i + u_i + \alpha_C k_i.$$

---

<sup>5</sup> using, for example, STATA's *sfcross* routine.

The one-sided error term  $\alpha_C k_i$  is an impediment to estimation since it alters the customary error structure of the stochastic cost frontier model. This may be remedied as follows. Let  $\overline{\ln C}$  denote the population mean of  $\ln C_i$ . It follows from (10) that

$$\overline{\ln C} = \overline{\ln C_p} + \bar{k}, \quad (11)$$

where  $\overline{\ln C_p}$  signifies the population mean of  $\ln C_{pi}$ , and  $\bar{k}$  that of  $k_i$ . Mean-center the regressor  $\ln C_i$  in (8) to give

$$\ln E_i = (\alpha_0 + \alpha_C \overline{\ln C}) + \alpha_C (\ln C_i - \overline{\ln C}) + \sum_{j=1}^n \alpha_j \ln P_{ij} + v_i + u_i. \quad (12)$$

Substituting (10) and (11) into (12) yields

$$\begin{aligned} \ln E_i &= (\alpha_0 + \alpha_C \overline{\ln C_p} + \alpha_C \bar{k}) + \alpha_C (\ln C_{pi} - \overline{\ln C_p}) \\ &\quad + \sum_{j=1}^n \alpha_j \ln P_{ij} + \{v_i + \alpha_C (k_i - \bar{k})\} + u_i. \end{aligned} \quad (13)$$

Observe that the problematic  $\alpha_C k_i$  has thus been converted into the tractable  $\alpha_C (k_i - \bar{k})$ , which may be added to the error  $v_i$  to make up the model's normally distributed error term. Provided the sample were large, replacing the unknown population mean  $\overline{\ln C_p}$  by its sample counterpart would operationalize (13).

The standard translog cost function isn't amenable to this procedure for the reason that it contains the square of  $\ln C_i$ . The mean-centering of  $\ln C_i$  in (9) shall yield

$$\begin{aligned} \ln E_i &= A + B(\ln C_i - \overline{\ln C}) + \sum_{j=1}^n D_j \ln P_{ij} + \beta_{CC}(\ln C_i - \overline{\ln C})^2 + \\ &\quad \sum_{j=1}^n \beta_{Cj} \ln P_{ij} (\ln C_i - \overline{\ln C}) + \sum_{j=1}^n \sum_{k=1}^n \beta_{jk} \ln P_{ij} \ln P_{ik} + \\ &\quad v_i + u_i, \end{aligned} \quad (14)$$

where  $A$ ,  $B$ , and  $D_j$  are linear combinations of subsets of the coefficients in (9). Substituting (10) and (11) into (14) leads to

$$\begin{aligned} \ln E_i &= A + B(\ln C_{pi} - \overline{\ln C_p}) + \sum_{j=1}^n D_j \ln P_{ij} + \beta_{CC}(\ln C_{pi} - \overline{\ln C_p})^2 + \\ &\quad \sum_{j=1}^n \beta_{Cj} \ln P_{ij} (\ln C_{pi} - \overline{\ln C_p}) + \sum_{j=1}^n \sum_{k=1}^n \beta_{jk} \ln P_{ij} \ln P_{ik} + \end{aligned}$$

$$\begin{aligned} & \left[ v_i + \{B + 2\beta_{CC}(\ln C_{pi} - \overline{\ln C_p}) + \sum_{j=1}^n \beta_{Cj} \ln P_{ij}\}(k_i - \bar{k}) \right] +, \\ & u_i + \beta_{CC}(k_i - \bar{k})^2. \end{aligned} \quad (15)$$

Observe that the error in square brackets above, safely assumed to be normally distributed, is heteroskedastic. While this is easily accommodated in estimation, the term  $\beta_{CC}(k_i - \bar{k})^2$  alters the error structure of the stochastic cost frontier model, so that estimates of (15) that ignore it are suspect. Note that the square of log output is included in the single-output translog cost function so as to permit the elasticity of cost to output, that is, scale economies, to vary with output. Therefore, it might be omitted if it were assumed that scale economies were invariant in output. There is precedent for this. For example, Binswanger (1974), in an early exposition, derived, as a Taylor series expansion, a translog cost function from which the square of log output is omitted. If it were assumed that scale economies were invariant in output, (15) would be modified as

$$\begin{aligned} \ln E_i &= A + \beta_C(\ln C_{pi} - \overline{\ln C_p}) + \sum_{j=1}^n D_j \ln P_{ij} + \\ & \sum_{j=1}^n \beta_{Cj} \ln P_{ij} (\ln C_{pi} - \overline{\ln C_p}) + \sum_{j=1}^n \sum_{k=1}^n \beta_{jk} \ln P_{ij} \ln P_{ik} + \\ & \left[ v_i + \{\beta_C + \sum_{j=1}^n \beta_{Cj} \ln P_{ij}\}(k_i - \bar{k}) \right] + u_i. \end{aligned} \quad (16)$$

Again, the error in square brackets above, clearly heteroskedastic, may safely be assumed to be normally distributed. Observe that the error structure of the stochastic cost frontier model is preserved, so that consistent estimation of (16),  $\overline{\ln C_p}$  replaced by its large sample counterpart, is unhindered.

The National Sample Survey Organization (NSSO) of the Govt. of India conducts the larger of its household consumer expenditure survey at quinquennial intervals. This study utilizes consumer expenditure data from the 38th (1983) and 61st (2004-05) quinquennial rounds of the National Sample Survey (NSS). The period 1983 to 2004-05, spanning about six-tenths of the era of secular decline in per capita calorie consumption, was analyzed as well by Deaton and Dreze (2009). Hence, a focus upon it, though there have been quinquennial rounds since 2004-05, will facilitate comparison with Deaton and Dreze's (2009) widely cited findings. It is notable that, since 1983 lies

well within the era predating the economic reforms of 1991, and 2004-05 falls well within the post-reforms era, the period between the 38th and 61st rounds of the NSS witnessed rapid economic transformation.

The surveys elicited the quantities of the various food items purchased by a household within the past 30 days, or consumed out of home-produced stocks, as well as their monetary values. The sum of these reported values was considered the household's expenditure upon calories,  $E_i$ . Quantities of individual food items were converted into calories by the means of conversion factors supplied by Gopalan et al. (1974). The sum, across food items, of these calories was considered the household's purchased calories.

The price paid by a household for a food item is calculated as expenditure upon the item, which includes the reported value of consumption out of home-grown stocks, divided by the number of units obtained of the item. This household-specific price is potentially endogenous since, for example, households keen to acquire calories cheaply may bargain-hunt. Hence, the exogenous price facing a household is best calculated as the community-wide mean of household-specific prices. To reduce the number of prices in (7) in the interest of empirical tractability, food items were grouped into categories, such as 'eggs, meat, and fish', 'fresh fruit', and 'vegetables'. The household-specific price applicable to a category was calculated as the weighted geometric mean of the prices paid by the household for the food items within that category, the weight assigned an item being its share of expenditure upon the category. In other words, the household-specific price assigned a food category is its Stone price index. A community-wide geometric mean of these household-specific Stone price indices for a food category was considered the category's exogenous price facing the community's households. A household's community was, for the most part, the First-Stage Sampling Unit (FSU), that is, rural village or urban block, in which it resided<sup>6</sup>. Community-wide nominal prices in rural

---

<sup>6</sup> Households reside in FSUs, located within Strata, which are, roughly, districts, lying within Regions, hierarchical domains below the level of State or Union Territory. If the community-wide geometric mean of household Stone price indices returned a missing value for the reason that the particular food category wasn't locally consumed, a household's community was taken to be the rural or urban, depending on whether its residence was rural or urban, portion of the hierarchical domain above.

areas were converted to real prices (1999-2000 = 100) using the consumer price index for agricultural laborers (CPIAL), whereas the CPI for industrial workers (CPIIU) was applied to nominal prices in urban areas. Since vegetarianism in India is overwhelmingly by religious dictate, the price of eggs, meat, and fish ought not to play a role in vegetarian households' purchases of food. Hence, the log of the price of eggs, meat, and fish ought not to be an argument of the deterministic Caloric Cost Frontier of vegetarian households. Therefore, it is interacted with an indicator of non-vegetarianism before inclusion in the function.

The vector of determinants of the mean of the distribution of cost-inefficiency,  $\mathbf{z}_i$ , is taken to include a time dummy variable, to capture secular change over time. Since the disease environment and occupational strenuousness may be factors in the cost-efficiency of caloric purchases,  $\mathbf{z}_i$  also includes measures of the perniciousness of the disease environment and the strenuousness of the household's principal occupation. Following Duh and Spears (2017), the disease environment is gauged by the local infant mortality rate, the annual number of deaths before age 1 per 1000 live births. While the Census of India supplies data on infant mortality rates at the district level, districts are unidentifiable in the 38<sup>th</sup> (1983) round of the NSS. This compels the use of state-level rural and urban infant mortality data to gauge the perniciousness of a household's disease environment. Next, following Siddiqui, Donato, and Jumrani (2019), household principal occupations requiring vigorous activity were identified, and duly indicated by a dummy variable as strenuous. Further, households which purchased no calories, or calories per capita in excess of 10,000 per day, were dropped for fear of being outliers. This is akin to the dropping, by Siddiqui, Donato, and Jumrani (2019), of households in which daily per capita calorie consumption was either zero or in excess of 10,000.

Table 1 presents the weighted, by sampling weights, sample means of a subset of the variables employed in the analyses, the arguments of the concerned translog frontier cost function being too numerous for concise presentation. It is indicated that mean purchased kilocalories per month was greater in the 38<sup>th</sup> (1983) than in the 61<sup>st</sup> (2004-05) round of the NSS, in keeping with the Calorie Consumption Puzzle. It is notable that mean household real monthly expenditure per



capita was 53% higher in 2004-05 than in 1983, which attests to the rapid growth of the Indian economy since its liberalization. It is also indicated that the nature of work substantially changed over this period, the fraction of households whose principal occupation was strenuous being distinctly lower in 2004-05. Table 1 also reveals that the disease environment, as measured by the infant mortality rate, was much improved over this period.

#### 4. Results

Since the focus of this study is time variation in the cost-efficiency of calorie purchases, the following discussion dwells on the estimated coefficients of the variables determining the mean of the distribution of cost-inefficiency. Table 2 presents estimates of the Cobb-Douglas model described by (13), in which the sample mean of log purchased calories substitutes its population counterpart. It accounts for the role of the strenuousness of occupations and the perniciousness of the disease environment in the cost-efficiency of calorie purchases. Estimation incorporates NSS sampling weights<sup>7</sup>. It is found that mean cost-inefficiency is lower in households whose primary occupation was strenuous, though this was less pronounced by 2004-05. As argued, the higher caloric needs of workers in strenuous occupations may be considered to raise their marginal utility from calories, potentially compelling less cost-inefficient (more cost-efficient) calorie purchases. That this was less pronounced by 2004-05 is consistent with strenuous occupations becoming less calorie-intensive between the 38th and 61st rounds of the NSS. For example, Deaton and Dreze (2009) observe that the mechanization of agricultural work has contributed to reduced calorie requirements.

Since illness both inhibits the absorption of nutrients and induces immunological responses that are calorie-intensive, the resulting higher calorie requirements of the habitants of areas more rife with diseases, that is, their greater marginal utility from calories, may turn their calorie purchases

---

<sup>7</sup> As a practical matter, the constant term was dropped from the linear expression for the mean of the distribution of cost-inefficiency, as its inclusion when sampling weights were incorporated in estimation prevented convergence of the ML estimator.

more cost-efficient. This too appears borne out, in that residents of states with higher infant mortality rates are less cost-inefficient in their calorie purchases on average.

It is notable that secular decline between 1983 and 2004-05 in the mean of the distribution of cost-inefficiency, as indicated by the negative coefficient of *61<sup>st</sup> Round of the NSS (2004-05)*, is robust to controlling for the strenuousness of occupations and the perniciousness of the disease environment. Owing to it, households whose principal occupation wasn't strenuous, residing in a state whose infant mortality rate in 1983 equaled the sample mean value of the variable that year, 115.854, and whose infant mortality rate in 2004-05 equaled the variable's improved sample mean value by that year, 66.187, would, *ceteris paribus*, have brought down the mean cost-inefficiency of their calorie purchases (since  $-0.128 - 0.0004 \times (66.187 - 115.854) < 0$ )<sup>8</sup>. Further, their mean cost-inefficiency would have decreased even had their principal occupation been strenuous in both periods (since  $-0.128 + 0.033 - 0.0004 \times (66.187 - 115.854) < 0$ )<sup>8</sup>, or strenuous in 1983 but not so by 2004-05 (since  $-0.128 - (-0.040 + 0.033) - 0.0004 \times (66.187 - 115.854) < 0$ )<sup>8</sup>. In sum, there seems to have been a *net* squeeze in the food budget in the period 1983-2005, income held constant, despite reduction in the calorie intensity of strenuous work, reduced participation in strenuous work, and an improved disease environment.

Might widening of the menu of services and non-food goods available to consumers, leading to rise in non-food spending at the expense of the food budget, have been partly responsible? Note that, since it is endogenous, household non-food expenditure may not simply be considered a determinant of the mean of the distribution of cost-inefficiency, for if it were, its endogeneity would lead to correlation between the error terms  $v_i$  and  $u_i$  in (7), a violation of the assumptions of the stochastic frontier model. After all, by (7),  $v_i$  affects  $E_i$ , expenditure on food, which, given the household budget constraint, affects non-food expenditure. Would it be possible to consider non-food expenditure an environmental variable if its endogeneity were addressed by instrumental variables methods such as that developed by Karakaplan and Kutlu (2017)? An identifying

---

<sup>8</sup> The null hypothesis of this linear combination of coefficients equaling zero is rejected at the 1% level.

instrument must be an exogenous determinant of non-food expenditure excluded from the model. Given the household budget constraint, such a variable would be impossible to find. Since expenditure on food (calories) is simply household income less non-food expenditure, any determinant of non-food expenditure will automatically be a factor in expenditure on food, that is, either an argument of the deterministic frontier cost function or a determinant of the mean of the distribution of cost-inefficiency. In sum, it is impossible to consider household non-food expenditure an environmental variable. In any case, if the ultimate reason for the squeeze in the food budget is the development-led widening of the menu of services and non-food goods available to consumers, an exogenous measure of the local availability of services and non-food goods is called for.

Since it is plausible that this widening was more rapid in urban than in rural India, the squeeze in the food budget may have been more acute in urban areas. The estimates in Table 3 indicate that decline in cost-inefficiency was indeed more pronounced in urban India, in that the estimated coefficient of *urban*×61<sup>st</sup> Round is negative and the variable significant. Since the community-wide mean of per capita non-food expenditure was likely higher the greater was local access to services and non-food goods, it may be considered a measure of this access. The estimates in Table 3 indicate that cost-inefficiency in the purchase of calories was indeed lower, *ceteris paribus*, in communities in which there was greater expenditure per capita on services and non-food goods. In sum, the estimates support a connection between increased local access to services and non-food goods and squeezed food budgets in the period 1983-2005.

In addition, they indicate that urban households, at least in 1983, and those headed by the literate purchased calories more cost-inefficiently on average, while the opposite was true of households headed by women and older individuals. Interestingly, mean cost-inefficiency increased in household size. This is consistent with Deaton and Paxon's (1998) argument that calorie requirements are lower in larger households, there being more members to shoulder fixed chores. As expected, mean cost-inefficiency increased in household per capita expenditure, a proxy for income, since a rise in income serves to loosen food budgets. It may be observed as well that expenditure

elasticities, the coefficients of log prices, were positive for all food items save beverages and oil. Since the price elasticity of demand for a good equals its expenditure elasticity less 1, it appears that demand for each food item, save beverages and oil, was price-inelastic. It is not surprising that price-inelasticity was most pronounced in the case of cereals, Indians' dietary staple.

Columns 1 and 2 of Table 4 presents estimates pertaining to the translog stochastic frontier cost function (15). These ought to be viewed with caution since estimation ignores the error  $\beta_{CC}(k_i - \bar{k})^2$ . Note that log prices are mean-centered as well to permit interpretation of their coefficients as expenditure elasticities at the sample geometric means of prices. By these estimates as well, flawed as they are, there seems to have been a net squeeze in the food budget between 1983 and 2005, contributed to by widening of the menu of services and non-food goods available to consumers. Column 3 and 4 of Table 4 presents estimates, these reliable, concerning (16), a translog frontier cost function in which scale economies are invariant in output. They too corroborate a net squeeze in the food budget, or declining food security, in which the greater availability of services and non-food goods played a part.

## 5. Conclusion

Stochastic cost frontier analyses of Indian households' purchases of calories indicates that these became less cost-inefficient (more cost-efficient) on average between the 38<sup>th</sup> (1983) and 61<sup>st</sup> (2004-05) rounds of the NSS, this despite reduction in the calorie intensity of strenuous work, reduced participation in strenuous work, and an improved disease environment. As argued, this is consistent with a net squeeze in the household food budget. That this squeeze appears more pronounced in localities in which per capita non-food expenditure was higher, suggests that it was at least partly associated with greater consumer access to services and non-food goods. This evidence of a net squeeze in the food budget suggests that it is a viable explanation of India's Calorie Consumption Puzzle.

Besides attempting an analysis of the Calorie Consumption Puzzle, this study demonstrates a novel application of Stochastic Cost Frontier Analysis, to consumption instead of the more common production. Stochastic Cost Frontier Analysis has not previously been applied to the consumption of multiple items treated as inputs yielding an output. Besides, the application isn't merely novel. It makes for an innovative tool in the diagnosis of changes over time in food security, in that increased cost-efficiency in calorie purchases may be indicative of a squeeze in the food budget or declining food security.

As defined by the UN's Committee on World Food Security, food security is achieved when "all people, at all times, have physical, social, and economic access to sufficient, safe, and nutritious food that meets their food preferences and dietary needs for an active and healthy life". By this definition, food must be nutritious as well as meet food preferences to assure food security. Therefore, it might be argued that the more economical purchase of only calories obtained from nutritious and preferred food items, rather than calories as a whole, be considered symptomatic of declining food security. While this is a substantive argument, identification of such food items wouldn't be straightforward, nutrition being so multi-dimensional that a food item rich in a dimension of nutrition may be poor in others, and food preferences exceedingly heterogeneous in nations as culturally diverse as India. Note, however, that Indians' diets are mostly traditional, which suggests that the food items consumed are both to regional tastes, and, having nurtured millennia of generations, nutritious. In sum, considering the more cost-efficient purchase of calories as a whole to be indicative of declining food security in a traditional society such as India, is defensible.

Table 1  
Sample Means (Weighted)  
Subset of Variables

| Variable   | Mean               | S.D.    | Mean                                  | S.D.     | Mean                                     | S.D.     |
|--|--------------------|---------|---------------------------------------|----------|--|----------|
|  | <i>full sample</i> |         | <i>38<sup>th</sup> Round<br/>1983</i> |          | <i>61<sup>st</sup> Round<br/>2004-05</i> |          |
| <b>Dependent</b>   |                    |         |                                       |          |  |          |
| monthly real expenditure on food (1999-2000 Rs.)                       | 1762.725           | 4.225   | 1771.863                              | 6.334    | 1756.941                                 | 5.617    |
| <b>Frontier</b>  |                    |         |                                       |          |  |          |
| kilocalories purchased last 30 days                                    | 369,447.7          | 876.904 | 410,202.6                             | 1407.315 | 343,647.1                                | 1119.399 |
| price of beverages (real Rs./liter)                                    | 9.700              | 0.040   | 2.219                                 | 0.022    | 14.436                                   | 0.055    |
| price of spices (real Rs./gram)  | 0.062              | 0.0001  | 0.049                                 | 0.0001   | 0.069                                    | 0.0001   |
| price of salt (real Rs./kg)  | 3.341              | 0.006   | 2.021                                 | 0.003    | 4.177                                    | 0.009    |
| price of sugar (real Rs./kg)   | 15.760             | 0.010   | 14.144                                | 0.012    | 16.783                                   | 0.014    |
| price of dried fruit (real Rs./kg)                                     | 68.383             | 0.434   | 59.687                                | 0.344    | 73.888                                   | 0.673    |
| price of fresh fruit (real Rs./kg)                                     | 6.843              | 0.020   | 5.063                                 | 0.020    | 7.969                                    | 0.029    |
| price of vegetables (real Rs./kg)                                      | 7.871              | 0.009   | 7.189                                 | 0.011    | 8.303                                    | 0.012    |
| price of eggs, meat, and fish (real Rs./kg)                            | 46.885             | 0.082   | 41.329                                | 0.073    | 50.403                                   | 0.122    |
| price of edible oils (real Rs./kg)                                     | 53.241             | 0.029   | 58.337                                | 0.035    | 50.015                                   | 0.036    |
| price of milk (real Rs./liter)   | 18.196             | 0.067   | 19.843                                | 0.152    | 17.152                                   | 0.050    |
| price of pulses (real Rs./kg)  | 22.787             | 0.015   | 18.788                                | 0.017    | 25.319                                   | 0.017    |
| price of cereal substitutes (real Rs./kg)                              | 15.155             | 0.033   | 12.914                                | 0.044    | 16.574                                   | 0.046    |
| price of cereals (real Rs./kg)   | 9.183              | 0.008   | 9.731                                 | 0.011    | 8.837                                    | 0.010    |
| <b>Inefficiency</b>  |                    |         |                                       |          |  |          |
| urban  | 0.228              | 0.001   | 0.181                                 | 0.001    | 0.257                                    | 0.002    |
| household head is literate   | 0.563              | 0.002   | 0.483                                 | 0.002    | 0.614                                    | 0.002    |
| household head is female   | 0.074              | 0.001   | 0.067                                 | 0.001    | 0.078                                    | 0.001    |
| age (years) of household head  | 45.573             | 0.045   | 44.772                                | 0.058    | 46.080                                   | 0.063    |
| household size (no. of members)  | 6.182              | 0.012   | 6.507                                 | 0.017    | 5.977                                    | 0.017    |
| monthly real expenditure per capita                                    | 571.638            | 1.610   | 431.388                               | 1.078    | 660.425                                  | 2.508    |
| strenuous occupation   | 0.681              | 0.002   | 0.744                                 | 0.002    | 0.641                                    | 0.002    |
| state infant mortality rate (deaths before age 1 per 1000 live births) | 85.441             | 0.101   | 115.854                               | 0.128    | 66.187                                   | 0.082    |
| FSU mean monthly real non-food expenditure per capita                  | 306.96             | 0.917   | 163.341                               | 0.481    | 397.88                                   | 1.361    |
| N  | 220,476            |         | 101,920                               |          | 118,556                                  |          |

weights = NSS sampling weights; variables appearing in logs in the analysis expressed in levels

Table 2  
The Strenuousness of Work, the Disease Environment, and the Cost-Efficiency of Calorie Purchases  
Cobb-Douglas Stochastic Cost Frontier

| Variable   | Coeff.                  | S.E.           |
|--|-------------------------|----------------|
| <i>Subset of Estimates of Cost Frontier: dependent variable = log monthly real expenditure on food</i> |                         |                |
| constant   | -5.216***               | 0.045          |
| log kilocalories (mean-centered)   | 0.657***                | 0.004          |
| log price of beverages   | -0.001***               | 0.0004         |
| log price of spices  | 0.017***                | 0.002          |
| log price of salt  | 0.033***                | 0.002          |
| log price of sugar   | 0.074***                | 0.003          |
| log price of dried fruit   | 0.004***                | 0.001          |
| log price of fresh fruit   | 0.002***                | 0.001          |
| log price of vegetables  | 0.074***                | 0.002          |
| log price of eggs, meat, and fish × nonvegetarian  | 0.008***                | 0.0003         |
| log price of edible oils   | 0.002                   | 0.005          |
| log price of milk  | 0.028***                | 0.001          |
| log price of pulses  | 0.098***                | 0.003          |
| log price of cereal substitutes  | 0.009***                | 0.001          |
| log price cereals  | 0.288***                | 0.003          |
| <i>Determinants of the mean of the distribution of cost-inefficiency</i>                               |                         |                |
| <b>61<sup>st</sup> Round of the NSS (2004-05)</b>  | <b>-0.128***</b>        | <b>0.003</b>   |
| urban  | -0.012***               | 0.002          |
| household head is literate   | 0.022***                | 0.001          |
| household head is female   | -0.029***               | 0.002          |
| age (years) of household head  | -0.0004***              | 0.00005        |
| household size   | 0.053***                | 0.001          |
| log monthly real expenditure per capita  | 0.444***                | 0.002          |
| <b>strenuous occupation</b>  | <b>-0.040***</b>        | <b>0.002</b>   |
| <b>strenuous occupation × 61<sup>st</sup> Round</b>  | <b>0.033***</b>         | <b>0.002</b>   |
| <b>state infant mortality rate</b>   | <b>-0.0004***</b>       | <b>0.00003</b> |
| log-likelihood   | 6.531 × 10 <sup>8</sup> |                |
| N  | 220,476                 |                |

\*, \*\*, and \*\*\* denote, respectively, significance at the 10%, 5%, and 1% levels

Table 3  
Widening of the Menu of Services and Non-Food Goods and the Cost-Efficiency of Calorie Purchases  
Cobb-Douglas Stochastic Cost Frontier

| Variable   | Coeff.                  | S.E.         |
|--|-------------------------|--------------|
| <i>Estimates of Cost Frontier: dependent variable = log monthly real expenditure on food</i> |                         |              |
| constant   | -5.060***               | 0.046        |
| log kilocalories (mean-centered)   | 0.651***                | 0.004        |
| log price of beverages   | -0.001                  | 0.0004       |
| log price of spices  | 0.019***                | 0.002        |
| log price of salt  | 0.038***                | 0.002        |
| log price of sugar   | 0.064***                | 0.003        |
| log price of dried fruit   | 0.006***                | 0.001        |
| log price of fresh fruit   | 0.005***                | 0.001        |
| log price of vegetables  | 0.090***                | 0.002        |
| log price of eggs, meat, and fish × nonvegetarian  | 0.008***                | 0.0003       |
| log price of edible oils   | -0.007                  | 0.005        |
| log price of milk  | 0.032***                | 0.001        |
| log price of pulses  | 0.110***                | 0.003        |
| log price of cereal substitutes  | 0.010***                | 0.001        |
| log price cereals  | 0.285***                | 0.003        |
| <br><i>Determinants of the mean of the distribution of cost-inefficiency</i>                 |                         |              |
| 61 <sup>st</sup> Round of the NSS (2004-05)  | -0.100***               | 0.003        |
| urban  | 0.015***                | 0.002        |
| <b>urban × 61<sup>st</sup> Round</b>   | <b>-0.015***</b>        | <b>0.003</b> |
| household head is literate   | 0.022***                | 0.001        |
| household head is female   | -0.028***               | 0.002        |
| age (years) of household head  | -0.0004***              | 0.00004      |
| household size   | 0.054***                | 0.001        |
| log monthly real expenditure per capita  | 0.470***                | 0.003        |
| strenuous occupation   | -0.037***               | 0.002        |
| strenuous occupation × 61 <sup>st</sup> Round  | 0.026***                | 0.002        |
| state infant mortality rate  | -0.0005***              | 0.00003      |
| <b>log FSU mean monthly real non-food expenditure per capita</b>                             | <b>-0.054***</b>        | <b>0.001</b> |
| log-likelihood   | 6.650 × 10 <sup>8</sup> |              |
| N  | 220,476                 |              |

\*, \*\*, and \*\*\* denote, respectively, significance at the 10%, 5%, and 1% levels



Table 4  
Translog Stochastic Cost Frontier

| Variable   | (1)<br>Coeff. | (2)<br>S.E.             | (3)<br>Coeff. | (4)<br>S.E.             |
|--|---------------|-------------------------|---------------|-------------------------|
| <i>Subset of Estimates of Cost Frontier: dependent variable = log monthly real expenditure on food</i> |               |                         |               |                         |
| constant   | -5.131***     | 0.269                   | -3.812***     | 0.038                   |
| log kilocalories (mean-centered)   | 0.882***      | 0.044                   | 0.667***      | 0.004                   |
| log kilocalories (mean-centered) squared   | -0.010***     | 0.002                   |               |                         |
| log price of beverages (mean-centered)   | 0.002         | 0.014                   | 0.016         | 0.013                   |
| log price of spices (mean-centered)  | -0.123        | 0.094                   | -0.045        | 0.074                   |
| log price of salt (mean-centered)  | -0.255***     | 0.053                   | -0.148***     | 0.054                   |
| log price of sugar (mean-centered)   | 0.100         | 0.104                   | 0.203*        | 0.110                   |
| log price of dried fruit (mean-centered)   | -0.007        | 0.024                   | -0.076***     | 0.027                   |
| log price of fresh fruit (mean-centered)   | -0.007        | 0.021                   | 0.018***      | 0.021                   |
| log price of vegetables (mean-centered)  | -0.016        | 0.076                   | -0.113        | 0.069                   |
| log price of eggs, meat, and fish (mean-centered) × nonvegetarian                                      | 0.200***      | 0.025                   | 0.195***      | 0.023                   |
| log price of edible oils (mean-centered)   | 0.571***      | 0.158                   | 0.687***      | 0.151                   |
| log price of milk (mean-centered)  | 0.077**       | 0.036                   | 0.058*        | 0.034                   |
| log price of pulses (mean-centered)  | -0.179*       | 0.103                   | -0.348***     | 0.097                   |
| log price of cereal substitutes (mean-centered)  | 0.106***      | 0.026                   | 0.089***      | 0.025                   |
| log price cereals (mean-centered)  | 0.419***      | 0.088                   | 0.232***      | 0.079                   |
| <i>Determinants of the mean of the distribution of cost-inefficiency</i>                               |               |                         |               |                         |
| 61 <sup>st</sup> Round of the NSS (2004-05)  | -0.092***     | 0.003                   | -0.089***     | 0.004                   |
| urban  | 0.006**       | 0.002                   | 0.008***      | 0.003                   |
| urban × 61 <sup>st</sup> Round   | -0.003        | 0.003                   | 0.00005       | 0.003                   |
| household head is literate   | 0.020***      | 0.001                   | 0.022***      | 0.001                   |
| household head is female   | -0.024***     | 0.002                   | -0.028***     | 0.002                   |
| age (years) of household head  | -0.0005***    | 0.00004                 | -0.0004***    | 0.00004                 |
| household size   | 0.063***      | 0.001                   | 0.053***      | 0.001                   |
| log monthly real expenditure per capita  | 0.487***      | 0.003                   | 0.465***      | 0.002                   |
| strenuous occupation   | -0.030***     | 0.002                   | -0.032***     | 0.002                   |
| strenuous occupation × 61 <sup>st</sup> Round  | 0.019***      | 0.002                   | 0.018***      | 0.002                   |
| state infant mortality rate  | -0.0004***    | 0.00003                 | -0.0004***    | 0.00003                 |
| log FSU mean monthly real non-food expenditure per capita  | -0.059***     | 0.001                   | -0.054***     | 0.001                   |
| log-likelihood   |               | 7.290 × 10 <sup>8</sup> |               | 7.156 × 10 <sup>8</sup> |
| N  |               |                         |               | 220,476                 |

\*, \*\*, and \*\*\* denote, respectively, significance at the 10%, 5%, and 1% levels; estimates in cols. 3 & 4 pertain to a model in which scale economies are invariant in calories

## References

- Banerjee, A. V. & Duflo, E. (2007). The Economic Lives of the Poor. *Journal of Economic Perspectives*, 21(1), pp. 141-168.
- Battese, G. & Coelli, T. (1995). A Model for Technical Efficiency Effects in a Stochastic Frontier Production Model for Panel Data. *Empirical Economics*, 20(2), pp. 325-332.
- Binswanger, H. P. (1974). A Cost Function Approach to the Measurement of Elasticities of Factor Demand and Elasticities of Substitution. *American Journal of Agricultural Economics*, 56(2), pp. 377-386.
- Chandrasekhar, C. P. & Ghosh, J. (2003). The Calorie Consumption Puzzle. *The Hindu Business Line*, 11, Retrieved from:  
<http://www.thehindubusinessline.in/2003/02/11/stories/2003021100210900.htm>.
- Deaton, A. & Dreze, J. (2009). Nutrition in India: Facts and Interpretations. *Economic and Political Weekly*, 44(7), pp. 42-65.
- Deaton, A. & Paxon, C. (1998). Economies of Scale, Household Size, and the Demand for Food. *Journal of Political Economy*, 106(5), pp. 897-930.
- Duh, J. & Spears, D. (2017). Health and Hunger: Disease, Energy Needs, and the Indian Calorie Consumption Puzzle. *The Economic Journal*, 127(606), pp. 2378-2409.
- Eli, S. & Li, N. (2015). Calorie Requirements and Food Consumption Patterns of the Poor. NBER Working Paper 21697, Cambridge: National Bureau of Economic Research.
- Filippini, M. & Hunt, L. C. (2012). US Residential Energy Demand and Energy Efficiency: A Stochastic Demand Frontier Approach. *Energy Economics*, 34(5), pp. 1484-1491.
- Gopalan, C., Sastri, B. R., & Balasubramanian, S. (1974). *Nutritive Value of Indian Foods*. National Inst. of Nutrition, Indian Council of Medical Research, Hyderabad, India.
- Karakaplan, M. U. & Kutlu, L., (2017). Handling Endogeneity in Stochastic Frontier Analysis. *Economics Bulletin*, 37(2), pp. 889-901.
- Lee, J., Park, C., Oh, D., & Kim, T. (2008). Measuring Consumption Efficiency with Utility Theory

- and Stochastic Frontier Analysis. *Applied Economics*, 40(22), pp. 2961-2968.
- Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, 341, pp. 976-980.
- Mehta, J. & Venkatraman, S. (2000). Poverty Statistics: Bermicide's Feast. *Economic and Political Weekly*, 35(27), pp. 2377-82.
- NSSO. (2014). Nutritional Intake in India. Report Number 560. NSS 68<sup>st</sup> Round, July 2011–June 2012. National Statistical Office, Ministry of Statistics & Programme Implementation, Government of India, October.
- Parker, I. (2010). The Poverty Lab. *The New Yorker*, May 12 2010 Issue.
- Patnaik, U. (2004). The Republic of Hunger. *Social Scientist*, 32(9/10), pp. 9–35.
- Patnaik, U. (2007). Neoliberalism and Rural Poverty in India. *Economic and Political Weekly*, 42(30), pp. 3132–3150.
- Sen, P. (2005). Of Calories and Things: Reflections on Nutritional Norms, Poverty Lines and Consumption Behaviour in India. *Economic & Political Weekly*, 40(43), pp. 4611-8.
- Siddiqui, M, Z., Donato, R., & Jumrani, J. (2019). Looking Past the Indian Calorie Debate: What is Happening to Nutrition Transition in India. *The Journal of Development Studies*, 55(11), pp. 2440-2459.

### **Compliance with Ethical Standards**

We, the authors, solemnly state that:

- this research was not support by any funding source, and that there are no potential conflicts of interest, and that
- it did not involve human or animal participants.