Yale University
EliScholar – A Digital Platform for Scholarly Publishing at Yale

Discussion Papers | Economic Growth Center

6-1-2013

The Caloric Costs of Culture: Evidence from Indian Migrants

David Atkin

Follow this and additional works at: https://elischolar.library.yale.edu/egcenter-discussion-paper-series

Recommended Citation

Atkin, David, "The Caloric Costs of Culture: Evidence from Indian Migrants" (2013). *Discussion Papers*. 1036.
https://elischolar.library.yale.edu/egcenter-discussion-paper-series/1036

This Discussion Paper is brought to you for free and open access by the Economic Growth Center at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

DEPARTMENT OF ECONOMICS
YALE UNIVERSITY
P.O. Box 208268
New Haven, CT 06520-8268

http://www.econ.yale.edu/

Economic Growth Center Discussion Paper No. 1028

Economics Department Working Paper No. 119

The Caloric Costs of Culture: Evidence from Indian Migrants

David Atkin
Yale University

June 2013

Notes: Center discussion papers are preliminary materials circulated to stimulate discussion and critical comments.

This paper can be downloaded without charge from the Social Science Research Network Electronic Paper Collection: http://ssrn.com/abstract=2306113.

The Caloric Costs of Culture: Evidence from Indian Migrants*

David Atkin†

June 2013

Abstract

Anthropologists have long documented substantial and persistent differences across social groups in the preferences and taboos for particular foods. One natural question to ask is whether such *food cultures* matter in an economic sense. In particular, can culture constrain caloric intake and contribute to malnutrition? To answer this question, I first document that inter-state migrants within India consume fewer calories per Rupee of food expenditure compared to their non-migrant neighbors, even for households with very low caloric intake. I then form a chain of evidence in support of an explanation based on culture: that migrants make nutritionally-suboptimal food choices due to cultural preferences for the traditional foods of their origin states. First, I focus on the preferences themselves and document that migrants bring their origin-state food preferences with them when they migrate. Second, I link together the findings on caloric intake and preferences by showing that the gap in caloric intake between locals and migrants is related to the suitability and intensity of the migrants' origin-state food preferences: the most adversely affected migrants (households in which both husband and wife migrated to a village where their origin-state preferences are unsuited to the local price vector) would consume 7 percent more calories if they possessed the same preferences as their neighbors.

JEL CODES: I10, O10, Z10, D12.

*Thanks to Keith Chen, Angus Deaton, Adrian de Froment, Penny Goldberg, Gene Grossman, Dan Keniston, Nathan Nunn, Nancy Qian, Gabriella Santangelo, Chris Udry and Jacqueline Yen. I also thank participants at numerous seminars and the NBER Economics of Culture and Institutions workshop for excellent comments.

†Department of Economics, Yale University, 27 Hillhouse Ave., CT 06511. E-mail: david.atkin@yale.edu

1 Introduction

Anthropologists have long documented substantial and persistent differences across social groups in the preferences and taboos for particular foods. For example, Harris (1985) analyzes the historic origin of the taboo on beef consumption that persists among Hindus today, while Prakash (1961) notes that the relative preference for wheat in Northwest India and rice in East India dates back to the first millennium A.D.. One natural question to ask is whether such *food cultures* matter in an economic sense.[1] In particular, can culture constrain caloric intake and contribute to malnutrition? Such a question is of interest both for understanding the value that households place on their culture, and for designing effective policies to improve nutrition.

A stark example of the willingness of households to trade off cultural food preferences for nutrition, and an ineffective policy that did not take such preferences into account, comes from the report of the Famine Inquiry Commission. The commission was established in the aftermath of the 1943 Bengal Famine in which between 1.5 and 4 million Bengalis died. The final chapter of the commission's report centered on the role of regional preferences in exacerbating the famine:

> During the famine large supplies of wheat and millets were sent to Bengal ... but efforts to persuade people to eat them in their homes in place of rice met with little success ... we visited numerous grain stores in which quantities of wheat were deteriorating for lack of demand. ... The problem of how to wean rice-eaters from their determined preference from a food in short supply and reluctance to turn to alternative grains is not peculiar to Bengal, but is of all-India importance. (Famine Inquiry Commission, 1945)

The goal of this paper is to understand whether culture can constrain caloric intake and contribute to malnutrition. In order to do so, I require a setting where people are sufficiently hungry that reductions in caloric intake can have health, and hence economic, repercussions.[2] Accordingly, I focus on India, where I observe many households on the edge of malnutrition. This setting allows me to investigate whether culture can constrain caloric intake by observing the number of calories hungry households forgo in order to accommodate their food culture. My analysis compares the consumption patterns of interstate migrants with those of their non-migrant neighbors who face the same prices but possess different cultural food preferences. This methodology al-

[1]Since the types of food that a group of people traditionally consumes embody the preferences, beliefs and social attitudes of the group, I will describe differences in food preferences across groups as different *food cultures*. This definition accords the existing definitions of culture in the economics literature. For example, Fernández (2011) "consider differences in culture as systematic variation in beliefs and preferences across time, space or social groups", and Guiso, Sapienza and Zingales (2006) define culture as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation". Given the strong persistence in food preferences across generations described in Atkin (Forthcoming), both definitions of culture are applicable in my context.

[2]As low caloric intake directly reduces productivity, health, and immunity to diseases (e.g. Fogel, 1994), the economic consequences are obvious. From a welfare perspective, households may be optimally trading-off nutrition and cultural practices. Alternatively, since many of the gains from proper nutrition arise later in life (e.g. Almond and Currie, 2010), uninformed parents may undervalue such gains when making food choices for their children. In this scenario, culture can constrain both nutrition and welfare.

lows me to broadly quantify the "costs" that culture can impose. To the best of my knowledge this paper is the first to attempt such a quantification exercise.

To carry out such an analysis, I require information on household food consumption coupled with the migrant status of household members. The 1983 and 1987-88 Indian National Sample Surveys prove ideal for this purpose as 250,000 households were asked about their migration particulars and their purchases of 169 different food products.

My analysis proceeds in four stages. In the first stage, I provide the background to my study. I focus on India during a period when childhood malnutrition rates were above 50 percent and 64 percent of households consumed fewer calories than the nutritional adequacy requirements used to determine India's poverty line. If undernourished households are not maximizing nutrition in this context, it is important to understand why. The example of rice and wheat provides suggestive evidence that culture constrains nutrition in this context. Despite these two cereals providing a similar number of both calories and micro-nutrients per Rupee, there is dramatic regional variation in rice and wheat consumption. For example, the relative price of rice and wheat is similar in the states of Kerala and Punjab, yet Keralans consumed thirteen times more rice than wheat and Punjabis ten times more wheat than rice. As discussed in Atkin (Forthcoming), agro-climatic endowments coupled with habit formation can explain how these different food cultures developed. In terms of the costs of these cultural preferences, a crude counterfactual shows that mean Indian caloric intake could be 6.1 percent higher if households in some rice-loving states switched the quantity of rice and wheat that they were consuming (and vice versa for wheat-loving states), and spent any cost savings on the cheaper of the two grains.

Although suggestive, such an approach may be misleading if foods which are strong complements with rice or wheat have different prices across India. The behavior of inter-state migrants provides more compelling evidence that culture can constrain caloric intake. The key observation for my empirical strategy is that migrants face the same prices as their neighbors but are likely to bring their cultural food preferences with them when they migrate. A quirk of India's data collection procedure ensures that households are surveyed in groups of ten drawn from blocks of no more than 180 proximate households within a village, town or city. This feature allows me to compare migrant and non-migrant households who face very similar prices (an assumption I test using household unit values). In this setting, finding that migrants consume fewer calories than otherwise-similar locals provides the first piece of evidence that households are willing to forgo calories to accommodate their cultural preferences.

The second stage of the analysis presents this finding: migrant households consume fewer calories per person compared to non-migrant households in the same village (conditioning on household food expenditures, characteristics and demographics in a flexible manner). The average level of this "caloric tax" (the percentage gap in caloric intake between locals and migrants) is equal to 1.6 percent of caloric intake. Reassuringly, I find a similar caloric tax when I compare households where the wife migrated across a state boundary at the time of marriage to households where the wife also moved village at the time of marriage but stayed within her own state (a comparison in which the two sets of households appear very similar in terms of observables and hence are likely to be similar in terms of unobservables as well).[3] I also find no evidence that the caloric tax is restricted to well-nourished households for whom reductions in caloric intake have no nutritional consequences: the magnitude of the tax ranges between 1 and 1.7 percent when I only consider households that are particularly poor or undernourished, or households with small children for whom caloric shortfalls are particularly harmful.

The third stage of the analysis investigates why migrants consume fewer calories than non-migrants. I form a chain of evidence showing that migrants are making nutritionally-suboptimal food choices due to cultural preferences for the traditional foods of their origin states. First, I focus on the preferences themselves. I document that migrants bring their origin-state food preferences with them. In particular, I show that compared to other households in the same village, the food-budget shares of a migrant household are more-closely correlated with the average food-budget shares of their origin state. Furthermore, these preferences for the foods of their origin state are more pronounced when both husband and wife are migrants (as opposed to just one of these two being migrants). Second, I combine these preference results with caloric tax results of the second stage. I show that the heterogeneity in the size of the migrant caloric tax is related to the suitability and intensity of their origin-state food preferences: the caloric tax is only present when the average bundle of the migrant's origin state provides fewer calories than the local bundle (both priced at the village price vector), and increases in size when both husband and wife are migrants. In terms of magnitudes, the migrant households whose cultural preferences put them at the biggest disadvantage (i.e. both husband and wife migrated to a village where the typical origin-state bundle provides fewer calories per Rupee than the local bundle) face a caloric tax of 7.0 percent. The caloric tax for this group remains a substantial 5.2 percent when I restrict my analysis to undernourished households. These are substantial magnitudes. If this group of migrant households possessed the same preferences as locals, the percentage consuming nutritionally inadequate diets would fall from 58 to 47 percent.

Finally, the fourth stage of the analysis rules out two alternative explanations. Migrants may simply have poor information about the local alternatives to their origin-state foods. Alternatively, migrants may not possess the technologies, such as cooking equipment or recipes, needed to make high-quality meals from the locally-cheap foods. Both these explanations generate a link between the size of the caloric tax and the typical bundle of the migrant's origin state, but do not rely on migrant and non-migrant households having different cultural preferences. Under these alternative explanations, the caloric tax should not persist many years after migration or be present if there are non-migrants in the household who are familiar with the local foods. Similarly, the size of the tax should be smaller for more educated households. I find no evidence for any of these hypotheses. Finally, since women are typically in charge of food purchasing and preparation in Indian households, the tax should be smaller when only the husband is a migrant compared to

[3] Such a comparison also mitigates potential selection problems that arise when household heads are choosing to migrate for better employment opportunities or because they are particularly adaptable to different cultures.

when only the wife is a migrant. In fact, I find the opposite result, consistent with migrants bringing their origin-state preferences with them and husbands having greater bargaining power in household decision making.

This set of results suggests that migrants in India consume fewer calories than non-migrants because they prefer to purchase the traditional products from their origin state even when these products are relatively expensive compared to local alternatives. The finding that culture can have economically significant costs is likely to be true in many other contexts. However, there are also scenarios where culture can have positive effects on nutrition (an effect I find for the subset of migrants with preferences particularly well-suited to the local price vector). And, of course, the magnitudes I find are specific to the context of migrants within India.[4]

For policymakers, this paper shows that effective policies for combating malnutrition should take culture into account. I discuss the policy implications further in the conclusion.

This paper contributes to several literatures. First, it adds to the growing literature on the importance of culture, a topic surveyed in Guiso, Sapienza and Zingales (2006) and Fernández (2011). In using the behavior of migrants to examine the influence of culture on household decisions, it is particularly closely related to Fernández, Fogli and Olivetti (2004) and Fernández and Fogli's (2009) studies of female labor force participation and Giuliano's (2007) study of family living arrangements. In contrast to this strand of the literature, which typically demonstrates that culture can influence behavior, my approach allows me to quantify the costs that culture can impose. As nutrition impacts economic growth (Fogel, 1994), the paper is also closely related to Algan and Cahuc's (2010) study of the relationship between culture and economic growth; and Guiso, Sapienza and Zingales's (2004) study on the link between culture and financial development.

Second, I add to the literature on the persistence of food preferences initiated by Staehle (1934), with recent contributions by Bronnenberg, Dube and Gentzkow (2012) and Logan and Rhode (2010). Although these papers document that migrants bring their food preferences with them, none of them explore the nutritional consequences of such preferences. Finally, this study is related to Nunn and Qian's (2011) study of the impact of New World potatoes on the Old World. Their finding of large take up of a new crop over hundreds of years and consequent nutritional improvements suggests that the persistent food culture I find may weaken over many generations.

In a companion paper, I provide theoretical and empirical evidence for the existence of regional food preferences in India.[5] The two papers differ in that Atkin (Forthcoming) lays out a model in which the combination of agroclimatic endowments and habits generate regional food tastes that favor the locally-abundant foods, and then explores the implications for the gains from trade liberalization. In contrast, this paper takes India's regional food preferences as given, interprets these as cultural phenomena and investigates their influence on caloric intake.

I layout the remainder of the paper in the following manner. Section 2 introduces the dataset and provides a short review of the literature on cultural food preferences in India. Drawing on

[4] If migrants are more adaptable to other cultures than non-migrants, focusing on migrants may provide a lower bound measure of the calories households are willing to forgo to accommodate their cultural preferences.

[5] Both this paper and Atkin (Forthcoming) formed part of my Ph.D. thesis.

this literature, I provide motivating evidence that households are trading-off nutrition for culture. Section 3 explains how I use migrants in my empirical strategy to identify the caloric costs of culture. Section 4 presents my main result, that migrant households consume fewer calories than comparable non-migrant households. Section 5 presents evidence that this finding is driven by migrants making nutritionally-suboptimal food choices due to preferences for the favored foods of their origin states. Section 6 rules out the two non-cultural explanations. Finally, section 7 discusses the implications of my findings and concludes.

2 Background on consumption patterns and caloric intake across India

In this section, I present two findings that motivate the subsequent analysis. First, the median Indian household consumes far fewer calories than the recommended caloric intake, consistent with the extremely high levels of childhood malnutrition observed in health and nutrition surveys. Second, despite this setting, Indian households seem to be making food choices driven by culture even when these choices result in reduced caloric consumption.

2.1 Data description

My analysis draws on two cross-sections of the Indian National Sample Survey (NSS) collected by the National Sample Survey Organization (NSSO): the 38th round (1983) and the 43rd round (1987-1988). These are the only two rounds of publicly available surveys in which the same household is asked both about their consumption of a broad set of foods as well as about their migration particulars.[6] Each survey round contains approximately 80,000 rural households (located in 8,000 villages) and 45,000 urban households (located in 4,500 urban blocks). I stack the two cross-sections and create a combined data set containing 245,334 households. To simplify the exposition, I will use the word village to refer to the lowest geographic identifier (a village in a rural area but actually a block in an urban area).

The surveys record household expenditures and quantities purchased in the last 30 days (as well as quantities consumed out of home-grown stock and gifts, both valued at the prevailing local prices) for each food item. There are 169 different food items, including 12 products made from rice or wheat, 9 types of pulse, 7 milk products and many vegetables, spices and meats. I obtain calorie data for each household by multiplying each food's caloric content, estimated by the NSSO, by the quantity consumed over the previous 30 days. I use this number (along with the size of the household from the household roster) to calculate the daily caloric intake per household member.[7] The surveys also provide information on household demographics and characteristics,

[6] The migration questions are part of the employment and unemployment schedule, schedule 10. In contrast, the consumption data are collected in the Consumer Expenditure schedule, schedule 1. In more recent survey rounds, these two schedules were no longer filled out by the same households.

[7] These numbers are likely to overestimate actual caloric intake. Some of these calories are wasted (due to spoilage or simply thrown away) or fed to servants, pets and guests. If wastage rates are higher for migrant households as they are less familiar with local foods, my estimates in later sections may underestimate the difference in caloric intake be-

as well as expenditures on non-food items. Finally, the NSS provides survey weights in order to make the sample nationally representative, and all the statistics I report use these weights.

2.2 Malnutrition in India

Indian households in 1983 and 1987-88 consume a small number of calories. The mean caloric intake is 2224 per person per day across the two samples. In order to get a sense of magnitudes, recent figures covering the period 2006-2008 were 3750 calories per person per day for the United States, 2990 for China, and 2360 for India (FAO, 2008).

India drew the poverty line it still uses today based on the calorie norms required for a nutritionally adequate diet. These norms were set at 2400 calories per person per day for rural India and 2100 calories per person per day for urban India.[8] Using this simple indicator of household nutrition, 66.4 percent of rural households and 59.6 percent of urban households in my sample are undernourished. Many households lie substantially below these levels with 45.5 percent of households consuming fewer than 2000 calories and 35.6 percent consuming fewer than 1850 calories per person per day. The upper panel of Figure I shows the full distribution of caloric intake.

While there is an imperfect mapping between my measure of caloric intake and malnutrition, these low levels of caloric intake are consistent with the extremely high child-malnutrition rates in India. The first wave of the National Family Health Survey was administered in 1992-1993. The survey measured and weighed around 35,000 children under age 4 and found that 53.4 percent were moderately to severely underweight, and 52.0 percent were moderately to severely stunted. These numbers imply a higher prevalence of under nutrition than in Sub-Saharan Africa (Deaton and Drèze, 2009), and suggest that a substantial number of Indian households were living on the edge of malnutrition at the time of the surveys.

2.3 Cultural preferences for food in India

In this subsection, I provide suggestive evidence that despite these low levels of caloric intake and high child-malnutrition rates, households seem to be making nutritionally-suboptimal food choices in part due to cultural preferences for certain foods.

The first thing to note is that almost all households could have purchased a bundle providing the recommended caloric norms for a smaller outlay than they actually spent on food. For example, if every household spent all their income on the item in their consumption bundle that provided the most calories per Rupee, then the percentage of households below the recommended caloric norms would fall to 9.5 percent (compared to the 64.9 percent in the actual data). Alternatively, if every household spent the same amount per calorie as the household in their village with the highest caloric intake per Rupee, only 39.2 percent of households would be consuming

tween migrants and non-migrants. Actual caloric absorption will be even lower than actual caloric intake if household members have ailments such as diarrhea or gastroenteritis that prevent the body from fully absorbing the calories.

[8] Poverty lines were set in 1978 at the average per-capita expenditure of NSS households consuming this number of calories, and are now updated each year using the inflation rate.

below the caloric norms. However, both of these bundles are likely to be deficient in other nutrients such as proteins and vitamins, as well as highly monotonous. Therefore, the fact that some households are not purchasing the most calorically-efficient bundle does not imply that they are making nutritionally-suboptimal food choices based on cultural considerations. Households may simply be maximizing nutrition and a taste for variety.

A review of the literature on cultural food preferences in India provides important background to my paper and motivates a more convincing counterfactual.

2.3.1 Review of the literature on cultural food preferences in India

It will be important for my analysis that there are many different food cultures within India, and that these food cultures differ across states. The field of nutritional anthropology has identified many different food cultures across religious, caste and ethnolinguistic groups within India. I provide examples below of these different food cultures before explaining how the regional taste differences that these food cultures generate fit into economic definitions of culture.

Chakravarti (1974) combines fieldwork and survey evidence to categorize many dimensions of food culture in India, while Harris (1985) and Nair (1987) and Simoons (1994) focus narrowly on animal consumption. It is here, in the attitude towards animal products, where cultural food preferences vary most dramatically across India. Of the major religious groups in India, Jains and Buddhists are generally vegetarian due to a belief in non-violence towards animals. Christians, Sikhs and Muslims eat animal products although the latter will not eat pork. For the Hindu majority that comprises over 80 percent of the population, adherence to vegetarianism differs by caste.[9] Typically, members of the Brahmin (priest) and Vaishya (trading) castes are vegetarian while members of the Kshatriya (warrior) and Kayasthas (service) castes are non-vegetarian. Lower caste households vary in attitudes towards meat eating, with some groups even consuming beef, a taboo food for most Hindus.

However, this categorization by caste masks substantial regional heterogeneity. For example, Hindus of all castes eat meat in the parts of the far north states of Himchal Pradesh, Uttar Pradesh and Jammu and Kashmir. Brahmin Hindus in West Bengal consume both fish and goat but not chicken which is taboo. Meanwhile, Brahmins in Assam further east do eat chicken, as well as fish and goat. Chakravarti (1974) argues that it is difficult to provide a single explanation for this diversity of attitudes and hypothesizes that interactions with neighboring cultures, local geography and adherence to different Hindu deities all play roles.[10]

There are also differences in the acceptance of non-meat items. Vegetarians in some parts of India consume eggs but others, such as Gujaratis who follow Swaminarayan, will not. In addition, many North Indians avoid eggs in summer believing them to be a "hot" food that can harm the body by raising its temperature. Conversely, "cold" foods such as citrus fruits, are avoided in

[9] There are several justifications for Hindu vegetarian practices: the principle of non-violence towards animals (Simoons, 1994, p. 6), and the purifying "satvic" properties of cereals and most vegetables (Khare, 1992, p. 208).

[10] For example, the goddess Varahi is particularly revered in West Bengal and is depicted holding a fish, potentially explaining why Brahmins eat fish in that state.

winter when they may cause colds and pneumonia (Chakravarti, 1974; Pool, 1987). In general, balancing "hot" and "cold" foods is believed to be necessary for the body's well-being (except for pregnant women for whom hot foods are harmful and cold foods beneficial). However, as Nag (1994) catalogs in detail, which precise foods are classified as hot and which foods cold varies dramatically across regions. For example groundnut is perceived as hot in Tamil Nadu but cold in Gujarat. There is also substantial regional diversity in beliefs regarding milk consumption, even within areas where bovines were historically used for agricultural production (Simoons, 1970).[11]

Attitudes towards vegetables also vary considerably across India. High caste and strict Hindus in the north of India will refrain from eating plants in the Allium genus like onions and garlic which are thought to overexcite passions (an aversion shared by some Jains and Buddhists), while in the south of India onion is often permitted but not garlic (Behura, 1962; Simoons, 1998). There is similar variation in the consumption of mushrooms as some high caste Hindus consider them unclean since they grow in dung. Simoons (1998) even highlights regional variations in attitudes towards eating potatoes, salt and the legume urd.

Regional taste differences of the kinds described above fit squarely within the definitions of culture used in the economics literature. Fernández (2011) defines differences in culture as "systematic variation in beliefs and preferences across time, space or social groups". The variation in food preferences across states of India certainly fits this definition since state boundaries were drawn primarily along major ethnolinguistic divisions. Furthermore, many of India's religious minorities are concentrated in particular states. Guiso, Sapienza and Zingales (2006) define culture as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation". Since adult food preferences are determined in part by the foods consumed in childhood, and adults choose which foods are fed to their children, the variation in food preferences across states of India also fits this second definition of culture.[12] Accordingly, this paper treats regional and cultural food preferences as synonymous.

2.3.2 Rice and wheat consumption across India

I now turn to cereals, the major category of food consumption omitted in the above discussion and the one most relevant for nutrition. Rice and wheat are the two dominant carbohydrates in India, together accounting for an enormous 65.5 percent of total caloric consumption in my sample. The NSS data provides strong empirical evidence of regional differences in preferences for rice and wheat. The upper panel of Figure II plots the relative consumption of rice and wheat calories against the relative price of rice and wheat calories for the 22,148 villages where I observe both rice and wheat purchases.[13] Although the relationship between relative price and relative

[11]Of course, unlike every other example in this section, there is a strong genetic component to variation in milk consumption since some adult populations are lactose intolerant. In my empirical work, I will not be able to separate genetic variation in lactose tolerance across states from variation in cultural preferences for milk across states.

[12]Birch (1999) surveys the evidence from the psychology and nutrition literature which finds that adult food preferences form in part through consumption in childhood.

[13]I exclude 2698 villages because I observe zero purchases of either rice or wheat, and hence do not know the relative price. These villages would likely strengthen my findings if included since they possess extreme preferences

quantity is downward sloping, by drawing horizontal lines across the plot it is obvious that for most of the observed price ratios there are some villages consuming relatively large quantities of rice and others relatively large quantities of wheat. This observation applies both when rice is a relatively cheap calorie source compared to wheat and when it is relatively expensive.

Regional preferences for rice and wheat in India are well documented in the nutritional anthropology literature. As highlighted in the introduction, even in 300 A.D. inhabitants of Northwest India around 300 A.D. were known for their special liking of wheat while inhabitants of East India were known for their love of rice (Prakash, 1961). Chakravarti (1974) classifies modern India into three divisions based on food habits for cereals: rice consuming areas in east and along the south and south-west coastline of India, bread consuming areas in the north and northwest of the country, and rice and bread consuming areas in the center of the country.

As an illustrative example, the lower panel of Figure II highlights two states with substantial price overlap, Kerala in the south and Punjab in the north. Keralans consumed thirteen times more rice than wheat and Punjabis ten times more wheat than rice at similar relative prices (the Kerala fixed effect in a regression of rice budget shares on relative prices for households in the two states is a massive 0.80).

Do these regional preferences for rice and wheat in India mean that households are consuming fewer calories than they could if they only cared about nutrition and dietary variety? To answer this question, I perform the following counterfactual. For every household, I calculate the amount it would cost to purchase a bundle with the same total quantity of calories they currently obtain from rice and wheat but swapping the quantities of rice and wheat.[14] For 42 percent of households, this bundle would cost less than they currently spend on rice and wheat. These are the households in the north-east and south-west quadrants of the upper panel of Figure II (with 94 percent of the switchers in the north-west quadrant). For this subset of households, I calculate the hypothetical caloric intake if the household allocated the cost savings to the cheaper of the two foods. I set the hypothetical and actual caloric intake equal for the remaining 58 percent of households. For the average household, the hypothetical caloric intake is 6.2 percent higher than their actual intake. This gap between actual and hypothetical caloric intake provides a measure of the number of calories households forgo in order to accommodate their preferences for rice or wheat.

This counterfactual solves several of the problems associated with the cruder counterfactuals at the start of this subsection. First, by switching the quantities allocated to rice and wheat, the amount of dietary variety is not declining (in the sense that every household still obtains the same number of calories from the less-consumed food). Second, these two foods provide a similar number of calories per Rupee and contain a similar set of nutrients. Therefore, switching between rice and wheat does not alter household nutrition along other dimensions.[15]

for rice or wheat. Relative consumption is the share of rice calories in total rice and wheat calories consumed by sample households in the village. The relative price is the log of the ratio of the median caloric unit values (i.e. expenditures divided by calories) across purchasing households.

[14] For households that don't consume both rice and wheat, I use the village median price as the price for the unpurchased cereal. As before, I restrict attention to the villages where I observe both rice and wheat purchases.

[15] The average unit value for rice was 0.95 Rupees per 1000 calories in 1983 and 1.09 in 1987/88. The equivalent

Although informative, some of the switching households may prefer rice or wheat for idiosyncratic reasons, rather than because of the regional preferences which have a cultural dimension to them. In order to highlight the cultural dimension, I classify states as having either "rice-loving" or "wheat-loving" cultures. Atkin (Forthcoming) shows that, over many generations, the process of habit formation leads to regional preferences for the foods the region is relatively well-suited to produce. Therefore, I utilize measures of land suitability from the Global Agro-Ecological Zones project (GAEZ) to proxy for each state's rice/wheat culture. The GAEZ data report the suitability of each state in India for growing both rice and wheat on a scale of 0 (not suitable) to 1 (very high suitability).[16] Consistent with the preferences for wheat and rice highlighted above, Punjab is better suited to wheat production and Kerala to rice (the difference between the crop-suitability score for rice and wheat is 0.29 in Kerala and -0.13 in Punjab). Reassuringly for my proxy, land suitability strongly predicts relative consumption of rice and wheat across states, even after conditioning on relative prices: A simple OLS regression of the household rice budget share on both relative rice prices at the village level and the state-level relative suitability for growing rice (the difference between the score for rice and wheat) yields a coefficient of 0.97 with a t-value of 149.41.

I label states as possessing a rice-loving culture if their GAEZ suitability measures are higher for rice than wheat, and a wheat-loving culture otherwise. With these labels in hand, I redo the counterfactual but only allow two types of consumer to switch their consumption patterns: households who are consuming more rice than wheat in states with a rice-loving culture but where wheat is a cheaper calorie source than rice, and households consuming more wheat than rice in states with a wheat-loving culture but where rice is cheaper than wheat. Allowing only these two types of household to switch around their rice and wheat consumption, I estimate that Indian households are forgoing 6.1 percent higher caloric intake in order to accommodate their cultural preferences.[17] Figure I plots the full distribution of caloric intake under the counterfactual, as well as a histogram of the caloric gains generated. If the foregone calories were realized, the percentage of households consuming less than the recommended caloric norms would fall by 7.3 percent. Of course, if many households did switch across rice and wheat, local prices would change and the actual gain in caloric intake would likely be smaller.

numbers for wheat were 0.66 and 0.75. According to the NSS, rice contained 3460 calories per kilogram, 75 grams of protein and 5 grams of fat. The equivalent numbers for wheat where 3410, 121 and 17. Wheat has slightly more fiber and vitamins, particularly folic acid. Thus, if anything, wheat is more nutritious than rice. As the caloric gains in the counterfactual come almost entirely from households switching from rice to wheat, nutrition is also likely to improve along other dimensions with the switch.

[16]The particular measure I use is the "crop suitability index" for rain-fed agriculture using intermediate input usage. The GAEZ website http://www.iiasa.ac.at/Research/LUC/GAEZv3.0/ contains further details. Data are only available for 28 of the 31 states. For rice, my suitability measure is the maximum of the two state-level index values for wetland and dryland rice cultivation.

[17]The rice suitability measure is higher than the wheat suitability measure for 23 out of the 28 GAEZ states, and 39.5 percent of households switch in this counterfactual. Alternatively, I can classify a state as rice-loving if the state fixed-effect in a regression of rice budget shares on relative prices is greater than 0.5. Using this measure, 25 out of 31 states are classified as rice-loving. The counterfactual results are very similar under this definition, with Indian households forgoing 6.0 percent higher caloric intake in aggregate and 37.7 percent of households switching. In both cases, almost all of the gains come from households in rice-loving states switching into relatively cheap wheat since wheat is a cheaper calorie source on average.

This last exercise provides suggestive evidence that households in India make food choices based on cultural preferences that are suboptimal from a nutritional sense. However, since every village has a different vector of relative prices, some of these consumption patterns may be rationalized by complex substitution patterns. For example, if coconuts are much stronger complements with rice than they are with wheat, the relatively high consumption of rice in the south of India may be due to cheap coconuts rather than a preference for rice. In the next section, I discuss a methodology that sidesteps this issue by considering all 169 foods in my data set and by focusing on consumption differences between households that have different cultural preferences yet face the same prices.

3 Empirical methodology: examining the behavior of migrants

The behavior of inter-state migrants provides more compelling evidence that culture can constrain caloric intake. The key idea is that migrants bring the regional preferences of their origin state with them, yet they face the same relative prices as non-migrants in their destination state.[18] If I show that migrant households consume bundles that resemble the consumption bundle of households in their origin state and that these bundles provide fewer calories than the typical local bundle, I can interpret the finding that migrants consume fewer calories per Rupee of food expenditure than locals as evidence that migrants pay a "caloric tax" to accommodate their cultural food preferences. Such an interpretation is reasonable as long as migrants do indeed face the same prices as non-migrants and value variety and other dimensions of nutrition to the same degree as non-migrants (two issues I will address later in this section).

Inevitably, this methodology can only estimate the caloric tax that actual migrants pay. If potential migrants are aware of this cost of migration, actual migrants are likely to face smaller caloric taxes either because they avoid locations with particularly deleterious price vectors or because they possess particularly open-minded or flexible preferences. Hence, the potential size of this tax may be much larger for households that choose not to migrate.

Prior to discussing the assumptions behind my identification strategy, it is useful to describe the migration information contained in the NSS data. By design, the survey only records permanent migrations (as opposed to temporary migrations for seasonal work opportunities). The survey asks whether the enumeration village differs from the household member's "last usual residence". If so, the household member is asked the reason for migration, how long ago they migrated and the state in which their last usual residence was located. I define inter-state migrants as households in which either the household head or their spouse moved between one of the 31 states in India. Except where noted otherwise, I use the household head's migration information if both the household head and the spouse emigrated. Since the household head is male and the spouse is female in 99.7 percent of cases, I use the terms household head and husband

[18] There is a long history of using migrants to identify the effects of culture. Fernández (2011) describes the approach in detail and reviews this literature. The use of migrants to highlight the persistence of regional food preferences was pioneered by Staehle (1934).

interchangeably, and the terms spouse and wife interchangeably.

Table I provides summary statistics for the dataset. Under the migrant definitions above, about 6.1 percent of households are classified as migrant households. Most of these households are long-term migrants, with 41.3 percent having migrated 20 or more years ago and only 15.2 percent having migrated less than 5 years ago. Finally, in India there is a norm of "patrilocality" whereby wives move in with their husband's family upon marriage. This norm appears in the data with the largest category of migrant households being those in which only the wife is a migrant (41.3 percent of cases). In a further 32.5 percent of cases, both the husband and wife are migrants from the same origin state, and in only 13.2 percent of cases is the husband the sole migrant. I exploit these norms later in this section in order to control for potential confounding factors.

The identification argument in the empirical analysis relies on the following assumptions: migrants and non-migrants living in the same geographic location face the same prices and external environment, have the same desire for good nutrition and dietary variety, and differ only in their food preferences (after controlling for various household expenditure measures, household demographics and observable household characteristics). I argue that these assumptions are likely to be satisfied for the following reasons.

First, the NSSO uses a methodology that allows me to make this comparison at an extremely disaggregated geographic level, thereby making such an assumption more tenable. In each survey round the NSSO draws a sample of around 8,000 rural villages and 4,500 urban blocks from the 1981 census rolls and surveys 10 households in each village/block. In order to reduce the work load for the survey enumerators, any village or urban block with more than 1200 inhabitants (approximately 180 households) is subdivided into smaller geographical subgroups and only one subgroup is surveyed. Thus, in a village with one surveyed migrant, I compare the migrant household to the other nine non-migrant households surveyed in the same subgroup (a subgroup which never encompasses more than about 180 proximate households).

Second, Indian migration patterns are not concentrated along only a few specific migration routes. If this were the case, the assumption that migrants and non-migrants have the same desire for nutrition and dietary variety may be violated. For example, suppose most migrants in India come from Kerala and Keralans particularly value nutrition. In this scenario, if migrants (i.e. Keralans) consume fewer calories than locals (i.e. non-Keralans) I cannot infer that they also consume fewer nutrients. Such a concern is mitigated if migrants come from many origin states (decreasing the likelihood that all migrants place a high value on nutrition or variety) and if migrants move in both directions between states (and so migrants and non-migrants place an equal value on nutrition or variety). Table II displays the proportion of all migrants that moved between every pair of origin and destination states. Unsurprisingly, the larger states in India are either the source or destination of most of the migrant flows (and the city-state of Delhi is a major recipient). However, the routes are dispersed with migrants moving from many different states and often in both directions, mitigating the concerns stated above.

Third, migrants do face the same prices as non-migrants, at least after controlling for observable characteristics. Table III compares both household characteristics and prices paid across migrant and non-migrant households within the same village. I discuss the characteristic and price comparisons in turn.

The characteristics I focus on are the set of controls used by Subramanian and Deaton (1996) when estimating caloric elasticities using the 1983 NSS survey. I regress each characteristic on a village fixed effect and a migrant-household dummy and report the coefficient on the dummy. Compared to other households in their village, migrant households have 6.2 percent higher per-capita expenditures, 4.5 percent higher per-capita food expenditure, and consume 1.3 percent more calories per person. Migrant households are slightly smaller, contain a larger proportion of prime-age males, are less likely to be categorized as an agricultural laborer household, and are more likely to be categorized as urban self-employed. In my empirical specifications, I include explicit controls for all of these characteristics (the full set of controls are described in section 4).

The last three rows test the assumption that migrants face the same prices as non-migrants. I calculate household-level prices by dividing household expenditure on a food by the calories purchased. The first row shows the coefficient on a migrant-status dummy when the log price per calorie is regressed on a product-village fixed effect and a migrant-status dummy. The second row shows the same coefficient but including expenditure controls in the shape of a cubic in log household food expenditure per capita (allowing the coefficients on food expenditure to differ by survey round). The third row shows the same coefficient but with the full set of controls for food expenditure and household characteristics described in section 4 (essentially the other variables in Table III).[19] Migrants do seem to pay about one third of one percent more than non-migrant households in the same village buying the same product, presumably by buying slightly higher quality levels. However, this difference is due to migrants being wealthier than non-migrants in their village since the difference disappears once I control for household food expenditure. The coefficient on the migrant dummy is also small and not significantly different from zero when the characteristic controls are included in addition to food expenditure controls. Therefore, in order to ensure that migrants and non-migrants in the same village are paying the same prices, all my specifications will include these controls. Additionally, I reproduce my main findings after repricing all household purchases using the village median prices to confirm that my results on calories consumed per Rupee of expenditure are driven by differences in the foods consumed rather than the prices paid.

After controlling for observables, migrant and non-migrant households may still differ on unobservables. As an important robustness check, I draw on an alternative sample in which migrant and non-migrant households appear far more similar along observable dimensions, and hence are likely to be more similar along unobservable dimensions. This sample also mitigates the potential

[19] Since there are multiple products for each household, there are over 5 million observations in this regression. Unlike the previous characteristic regressions which are weighted by household survey weights, these regressions are weighted by the household survey weights multiplied by the food-budget share spent on that product. This weighting ensures that more important prices are weighted more heavily, as well as ensuring that the sum of weights for each household is equal to the household survey weight from the NSS.

selection problems that arise when household heads are choosing to migrate for better employment opportunities or because they are particularly adaptable to different cultures.[20]

I take advantage of the fact that a substantial proportion of migration in India is driven by women moving to their husband's village at the time of marriage (Srinivas, 1980). This norm of "patrilocality" is so prevalent that 57 percent of wives report that both their current village is not their last usual residence and list the reason for leaving that location as "on marriage".[21] Most of these moves occur within the same state, with wives crossing a state border at the time of marriage in only 6 percent of cases. I exploit this variation by focusing just on households in which the wife moved for marriage, and comparing households in which the wife moved inter-state (migrant households) to households in which the wife moved intra-state (non-migrant households). In the spirit of the exercise, I also exclude the households in which both the husband and the wife moved at the same time since these households may be moving for work opportunities.

Although the same proportion of households happen to be classified as migrants in both this "*wife moved for marriage*" sample and the main sample, the average migrant household differs substantially across the samples. Table I includes descriptive statistics for both samples. Migrant households in the *wife moved for marriage* sample are more likely to live in rural areas (and hence appear more similar to the general population which is predominantly rural) and are more likely to be long-term migrants.

Table III confirms the conjecture that migrant and non-migrant households appear more similar in the *wife moved for marriage* sample. Although migrant households still spend more on both food and all goods, the difference in expenditures between migrant and non-migrant households declines by a third. Migrant households are no longer smaller, nor do they contain a larger proportion of prime-age males. Finally, migrants pay less, not more, than non-migrants for the same product, and these differences are miniscule and insignificant with or without controls. Therefore, I reproduce all my main findings using this *wife moved for marriage* sample in order to convince the reader that my results are not likely to be driven by unobservable differences between migrants and non-migrants.

4 Migrants consume fewer calories per Rupee than non-migrants

In this section, I present the first empirical result: that migrant households pay a "caloric tax". In particular, I test the hypothesis that migrants consume fewer calories per Rupee of food expenditure compared to the non-migrant households living around them.

In order to test this hypothesis, I use the data on the consumption of all 169 foods to generate

[20] To produce my findings on caloric intake, migrants need to consume higher price per calorie foods than non-migrants with similar incomes for reasons unrelated to the tastes of their origin state. The selection bias likely works in the other direction. For example, migrants may be more likely to be manual laborers who consume diets heavy in cheap carbohydrates, or migrants may have unusually adaptable and adventurous tastes.

[21] In contrast, under 1 percent of male household heads moved location for the purpose of marriage. Among male household heads who do move, 48 percent cite employment reasons and only 18 percent cite marriage (these figures are 2 percent and 85 percent for wives).

$\ln calories_i$, the log of caloric intake per person per day, for every household (where i indexes households). I regress this measure on $migrant_i$, a dummy variable for a migrant household, and d_{vt}, a village-round fixed effect (where v denotes village or urban block and t denotes the survey round). The village-round fixed effect is equivalent to a village fixed effect since villages are anonymized and cannot be matched across the two survey rounds. Additionally, I include a vector of household-level controls, $\mathbf{X}_i$, containing a third-order polynomial in the log of the per-capita food expenditure over the previous 30 days, as well as a comprehensive set of demographics and characteristics that follow the specification used by Subramanian and Deaton (1996):

$$\ln calories_i = \beta_1 migrant_i + d_{vt} + \mathbf{\Pi}_t \mathbf{X}_i + \varepsilon_i. \tag{1}$$

The hypothesis $\beta_1 < 0$ tests whether migrants consume fewer calories than their neighbors in the same village, conditional on their food expenditure and other household-level controls. Given the inclusion of log food expenditure in the controls, this test is exactly equivalent to asking if migrants obtain fewer calories per Rupee of food expenditure than non-migrants in their village, conditioning on food expenditure and other household-level controls.

The characteristic and demographic variables control for the possibility that, compared to other households in the village, migrants may work in less physically-intensive jobs or have different demographic structures. Household demographics are captured by log household size as well as the proportion of household members that fall into five sex-specific age buckets.[22] The included household characteristics are indicator variables for the household's primary activity among the following categories: rural self-employed in agriculture, rural self-employed in non-agriculture, rural agricultural labor, rural other labor, rural other, urban self-employed, urban wage earner, urban casual labor and urban other. I allow the coefficients on all these controls to differ by survey round. Subramanian and Deaton (1996) also include indicators for religion and whether the household is a member of a scheduled caste. Since religious affiliation and caste membership may be cultural determinants of food preferences, I do not include these as controls.

The error terms may be correlated across households within the same village and across households that share the same origin state. Therefore, both here and in the regressions that follow, I two-way cluster the standard errors at both the village-round and origin-state level.

Column 1 of Table IV shows the results of this regression. I reject the null hypothesis, that migrants consume an equal or greater number of calories per Rupee than non-migrants, at the 1 percent level: inter-state migrant households are consuming 1.59 percent fewer calories than their non-migrant neighbors, controlling for food expenditure. In monetary terms, this caloric tax on migrants is commensurate with the caloric decline due to a 2.47 percent reduction in food expenditure for the average migrant household.[23]

As discussed in section 2, caloric intake is not equivalent to nutrition, and households may

[22]These age buckets are 0-4, 5-9, 10-15, 15-55 and over 55.

[23]I calculate these numbers using the round-specific coefficients on the expenditure controls combined with the mean log per-capita food expenditure of migrants in each round.

trade-off calorie-rich foods for protein- or vitamin-rich foods. However, there is no reason to think that migrants would trade-off these components of a nutritious diet in a different manner to non-migrants facing the same prices and spending the same amount on food (recall that migrants are moving between many different states and often in both directions and so any state variation in preferences for proteins or vitamins should sweep out in the aggregate). Thus, the smaller number of calories per Rupee that migrants consume likely implies a lower level of nutrition.

The magnitude of the caloric tax does not mean that cultural preferences for food can only have small impacts. First, the size of the caloric tax should depend on how costly it is for a migrant to accommodate their origin-state food preferences. If the origin-state preferences are well-suited to the local price-vector, migrants may actually consume more calories for a given level of food expenditure. The coefficient on $migrant_i$ merely summarizes the average caloric tax faced by migrants traveling along a multitude of routes and facing positive and negative caloric taxes.[24] Second, recall that for many of these households only one member of the household (usually the wife) migrated from another state. Any effects are likely to be more exaggerated if both husband and wife are migrants since a greater proportion of household decision-makers possess non-local preferences. In sections 5.3 and 5.4, I explore both these dimensions of heterogeneity and find substantially higher caloric taxes for the more adversely affected migrant groups.

4.1 Robustness checks

The remaining columns of Table IV report a variety of robustness checks. Column 2 of Table IV runs the specification in equation 1 on the *wife moved for marriage* sample, described in section 3, for which unobservable differences between migrant an non-migrant households were less of a concern. I compare households where wives moved intra-state at the time of marriage (a non-migrant household) with those where the wife moved inter-state (a migrant household). The caloric tax on migrants is still significantly negative for this sample, but is attenuated by 24 percent.[25] The decline in the size of the coefficient is not surprising. These wives are typically moving into their husband's households (often containing other extended family members such as the husband's parents). Any cultural preferences brought by the wife are likely to have a smaller impact on household spending decisions compared to the scenario where both husband and wife are migrants (a hypothesis I test in a more direct manner in sections 5.2 and 5.4).

Columns 3 to 5 of Table IV use alternative sets of expenditure controls in place of the polynomial in log per-capita food expenditure. Column 3 uses a third-order polynomial in log per-capita expenditure on all goods. I find a migrant caloric tax of 1.36 percent. Since the coefficients are of similar magnitude when I control for either food expenditure or total expenditure, migrant households are not simply substituting from non-food to food expenditure in order to accom-

[24] The average caloric tax faced by migrants will be negative if local preferences adapt through the process of habit formation to favor whichever foods are locally inexpensive as in Atkin (Forthcoming).

[25] I would also find $\beta_1 < 0$ if wives from more distant villages are more valued and fed higher quality foods. Alternatively, I would find $\beta_1 < 0$ if wives consume cheaper calorie sources than other household members, combined with wives from further away consuming less food (or controlling a smaller share of the household budget).

modate their food preferences. Results are similar in column 4, which uses a polynomial in log per-capita real food expenditure, and in column 5, which allows the food expenditure elasticities to vary by state.[26]

In column 6 of Table IV, I instrument the food expenditure polynomial with a polynomial in non-food expenditure to control for potentially correlated measurement error (since both calories and food expenditure are calculated using the same raw data).[27] Food expenditure may also be endogenous. For example, a shock that increases the demand for calories, such as changing work patterns, will also affect food expenditure and result in a positive correlation between food expenditure and the error term, biasing the coefficients on food expenditure upwards. However, there will be a negative or no correlation with non-food expenditure, and so the true value of the coefficient will be bounded between the instrumented and uninstrumented estimates. Since the estimated β_1 is only attenuated by a quarter and still significantly negative in the instrumental variables specification, both measurement error and the potential endogeneity of food expenditure do not seem to be a major problem in this context.

Finally, I address the concern that migrants and non-migrants may pay slightly different prices (recall from table III that although the difference in prices paid was insignificant after conditioning on food expenditures, it was still positive). This concern is most severe for the three foods (rice, wheat and sugar) that are commonly sold through the subsidized Public Distribution System (PDS).[28] Although the system was not restricted to households with Below Poverty Line cards until the 1990's, migrants may still have had worse access to this system and hence paid higher prices for these three foods even after conditioning on quality. Column 7 reproduces my main specification excluding these three food groups (i.e. replacing $\ln calories_i$ and the food expenditure measures with the calories from, and expenditure on, the remaining foods). Reassuringly, I actually find a larger caloric tax of 2.05 percent when I exclude these PDS foods. Columns 8 and 9 take a different approach to show that my results are driven by differences in the foods consumed rather than the prices paid. Column 8 replaces $\ln calories_i$ with $\ln calories\_per\_Rupee_i$, the calories per Rupee spent on food, and includes controls for per-capita expenditure (recall that the coefficient on $migrant_i$ would be identical to column 1 if per-capita food expenditure controls were included instead). Column 9 uses the same specification but calculates $\ln calories\_per\_Rupee_i$ by pricing each food at the village median price for that food. I find that migrants obtain 1.45 percent fewer calories per Rupee of food expenditure than their non-migrant neighbors, conditioning on total expenditure, and this caloric tax is essentially unchanged at 1.49 percent when I price each food at the village median price. Therefore, migrants consume fewer calories than locals through purchasing different consumption bundles rather than through paying different prices.

[26] Column 4 uses log per-capita food expenditure divided by a state-specific Stone price index (the sum of log prices weighted by state budget shares). Column 5 interacts log per-capita food expenditure with state-round fixed effects.

[27] The first stage is very strong with a Cragg-Donald F-statistic of 36.8.

[28] I can reproduce the prices paid regressions of table III just for foods in these three groups. Although still not significantly different from zero at the 10 percent level, migrants pay 0.24 percent more for these foods than non-migrants in the same village after conditioning on the full set of controls.

4.2 Migrants consume fewer calories even when on the edge of malnutrition

Migrants may be willing to accommodate their cultural preferences but only if they are sufficiently rich and well-nourished that any foregone calories are irrelevant (and may even be beneficial). Accordingly, Table V repeats the basic specification for various sub-populations that are poor and under-nourished (i.e. I compare the caloric intake of poor and undernourished migrant households to that of poor and undernourished non-migrant households in the same village).

Column 1 repeats the baseline specification. Columns 2 through 4 restrict attention to undernourished households (those consuming fewer than either 1850 or 2000 calories per person per day, or those consuming fewer calories than the 2400 rural/2100 urban calorie norms used to calculate Indian poverty lines). Columns 5 through 8 restrict attention to poorer households (those spending less than either the median or 25th percentile of either per-capita expenditure or per-capita food expenditure in that survey round). Although the size of the caloric tax paid by migrants is slightly smaller for these seven subpopulations, it still lies between 1.0 and 1.7 percent.[29]

Nutritional shortfalls at young ages have substantial scarring effects on productivity, earnings and health in adulthood (Almond and Currie, 2010). Thus, adequate nutrition is particularly important for households with young children. Columns 9 through 12 restrict attention to families with children below the age of 5 or below the age of 16. I refine both these samples further by restricting attention only to households that are spending less than the median level of per-capita food expenditure. In all four of these subpopulations, I find that migrant households consume significantly fewer calories per Rupee than non-migrant households, with magnitudes ranging between 1.3 and 1.7 percent. Given that India's child malnutrition rates were in excess of 50 percent around this time period, these poor households with children are very likely to be on the edge of malnutrition.

In summary, migrant households consume fewer calories per Rupee of food expenditure than non-migrant households, even when on the edge of malnutrition.

5 Why do migrants consume fewer calories than non-migrants?

The previous section showed that migrant households consume fewer calories than comparable non-migrant households and that this result holds even for households on the edge of malnutrition. In this section, I form a chain of evidence in support of an explanation based on culture: that migrants make nutritionally-suboptimal food choices due to strong preferences for the favored foods of their origin states.

First, I focus on the preferences themselves. In section 5.1 I document that migrant households bring their origin-state food preferences with them when they migrate and in section 5.2 I show that the intensity of these preferences depends on whether both husband and wife are migrants

[29] The coefficients are significantly negative in all cases except when considering the bottom quartiles of total expenditure or food expenditure. In these two cases, the sample size is dramatically smaller, and the standard errors are much higher (rather than the magnitude of the coefficients being much lower).

(as opposed to just one of these two being migrants). Second, I combine these preference results with the earlier results regarding caloric intake. Sections 5.3 and 5.4 find that the heterogeneity in the size of the migrant caloric tax is related to the suitability and intensity of these origin-state food preferences: the caloric tax is only present when the average bundle of the migrant's origin state provides fewer calories than the local bundle (both priced at the local price vector), and it increases in size when both husband and wife are migrants (again compared to just one of these two being migrants). This chain of evidence suggests that Indian migrants consume fewer calories than non-migrants because they prefer to purchase the favored products from their origin state even when these products are relatively expensive compared to local alternatives.

5.1 Migrants bring their food preferences with them

In this subsection, I present the first piece of evidence, that migrants bring their food preferences with them. In particular, I test the hypothesis that, compared to other households living in the same village, a migrant household's consumption bundle more closely resembles the average bundle of their origin state.

I first present a simpler specification that just focuses on the consumption of rice and wheat. I test whether the amount of rice a migrant household consumes is related to the amount of rice that households in their origin state consume. I can do this in two ways. As is typical in the economics of culture literature surveyed by Fernández (2011), I can just focus on migrants and test whether migrants who come from rice-loving states spend more on rice than migrants in the same village who come from wheat-loving states. Alternatively, I can test whether migrants who come from states that are more rice-loving than their current state spend more on rice than locals.

I regress rice's share of total household rice and wheat expenditure, $\frac{rice_i}{rice_i+wheat_i}$, on the average rice share of their origin state, $\frac{rice_i^o}{rice_i^o+wheat_i^o}$, a measure of the household's relative preference for rice and wheat based only on their origin state (where the origin-state average is denoted by an o superscript and is calculated using only non-migrant households interviewed in the same survey round as household i):[30]

$$\frac{rice_i}{rice_i + wheat_i} = \alpha_1 \frac{rice_i^o}{rice_i^o + wheat_i^o} + d_{vt} + \mathbf{\Pi}_t \mathbf{X}_i + \varepsilon_i. \tag{2}$$

For non-migrants, the origin state rice share is simply the average rice share of their current state. The regression specification also includes the same village fixed effects, d_{vt}, and vector $\mathbf{X}_i$ of household-level controls used in section 4.

I first restrict attention only to migrant households. Since I include village fixed effects, a positive α_1 coefficient indicates that migrants who moved from states that are more rice-loving than the origin states of other migrants within their village consume a larger share of rice than other migrants (and vice versa for migrants from more wheat-loving states). Column 1 of Table VI

[30] When calculating the total expenditure on either rice or wheat, I include all 12 of the rice and wheat-based products in the surveys (e.g. wheat, baking flour, cake flour, semolina flour, noodles and bread).

contains the results of this simple regression. I find support for the hypothesis with a positive and highly significant estimate of α_1 equal to 0.189. I can also include non-migrants in the regression. With all villagers included, a positive α_1 coefficient indicates that migrants who moved from states that are more rice-loving than their destination state consume a larger share of rice compared to the locals in their village. Column 2 of Table VI contains the results of this regression. Once more, I find a positive and highly significant estimate of α_1 (here equal to 0.123).[31]

Although informative, such an exercise only incorporates information on two types of food. If I wish to consider all 169 food items, I require a different approach to test whether migrants bring their food preferences with them. One option is to repeat the exercise above for all 169 foods and then aggregate the coefficients in some manner. However, the preponderance of zero quantities for many of the less-consumed foods means that it is difficult to compare consumption across households within the same village on a good-by-good basis.[32]

Instead, I propose an intuitive and transparent measure of preference similarity based on correlations between household consumption bundles and a reference consumption bundle for a particular state. I calculate $\rho_i^s = corr(\mathbf{bshare}_i, \mathbf{bshare}_i^s)$, the correlation between the vector of 169 food-budget shares of household i ($\mathbf{bshare}_i$) and the vector of average food-budget shares of a particular state s ($\mathbf{bshare}_i^s$). As with the rice-wheat specification, the state-averages are calculated using only non-migrant households interviewed in the same survey round as household i (hence the need for an i superscript on $\mathbf{bshare}_i^s$). This budget share correlation naturally over-weights the food items with high budget shares, a desirable property if I want to explore the link between preference differences and differences in total caloric intake.[33]

These ρ_i^s correlations provide a simple measure of the similarity between household i's preferences and the average preferences of non-migrants in state s.[34] I test whether migrant and non-migrant households possess the same preferences by comparing the size of these correlations

[31] If all households allocated expenditures on rice and wheat in the same proportions as the average household in their origin state, the α_1 coefficient would equal 1. The smaller coefficient may be a result of either migrant adaptation to local preferences, or preferences that are not Cobb-Douglas. In unreported results, I also run both specifications using rice calorie shares instead of expenditure shares and obtain estimates of 0.181 for the migrant-only sample and 0.118 for the full sample (both significantly different from zero at the 1 percent level).

[32] If I regress the budget share of each food in the total food budget on the average food-budget share for that food in the household's origin state (with the same controls as in equation 2), 145 of the 180 coefficients are positive and 72 are significantly greater than zero at the 5 percent level. There are more than 169 different foods since several food categories changed between the 38th and 43rd survey round.

[33] To see this fact, note that the correlation between vectors $\mathbf{x}$ and $\mathbf{y}$ is equal to $\frac{\sum(x_j y_j - \bar{x}\bar{y})}{(n-1)s_x s_y}$ where $\bar{x}$ and s_x denote the mean and standard deviation of vector $\mathbf{x}$. The mean budget share for any vector is equal to $\frac{1}{169}$. Therefore "outlier" foods with high average budget shares will typically have larger values for $\sum(x_j y_j - \bar{x}\bar{y})$ and hence be more influential. For example, take any $\mathbf{x}$ and swap any two elements x_1 and x_2. The correlation between the new and the old bundle equals $1 - \frac{(x_1-x_2)^2}{(n-1)s_x^2}$ and will only differ substantially from 1 if one of the swapped elements is a large part of the budget.

[34] If households have Cobb-Douglas preferences, $u = \Pi_{g=1}^{169} c_g^{\theta_g}$ with $\sum_{g=1}^{169} \theta_g = 1$ (where c_g is the consumption and θ_g the preference parameter for food g), my preference-similarity measure is the correlation between a household's preference parameters and the average preference parameters of non-migrants in state s. Atkin (Forthcoming) proposes an alternative Almost Ideal Demand System approach that allows for non-homotheticities and budget shares that respond to prices. Such a methodology is less feasible in this context since it is difficult to estimate migrant-specific preference when there are very few migrants from a particular origin state in a particular destination state. In the robustness analysis, I present a modification that allows for non-homotheticities in consumption.

across households that face the same price vector.

As a first step, I test whether, compared to other households living in the same village, a migrant's consumption bundle less closely resembles the average bundle of their current state of residence. I regress the correlation ρ_i^d of a household's bundle with their current state bundle (labeled state d as it is a migrant's "destination" state) on a migrant-household dummy:

$$\rho_i^d = \beta_1 migrant_i + d_{vt} + \mathbf{\Pi}_t \mathbf{X}_i + \varepsilon_{ist}. \tag{3}$$

As in previous specifications, I include village fixed effects, d_{vt}, and vector $\mathbf{X}_i$ of household-level controls. A negative value of β_1 indicates that migrant households consume bundles that are less similar to the current state bundle (in comparison to non-migrant households in the village). As shown in column 3 of Table VI, the data support this sign prediction. I find a estimated coefficient of -0.0111. I can reject the null hypothesis, that migrants do not differ from non-migrants in the manner described above ($\beta_1 \geq 0$), at the 1 percent level.[35]

The finding that migrants possess different preferences than non-migrants does not necessarily imply that migrants bring with them preferences for the specific foods of their origin state. I now test this hypothesis. I focus only on villages with migrants living in them, and compare the similarity of the bundles of both migrants and non-migrants in the village to the migrant's origin-state bundle. To do this, I switch correlation measures to the correlation $\rho_i^{o_v}$ of a household's bundle with the bundle of the origin state of migrants within their village (where o_v indicates the origin state of migrants in village v, distinct from the o superscript which indicates the origin state of the household itself). I regress this correlation on a dummy variable $migrant_i^{o_v}$ indicating a household that contains a migrant from state o_v:

$$\rho_i^{o_v} = \gamma_1 migrant_i^{o_v} + d_{vt}^{o_v} + \mathbf{\Pi}_t \mathbf{X}_i + \varepsilon_{ist}. \tag{4}$$

Villages may have multiple origin states o_v if there are migrants from more that one state living there. In this scenario, there are multiple observations per household, one for each origin state in the village. Therefore, I include a separate village fixed effect for each origin state in each village, $d_{vt}^{o_v}$, in addition to the set of household-level controls used in the previous specifications.[36] A positive value of γ_1 indicates that migrant households originally from origin-state o_v consume bundles that are more similar to the bundle of that particular origin state o_v (in comparison to how similar the bundles of neighboring households not from o_v are to the bundle of origin state o_v). As shown

[35]The negative coefficient on $migrant_i$ is partly mechanical since average budget shares of state d were calculated using only non-migrant households. Although this bias is likely to be small (the average state-round sample contains 3,700 non-migrant households), I reproduce the regression using average bundles calculated using all households. The β_1 coefficient remains significantly negative at the 1 percent level, rising only slightly to -0.0106.

[36]As previously, I two-way cluster at the village-round and origin state o of the household. Since I compare all households to the average bundle of a migrant's origin state, an alternative is to cluster at the village-round and o_v-state level. Clustering at the household level is also sensible since there are multiple observations per household. The standard errors are very similar under the first two clustering procedures, and smaller with household-level clustering. Therefore, I report the more conservative standard errors that use the first procedure.

in column 4 of Table VI, the data support this sign prediction. I estimate a positive coefficient of 0.0226 and can reject the null hypothesis at the 1 percent level. I find that, compared to other households living in the same village, a migrant's consumption bundle more closely resembles the average bundle of the migrant's origin state.[37]

I assess the magnitudes of the coefficients in the following manner. On average, migrants still consume bundles that are more closely correlated with the reference bundle of their current state than their origin state (the average correlations are 0.7270 and 0.6712 respectively). However, for comparable non-migrant households, the gap between the two correlations is substantially larger (the β_1 and γ_1 coefficients imply that the current-state correlation is 0.0111 higher and the migrant-state correlation is 0.2226 lower). Therefore, migrants close about 40 percent of this dissimilarity gap (i.e. the gap between the correlation with the current state bundle and the correlation with the migrant state bundle).

Columns 5 through 9 of Table VI run the main regression specified in equation 3 for the *wife moved for marriage* sample, for the alternative expenditure specifications detailed in section 4.1, and for the subset of non-PDS foods. In addition to these robustness checks, my findings are robust to using two alternative preference-similarity measures. First, migrant households may come from different parts of the income distribution than the average household in their origin state. Since migrants are not observed before their migration, any correction is necessarily imperfect. Column 10 presents one such correction. I recalculate the reference bundles using non-migrant households in the same national income quartile as household i (again in the round that the household was surveyed). Therefore, I compare the correlation between a household's bundle and the bundle consumed in state o_v by households at similar income levels. Second, although budget shares have the appealing feature that they map directly into parameters of the utility function if food preferences are of the Cobb-Douglas form, column 12 calculates the correlations using vectors of caloric shares instead (where a caloric share is a food item's share of household caloric consumption). Results are similar across all the robustness specifications, with γ_1 significantly positive in every regression.

Finally, Table VII reports these correlation results for the various subsamples of poor and undernourished households detailed in section 4.2. Mirroring the caloric tax results, I find positive (and significant) coefficients across the various subsamples.[38]

In summary, migrant households bring the cultural food preferences of their origin state with them when they migrate.

[37] There is some heterogeneity across both food types and household characteristics. If I focus on subsets of products, I find more substantial persistence among the 24 cereals or just rice and wheat products (γ_1 coefficients equal to 0.0273 and 0.0230 respectively) than among non-cereals (a γ_1 coefficient equal to 0.0159). If I include interactions between the migrant dummy and either a rural indicator or the proportion of children in the household, I find significantly less persistence among rural households and those with a higher proportion of children.

[38] The only two exceptions are the bottom quartile of food or total expenditure subsamples in which the sample size shrinks by more than 90 percent.

5.2 The number of migrants in the household increases the intensity of preferences

If the results of the previous subsection are driven by cultural preferences for the foods of a migrant's origin state, I expect more pronounced effects when there are multiple migrants within the household. In this subsection, I show that there are stronger preferences for origin-state foods when both husband and wife are migrants as opposed to only one of the two (since in the former scenario both primary decision makers in the household possess non-local preferences).

Table VIII explores the heterogeneity across these within-household migrant structures. I allow the coefficient on the migrant dummy in equation 4 to vary with household structure by interacting $migrant_i^{o_v}$ with dummies for the the migrant status of the household head and their spouse: (1) only one of either the head or spouse is a migrant ($onlyone_i$), or (2) both the head and spouse are migrants ($both_i$). I treat migrant households where the head has no spouse as a third category and also interact a no spouse dummy ($nospouse_i$) with the migrant dummy. However, since households with no spouse may differ from other households more generally, I also include the no spouse dummy in the controls. Equation 4 becomes:[39]

$$\begin{aligned}\rho_i^{o_v} \quad = \quad & \gamma_1 migrant_i^{o_v} \times onlyone_i + \gamma_2 migrant_i^{o_v} \times both_i \\ & + \gamma_3 migrant_i^{o_v} \times nospouse_i + d_{vt}^{o_v} + \pi_{1t} nospouse_i + \mathbf{\Pi}_t \mathbf{X}_i + \varepsilon_i. \qquad (5)\end{aligned}$$

The γ coefficients from this specification provide separate estimates of the similarity of migrant bundles to their origin-state bundle for each of these three types of migrant household. The hypothesis at the start of the subsection corresponds to $\gamma_2 > \gamma_1$: the similarity of migrant consumption bundles to their origin-state reference bundle (compared to the similarity of non-migrant bundles to the same reference bundle) is stronger when both husband and wife are migrants, and weaker when only one is a migrant. I find support for this hypothesis in Panel 1 of Table VIII. When only only one of the husband and wife is a migrant, I obtain a coefficient on the migrant dummy of 0.0079. In contrast the size of the caloric tax is significantly larger when both husband and wife are migrants (a coefficient of 0.0416). I can reject the null that the coefficients on $migrant_i^{o_v} \times onlyone_i$ and $migrant_i^{o_v} \times both_i$ are equal at the 1 percent level.

In summary, migrant households exhibit stronger preferences for the foods of their origin state if both husband and wife are migrants as opposed to only one of the two.[40] Section 6 dismisses alternative explanations for my findings based on information or technology rather than culture. In doing so, I explore further differences in the intensity of preferences based on the gender of the migrant, the time since migration, and the similarity of the origin-state and destination-state bundles.

[39]If both head and spouse are migrants but come from different origin states, I replace the $migrant_i^{o_v}$ indicator variable with the value of one half for each of the two origin states. Since there are very few such households, results are essentially unchanged if these households are dropped.

[40]The focus on the migrant status of the household head and spouse, as opposed to all household members, seems appropriate. If I supplement equation 4 with an interaction between the migrant dummy and the proportion of household members that are migrants, the interaction term is positive and highly significant. However, if I also include the migrant structure dummies used in equation 5, the proportion of migrants interaction is no longer significant.

5.3 The size of the caloric tax depends on the suitability of the migrant preferences

This subsection links together my two previous findings: that migrants bring their origin-state tastes with them and that migrants consume fewer calories per Rupee than locals. I show that the size of the caloric tax paid by migrants depends on how well-suited their origin-state preferences are to the local price vector. In particular, I test the hypothesis that the size of the caloric tax is larger if migrants move to a village where the preferences of their specific origin state place them at a caloric disadvantage relative to locals.

In order to test this hypothesis, I require a measure of how calorically advantageous a certain set of origin-state preferences is. Once more, I proxy the migrant's origin-state preferences with their origin-state reference bundle, $\mathbf{bshare}_i^o$, a vector of average food-budget shares of non-migrants in their origin state. I then calculate $\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v)$, the log of calories derived from 1 Rupee allocated in the same proportions as this origin-state reference bundle $\mathbf{bshare}_i^o$ but with foods priced at the destination-village price vector $\mathbf{P}_i^v$.[41] Similarly, I calculate $\ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)$, the log of calories derived from 1 Rupee allocated in the same proportions as the average bundle $\mathbf{bshare}_i^v$ of non-migrant households in the migrant's destination village (also at destination-village prices). The log difference between the calories derived from each of these 1 Rupee bundles measures the caloric advantage of a migrant's origin-state preferences over the local preferences. Migrants who move to villages where their origin-state average bundle is a relatively expensive method of obtaining calories compared to the local bundle have a negative value for this log difference. These migrant households have particularly disadvantageous preferences and should face a larger caloric tax compared to a migrant household for whom this log difference is positive.

To implement this test, I rerun my calorie regression, equation 1, except I now interact the migrant dummy with an indicator variable, $\mathbf{1}[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$, that takes the value of 1 for negative values of the log difference described above:

$$\ln calories_i = \beta_1 migrant_i + \beta_2 migrant_i \times \mathbf{1}[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)] + d_{vt} + \mathbf{\Pi}_t \mathbf{X}_i + \varepsilon_i. \quad (6)$$

The values of the log differences range from -0.62 for the 1st percentile of migrants to 0.92 for the 99th percentile, with negative values for one third of migrant households. As before, I include village-round fixed effects, d_{vt}, and the same vector $\mathbf{X}_i$ of controls for expenditure and household demographics described in section 4.

The hypothesis at the start of the subsection corresponds to $\beta_2 < 0$: the caloric tax is more negative if a migrant's origin-state bundle provides relatively few calories per Rupee compared to the local bundle. Column 1 of Table IX presents the results of this regression. I can reject the null hypothesis at the one percent level. The estimated β_2 coefficient is significantly negative and equal

[41] I obtain the vector $\mathbf{P}_i^v$ by treating unit values (the expenditure on a food divided by the quantity purchased) as price data. Unit values are not actual prices since quality varies. In part because of this concern, I use median village prices as my price measure. These prices are robust to outliers and are less contaminated by quality effects. If none of the village sample purchase a good, I use the median price at an incrementally higher level of aggregation.

to -0.0283. The main effect, β_1, is insignificant and close to zero. Migrants only pay a caloric tax if they live in a village where purchasing their origin-state reference bundle provides fewer calories than the local bundle. Summing the two coefficients, I find that migrant households living in villages where their preference are badly suited to the local price vector consume 3.33 percent fewer calories than comparable non-migrant households.

Columns 2 and 3 of Table IX present alternative specifications for the $\ln K(.,.)$ interaction. Column 2 allows the caloric tax to vary for migrants in each quartile of the $\ln K(.,.)$ difference. The largest caloric tax of 3.93 percent is faced by migrant households in the bottom quartile. The size of the caloric tax becomes progressively smaller for households in the second and third quartiles and becomes significantly positive for the top quartile.[42] This top quartile of migrants have the most advantageous origin-state preferences and receive a caloric dividend rather than pay a caloric tax. In terms of magnitudes, these households consume 2.27 percent more calories per Rupee than their non-migrant neighbors. Column 3 interacts the migrant dummy with $\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) - \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)$, a continuous measure of a migrant's caloric advantage over locals. Unsurprisingly, I find a positive and significant coefficient of 0.0629: the size of the caloric tax increases with the caloric disadvantage of a migrant's origin-state preferences.[43]

As in previous sections, I present many robustness checks and results for poor and undernourished subpopulations. The other columns of Table IX report my findings using the *wife moved for marriage* subsample, alternative expenditure controls, only non-PDS foods, income-quartile adjusted reference baskets, instrumented food expenditure, and replacing calories with calories per Rupee calculated with both actual and village median prices. Table X contains results for poor and undernourished subpopulations. In all 20 specifications, I find a significantly negative β_2 coefficient. The magnitudes for migrants with disadvantageous preferences (e.g. for whom $\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)$) range from a caloric tax of 1.87 percent for households below median income, to a caloric tax of 3.46 percent for households consuming fewer than 2000 calories per person per day.

I perform two additional robustness tests that address particular concerns with this exercise. One concern is that the $\ln K(.,.)$ difference is related to the distance migrants have traveled and

[42]The three quartile boundaries are -0.071, 0.118 and 0.322. It is not surprising that I find a negative caloric tax for small positive values of the $\ln K(.,.)$ difference (i.e. for the third quartile). First, the origin-state bundle is an imperfect measure of migrant preferences unless the utility function takes a Cobb-Douglas form described in footnote 34, is fixed for life and is identical for every person born in that state. One departure that produces caloric taxes for positive $\ln K(.,.)$ differences is if richer households obtain fewer calories per Rupee due to non-homotheticities. In this scenario, the true $\ln K(.,.)$ difference is more negative than the measured difference if migrant incomes rise upon migration. Second, if migrants pay slightly higher prices than non-migrants, the caloric advantage of the origin-state bundle over the local bundle is actually a small disadvantage when priced at the prices migrants actually pay rather than village median prices I use in $\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v)$. In support of this explanation, in columns 5 and 8 I find no negative caloric tax for the third quartile, yet similar results for other quartiles, if I exclude PDS foods or if I calculate calories per Rupee priced at the median village prices (the two robustness specifications dealing with the concern that migrants paid slightly higher prices).

[43]The fact that the coefficient is substantially smaller than one implies that either migrants adapt their preferences after migrating, or not all household members possess such preferences, or preferences are not of the Cobb-Douglas/identical-within-state form (or some combination of these three explanations). I will provide some explicit evidence for the first two of these explanations in later sections.

migrants from far-off places differ from other migrants in some unobservable way. I address this concern by including an additional interaction between $migrant_i$ and the log distance between the migrant's destination region (a subset of their state) and their origin state. The coefficient on the distance interaction is insignificant while the β_2 coefficient is essentially unchanged (column 13 of Table IX). A second concern is that measurement error in $\ln calories_i$ will be correlated with measurement error in $\ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)$ for non-migrant households (recall that $\ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)$ is the average calories per Rupee for non-migrants in the village). The average state-round sample contains 3,700 non-migrant households and so any bias due to measurement error should be small at higher levels of disaggregation. Accordingly, column 14 calculates the $\ln K(.,.)$ difference using average bundles at the state level instead of at the village level (still pricing the bundles at destination-village prices). The β_2 coefficient remains negative and significant in this specification.

In summary, I establish a clear link between the specific preferences of a migrant's origin-state and the caloric tax paid by migrants. The size of the tax is larger when a migrant's origin-state preferences are badly-suited to the local price vector.

5.4 The size of the caloric tax depends on the intensity of the migrant preferences

In this subsection, I provide further support for a cultural explanation by showing that the caloric tax paid by migrants is related to the intensity of their preferences for origin-state foods. Since household preferences for origin-state foods are more intense if multiple household members possess those preferences (as shown in section 5.2), I test the hypothesis that the size of the caloric tax paid by migrants is larger if both husband and wife are migrants as opposed to just one of the two.

I interact the migrant terms in the caloric tax specification, equation 1, with the same set of migrant-structure dummies I used in section 5.2:

$$\ln calories_i = \beta_1 migrant_i \times onlyone_i + \beta_2 migrant_i \times both_i + \beta_3 migrant_i \times nospouse_i + d_{vt} + \pi_{1t} nospouse_i + \mathbf{\Pi}_t \mathbf{X}_i + \varepsilon_i. \quad (7)$$

The β coefficients provide separate estimates for the caloric tax faced by migrants for each of these three structures. Panel 2 of Table VIII reports these regression coefficients. The results mirror the findings of section 5.2. I find support for the hypothesis that $\beta_2 > \beta_1$. When only only one of the husband and wife is a migrant, I obtain a coefficient on the migrant dummy of -0.0125. In contrast, the size of the caloric tax is significantly more negative when both husband and wife are migrants (a coefficient of -0.0228).[44]

Panel 3 of Table VIII performs a similar breakdown for equation 6 of the previous subsection.[45] For each of the three migrant structures, I find that the size of the caloric tax is larger when the

[44] I reject the null that the coefficients on $migrant_i \times onlyone_i$ and $migrant_i \times both_i$ are equal with a p-value of 5.3.

[45] If both head and spouse are migrants but come from different origin states, I take the average value of $\mathbf{1}[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$ across the two origin states. Since there are very few such households, results are essentially unchanged if these households are dropped.

migrant's origin-state reference bundle provides fewer calories per Rupee than the local bundle (both priced at the local price vector). The ordering of the size of the tax is also consistent with the hypothesis above.[46] The most adversely affected households (households in which both husband and wife migrated to a village where their origin-state reference bundle provides fewer calories than the local bundle) face a caloric tax of 7.0 percent.

The magnitude of this caloric tax is substantial. The median caloric intake for this migrant subgroup is 2134 calories per person per day with 58 percent of households consuming less than the recommended calorie norms (2400 calories in rural areas, 2100 in urban). If these migrants had the same preferences as locals, the median would rise to 2292 calories and the percentage of households below the caloric norms would fall to 47 percent. As with the earlier specifications, these effects are not limited to better-nourished households. For example, panels 4 through 6 reproduce the specifications in both this subsection and section 5.2 on the subsample of households consuming fewer than 2000 calories per person per day. The size of the caloric tax for this same group of households, those in which both husband and wife are migrants with unsuitable preferences, remains a sizable 5.2 percent (column 2 of panel 6).

In summary, I find strong evidence that culture can constrain caloric intake. I find that migrants are bringing their food preferences with them, and that the caloric tax is larger when the favored foods of their origin-state are expensive compared to local alternatives. Further corroborating a food-culture explanation, the migrant households that pay the largest caloric tax are those with multiple migrants that possess these unsuitable preferences.

6 Alternative explanations

Up to this point, I have presented a chain of evidence in support of the hypothesis that migrants in India are making nutritionally-suboptimal food choices due to strong preferences for the favored foods of their origin states. First, I showed that migrants bring their food preferences with them. Second, I showed a strong link between the size of the caloric tax and the local cost of the reference bundle of the migrant's origin state. These findings are inconsistent with a story in which migrant preferences differ from those of non-migrants but in a manner unrelated to their cultural origins. However, these findings do not contradict a story where migrants possess better information or technology, rather than stronger preferences, for the foods of their origin state. I now discuss these two alternative explanations.

6.1 An information story

The first alternative explanation is that migrants have poor information about local prices or about the availability and nutritional properties of local alternatives to their origin-state foods. Under these scenarios, migrants would consume fewer calories per Rupee than non-migrants as

[46] I can reject at the 1 percent level the null that the size of the caloric tax for migrants with unsuited preferences is the same if only one of the head and spouse are migrants or if both are migrants.

they are unaware of cheaper alternatives. Migrants may also consume bundles that more closely resemble their origin-state reference bundle since they are more familiar with these foods. In this subsection, I provide five pieces of evidence that contradict this information-based explanation for my findings.

First, I document that the caloric tax is persistent and remains many years after migration. Even if migrants are initially uninformed, after many years in the destination village they would become familiar with the local foods and prices. I exploit the data on years since migration and rerun my main regression specifications on subpopulations that exclude recent migrants. Specifically, I exclude migrant households where the most recent migrant arrived less than 5, 10 or 20 years prior to the survey. Columns 13 to 15 of Table V presents these three regressions for the basic calorie specification, equation 1. The caloric tax remains significantly negative for the first two long-term migrant specifications, although the size of the tax declines. When I exclude all migrants who arrived less than 20 years previously, the tax disappears altogether. However, the specifications from section 5 tell a more complete story. Although the coefficients are progressively attenuated as I remove the more recent migrants, long-term migrants still consume bundles more closely related to their origin-state bundle than locals do (columns 13 to 15 of Table VII), and still pay a caloric tax if they move to locations where their origin-state bundle provides fewer calories per Rupee than the local bundle (columns 13 to 15 of Table X).[47] Therefore, even migrants who have had many years to learn about local foods and prices pay a caloric tax when their origin-state preferences are unsuited to the local price vector.

Second, I find evidence of a caloric tax on migrants when only one of the husband or wife are migrants (Panel 1 of Table VIII), and even when wives are moving to their husband's village (column 2 of Table IV). In these cases, other household members already possess information about local foods and prices yet the caloric tax remains.

Third, in Indian society it is typically women who are in charge of the purchase and preparation of foods. Therefore, under an information-driven story the caloric tax should be stronger if wives rather than husbands are migrants. On the other hand, in traditional societies such as India, men typically have greater bargaining power in household decision making; therefore, under a preference-driven story, the caloric impacts due to a migrant in the household will be stronger if the husband is a migrant as opposed to the wife. I evaluate these two competing hypotheses. Panels 1 to 3 of Table XI present similar specifications to Table VIII but break the "only one" category into two categories: only the head is a migrant and only the spouse is a migrant. I find no support for the information-driven prior. In fact, across each of the three regressions, the ordering of the coefficients across these two categories is in accordance with the preference-driven prior above (i.e. I find larger correlations and caloric taxes if the husband is a migrant as opposed to the wife). For example, in Panel 2, the size of the caloric tax if only the spouse is a migrant is around 1.0 percent and rises to 1.9 percent when only the husband is a migrant.[48] One possible explanation for

[47] For migrants who left 20 or more years ago the size of the tax is 1.79 percent (significant at the 1 percent level).

[48] I can reject the hypothesis that the coefficient on only spouse is equal or greater that the coefficient on only head at the 6 percent level. Similarly, I can reject at the 5 percent level the same hypothesis in panel 3 for migrants who move

these findings is that the husband's mother is in charge of household food purchases and preparation and that husbands bring their mothers with them when they migrate. Under this scenario, I would find larger caloric taxes if the husband is a migrant even under a pure information story. However, the husband's mother is only present in 13.6 percent of households where the husband and not the spouse is a migrant and, as shown in panels 4 to 6, results are essentially unchanged when these households are excluded.

Fourth, if the explanation is that migrants have poor information, the migrant tax is likely to be smaller among literate segments of the population who can acquire information more easily. I find the opposite relationship in the data. Column 16 of Tables V, VII and X restrict attention to households in which the household head is literate. The size of the caloric tax actually grows larger when I focus on this subpopulation.

Finally, inconsistent with a story where migrants are simply unaware of local alternatives, I present evidence that migrants do adjust their purchasing behavior when their origin-state preferences are particularly unsuited to the local price vector (e.g. when their origin-state bundle is more costly than the local bundle). I return to the preference-similarity regression, equation 4, and interact the migrant dummy with an indicator for a negative value of $[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) - \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$. I report this regression in column 11 of table VI. I find a significantly negative coefficient on the double interaction $migrant_i^{o_v} \times \mathbf{1}[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$ corresponding to a 50 percent decline in the effect size. Therefore, migrants seem to be aware that substituting away from their origin-state foods can improve nutrition since they moderate their consumption choices in contexts where consumption of these foods is most disadvantageous. However, the adaptation is incomplete (i.e. the sum of the coefficient on the double interaction and the coefficient on $migrant_i^{o_v}$ is still positive and significantly different from zero at the 1 percent level). Even in contexts where their origin-state preferences are calorically disadvantageous, migrant households still consume bundles that more closely resemble the bundles consumed in their origin state (consistent with my previous finding that these migrant households consume fewer calories than locals).

6.2 A technology story

The second alternative explanation is that migrants do not possess the technologies to make high-quality meals using the locally-cheap foods. These technologies encompass cooking and food-preparation equipment as well as recipes and techniques that turn raw foods into enjoyable meals. For example, a family in Punjab may be expert at transforming wheat into delicious roti (a flat bread), but may lack the training or equipment to make a tasty dosa (a rice-based pancake). If the family migrated to Kerala, they may continue to consume wheat as they enjoy well-prepared meals over badly-prepared meals rather than wheat over rice.

Once more the evidence from various subpopulations in the data contradicts a story in which technology is the sole explanation for my findings. If technology was responsible, the caloric

to locations where their origin-state bundle provides fewer calories than the local bundle.

tax should disappear for long-term migrants. These households have spent many years away, providing a sufficient time frame over which to purchase new equipment as well as learn new recipes and techniques. Similarly, there should be no tax for migrants who are moving into a non-migrant household since in these households there should be appropriate kitchen equipment already present and migrants can learn recipes and preparation techniques from other household members. Finally, as discussed above, women are typically in charge of food preparation in Indian households. Therefore, if a lack of recipes and preparation techniques were the cause of the caloric tax, the tax should be smaller when only the husband is a migrant compared to when only the wife is a migrant. As shown in the previous subsection, I find substantial caloric taxes for all these subpopulations and a larger tax for migrant husbands than for migrant wives.

7 Conclusion and policy implications

This paper sets out to answer a simple question: do food cultures matter in an economic sense, and in particular, can culture constrain caloric intake and contribute to malnutrition? I address this question by exploiting the fact that migrants and non-migrants face the same relative prices, yet possess very different preferences. Drawing on detailed household survey data from India, I find that inter-state migrants consume fewer calories per Rupee of food expenditure compared to their non-migrant neighbors. This caloric tax on migrants corresponds to 1.6 percent of caloric intake and is evident even for households on the edge of malnutrition. I then provide a chain of evidence in support of an explanation based on culture: that migrants make nutritionally-suboptimal food choices due to strong preferences for the favored foods of their origin states. First, I document that migrants bring their origin-state food preferences with them when they migrate and that these preferences are stronger when there are more migrants in the household. Second, I show that the heterogeneity in the size of the migrant caloric tax is related to the suitability and intensity of these origin-state food preferences. The most adversely affected migrants (households in which both husband and wife migrated to a village where their origin-state preferences are unsuited to the local price vector) would consume 7 percent more calories if they possessed the same preferences as their neighbors.

These results provide insight into the value that households place on their culture. Even households on the edge of malnutrition, a population for which reductions in caloric intake have serious repercussions for both health and economic well-being, are willing to substantially reduce their caloric intake in order to accommodate their cultural food preferences.

In terms of policy, the finding that culture can constrain caloric intake has important implications for tackling hunger and malnutrition. The cultural causes of hunger need to be understood when designing programs to alleviate malnutrition. Three types of program are particularly relevant: programs that provide food aid or price subsidies to consumers; programs that reduce tariffs or use other trade policies to increase food imports; and programs that aim to develop bio-fortified or high-yield crop varieties. In all three cases, the programs will be more effective if the targeted foods are those favored by households on the edge of malnutrition.

As a concrete example, white maize is greatly preferred to yellow maize in much of Africa.[49] However, much food aid to Africa comes in the form of imported yellow maize, and vitamin-A bio-fortification currently involves the addition of carotenes which turn the maize yellow-orange. Programs that provide cheap yellow maize to hungry communities, or try to reduce vitamin-A deficiency through wider availability of bio-fortified maize, are less effective in contexts where there are cultural preferences for white maize. Food vouchers that allow consumers to choose their favored foods or bio-fortification of traditional foods may prove more successful in such cases. Similarly, the introduction of high-yield varieties (HYV) of rice, wheat and yellow maize spurred "the green revolution" in much of the developing world. However, this revolution bypassed Sub-Saharan Africa. Alongside a range of other factors, adoption of these three HYV crops was held back by strong local preferences for Sub-Saharan staples such as sweet potato, cassava, sorghum, teff and white maize.[50]

Another potential remedy, and one mooted by the Bengal Famine Inquiry Commission, involves facilitating preference changes through campaigns that encourage the consumption of alternative foods. The commission notes that such a campaign was implemented in Ceylon with some success in order to increase Australian wheat consumption following the blockage of rice imports during World War II. However, efforts may be better targeted at children who have less-rigid preferences and even then may be slow to yield results:

> As long as rice is available, rice eaters in general will consume it in preference to other grains and in such circumstances "eat more wheat" campaigns are not likely to be very effective. ... If school-feeding schemes are developed, alternative cereals could be used for school meals. ... Further, if children learn to take such foods, they may carry the preference into later life. Children are more flexible in their dietary habits than adults. Whatever methods are adopted in the attempt to encourage the use of wheat in place of rice, progress is likely to be slow. (Famine Inquiry Commission, 1945)

A fruitful avenue for further research would be to explore the dynamics of food cultures and to better understand how nutritionally-beneficial preferences develop.

[49] See (McCann, 2005) for the historical origins of this preference for white maize. Muzhingi et al. (2008) and De Groote and Kimenju (2008) provide empirical evidence for this preference ordering.

[50] See Paarlberg (2010) for a more complete discussion of the reasons for the failure of Africa's green revolution.

References

Algan, Yann, and Pierre Cahuc. 2010. "Inherited Trust and Growth." *The American Economic Review*, 100(5): 2060–92.

Almond, Douglas, and Janet Currie. 2010. "Human Capital Development Before Age Five." NBER Working Papers 15827.

Atkin, David. Forthcoming. "Trade, Tastes and Nutrition in India." *The American Economic Review*.

Behura, NK. 1962. "Meals and Food Habits in Rural India." *Bulletin of the Anthropological Survey of India*, 11(4): 111–137.

Birch, Leann L. 1999. "Development of Food Preferences." *Annual Review of Nutrition*, 19(1): 41–62.

Bronnenberg, Bart J., Jean-Pierre H. Dube, and Mathew Gentzkow. 2012. "The Evolution of Brand Preferences: Evidence from Consumer Migration." *The American Economic Review*, 102(6): 2472–2508.

Chakravarti, AK. 1974. "Regional Preference for Food: Some Aspects of Food Habit Patterns in India." *The Canadian Geographer/Le Géographe canadien*, 18(4): 395–410.

Deaton, Angus, and Jean Drèze. 2009. "Nutrition in India: Facts and Interpretations." *Economic & Political Weekly*, 44(7): 43.

De Groote, Hugo, and Simon Chege Kimenju. 2008. "Comparing Consumer Preferences for Color and Nutritional Quality in Maize: Application of a Semi-Double-Bound Logistic Model on Urban Consumers in Kenya." *Food Policy*, 33(4): 362–370.

Famine Inquiry Commission. 1945. *Report on Bengal.* Government of India. p. 176–178.

FAO. 2008. "Food Consumption Statistics." *http://www.fao.org/fileadmin/templates/ess/documents/food_security_statistics/FoodConsumptionNutrients_en.xls*, Accessed: 2013-04-30.

Fernández, Raquel. 2011. "Does Culture Matter?" In . Vol. 1 of *Handbook of Social Economics*, , ed. Alberto Bisin Jess Benhabib and Matthew O. Jackson, 481 – 510. North-Holland.

Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." *The Quarterly Journal of Economics*, 119(4): 1249–1299.

Fernández, Raquel, and Alessandra Fogli. 2009. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility." *American Economic Journal: Macroeconomics*, 1(1): 146–77.

Fogel, Robert W. 1994. "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy." *The American Economic Review*, 84(3): 369–395.

Giuliano, Paola. 2007. "Living Arrangements in Western Europe: Does Cultural Origin Matter?" *Journal of the European Economic Association*, 5(5): 927–952.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2004. "The Role of Social Capital in Financial Development." *The American Economic Review*, 94(3): 526–556.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2006. "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives*, 20(2): 23–48.

Harris, Marvin. 1985. *Good to Eat: Riddles of Food and Culture.* Simon and Schuster.

Khare, R.S. 1992. "Annambrahman: Cultural Models, Meanings, and Aesthetics of Hindu Food." In *The Eternal Food: Gastronomic Ideas and Experiences of Hindus and Buddhists.* , ed. R.S. Khare, 201–220. SUNY Press.

Logan, Trevon, and Paul Rhode. 2010. "Moveable Feasts: A New Approach to Endogenizing Tastes." Working Paper, Ohio State University.

McCann, James. 2005. *Maize and Grace: Africa's Encounter with a New World crop, 1500-2000.* Harvard University Press.

Muzhingi, Tawanda, Augustine S. Langyintuo, Lucie C. Malaba, and Marianne Banziger. 2008. "Consumer Acceptability of Yellow Maize Products in Zimbabwe." *Food Policy*, 33(4): 352 – 361.

Nag, Moni. 1994. "Beliefs and Practices about Food during Pregnancy: Implications for Maternal Nutrition." *Economic and Political Weekly*, 29(37): pp. 2427–2438.

Nair, K Narayanan. 1987. "Animal Protein Consumption and the Sacred Cow Complex in India." In *Food and Evolution: Toward a Theory of Human Food Habits.* , ed. Marvin Harris and Eric B. Ross, 445–454. Temple University Press.

Nunn, Nathan, and Nancy Qian. 2011. "The Impact of Potatoes on Old World Population and Urbanization." *The Quarterly Journal of Economics*.

Paarlberg, Robert. 2010. *Food Politics: What Everyone Needs to Know.* Oxford University Press.

Pool, R. 1987. "Hot and Cold as an Explanatory Model: The example of Bharuch district in Gujarat, India." *Social Sciences and Medicine*, 25(4): 389–399.

Prakash, Om. 1961. *Food and Drinks in Ancient India: from Earliest Times to c. 1200 AD.* Munshi Ram Manohar Lal. p. 1–33.

Simoons, Frederick J. 1970. "The Traditional Limits of Milking and Milk Use in Southern Asia." *Anthropos*, 65(3/4): 547–593.

Simoons, Frederick J. 1994. *Eat Not This Flesh: Food Avoidances from Prehistory to the Present.* University of Wisconsin Press.

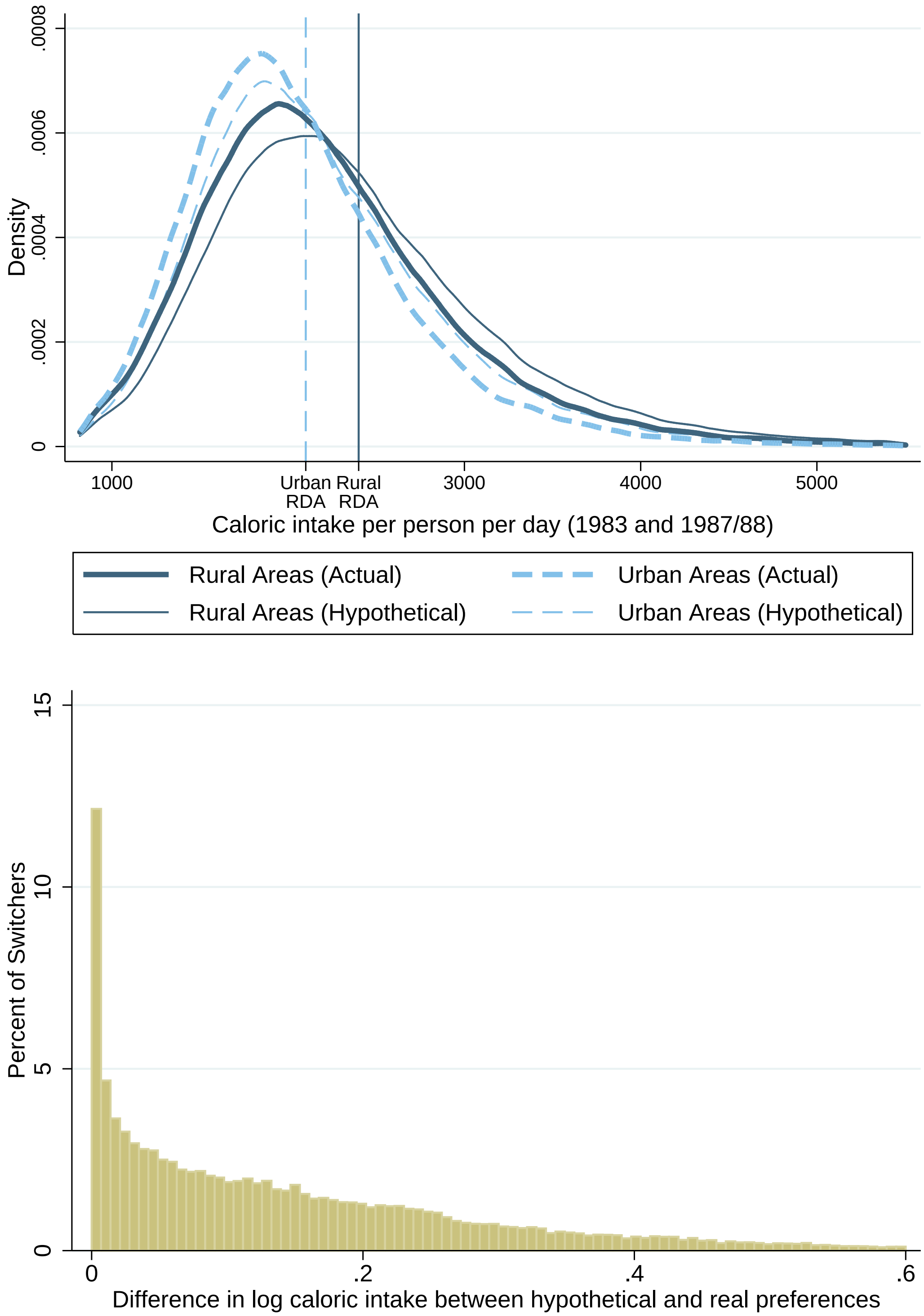
Simoons, Frederick J. 1998. *Plants of Life, Plants of Death.* University of Wisconsin Press.

Srinivas, M.N. 1980. *India: Social Structure.* Transaction Publishers.

Staehle, Hans. 1934. "The Reaction of Consumers to Changes in Prices and Income: A Quantitative Study in Immigrants' Behavior." *Econometrica*, 2(1): 59–72.

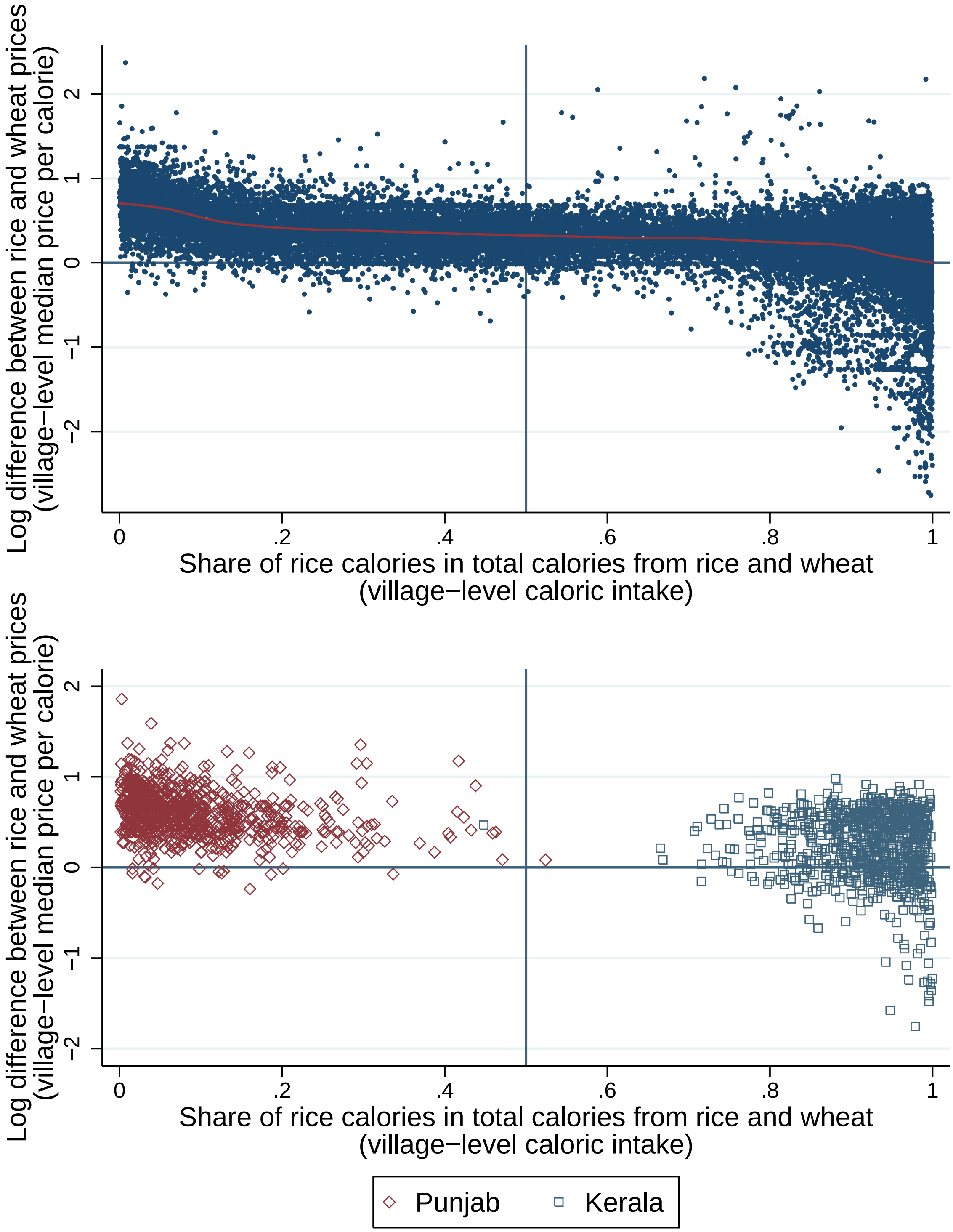
Subramanian, Shankar, and Angus Deaton. 1996. "The Demand for Food and Calories." *Journal of Political Economy*, 104(1): 133.

Figure I: Caloric intake under actual and hypothetical preferences for rice and wheat



Note: The upper panel displays the counterfactual distribution of caloric intake if households switched the quantities purchased of rice and wheat where advantageous, spending any cost savings on the cheaper of the two foods. The distribution includes all households in the 22,148 villages where I observe both rice and wheat purchases in the two survey rounds. The lower panel displays a histogram of the caloric gains available to the switchers in the counterfactual exercise.

Figure II: Regional consumption patterns for rice and wheat



Note: The upper panel plots the relative price of rice and wheat calories against the share of rice calories in the total calories from rice and wheat for the 22,148 villages where I observe both rice and wheat purchases in the two survey rounds. The lower panel highlights the observations for villages in Kerala and Punjab.

Table I: Sample descriptive statistics

	Number of households (unweighted)	Proportion of full sample	Proportion of sample -rural	Proportion of sample - migrant households	Proportion of migrants - rural households
Full sample	240,081	1.000	0.778	0.061	0.459
Wife moved for marriage sample	93,823	0.487	0.847	0.061	0.576

Migrant breakdown by within-household migrant structure	No spouse	Only spouse is migrant	Only head is migrant	Both migrants (same state)	Both migrants (different state)
Full Sample	0.112	0.413	0.132	0.324	0.019
Wife moved for marriage sample	0.000	0.807	0.000	0.173	0.020

Migrant breakdown by years since migration (most recent migrant in household)	0 to 4 years	5 to 9 years	10 to 20 years	20 or more years
Full Sample	0.152	0.141	0.293	0.413
Wife moved for marriage sample	0.072	0.137	0.320	0.470

Migrant breakdown by per capita expenditure and nutrition	Consuming below 1850 calories per capita per day	Consuming below 2000 calories per capita per day	Consuming below 2400/2100 calories (rural/urban)	Below median per capita expenditure	Below 25th percentile of per capita expenditure
Full Sample	0.338	0.433	0.566	0.307	0.126
Wife moved for marriage sample	0.338	0.436	0.591	0.369	0.164

Note: Table shows the proportion of the sample households in various categories. All proportions use the household weights provided by the NSS.

Table II: Proportion of migrant households by migration route

Destination State: State Name	State No.	Origin State: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Total	State Size
A&N Islands	1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.2	0.0
Andhra Pradesh	2	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	1.5	0.1	0.0	0.1	0.8	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.1	0.0	0.9	0.0	0.0	0.0	4.4	8.2
Arunachal	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1
Assam	4	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.5	2.2
Bihar	5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.1	0.0	0.0	0.0	0.3	0.8	1.8	10.1
Chandigarh	6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.1	0.0	0.9	0.1
Dadra & N Haveli	7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Delhi	8	0.0	0.1	0.0	0.1	0.3	0.0	0.0	0.0	0.0	0.1	1.8	0.1	0.1	0.0	0.1	0.0	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.5	0.0	0.1	0.1	4.1	0.2	8.8	0.9
Goa, Daman & Diu	9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.2
Gujarat	10	0.0	0.1	0.0	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.3	0.8	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.7	0.0	0.0	0.0	0.5	0.1	3.2	4.7
Haryana	11	0.0	0.0	0.0	0.0	0.2	0.1	0.0	0.8	0.0	0.1	0.0	0.1	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	2.1	0.0	0.0	0.0	1.7	0.0	6.9	2.0
Himachal Pradesh	12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.6
Jamu & Kashmir	13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.9
Karnataka	14	0.0	1.5	0.0	0.0	0.2	0.0	0.0	0.0	0.1	0.1	0.0	0.1	0.0	0.0	0.5	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	2.0	0.0	0.1	0.0	6.4	5.4
Kerala	15	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.0	1.8	3.4
Lakshdweep	16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Madhya Pradesh	17	0.0	0.2	0.0	0.0	0.3	0.0	0.0	0.1	0.0	0.2	0.1	0.0	0.0	0.0	0.1	0.0	0.0	1.6	0.0	0.0	0.0	0.0	0.1	0.0	0.2	1.3	0.0	0.0	0.0	2.6	0.2	7.4	8.0
Maharashtra	18	0.0	0.9	0.0	0.1	0.1	0.0	0.0	0.1	0.3	2.5	0.1	0.1	0.1	2.9	0.3	0.0	1.5	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.2	0.5	0.0	0.3	0.0	1.9	0.2	12.4	8.9
Manipur	19	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
Meghalaya	20	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2
Mizoram	21	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Nagaland	22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Orissa	23	0.0	0.4	0.0	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	1.0	3.4	4.0
Pondicherry	24	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.8	0.1
Punjab	25	0.0	0.0	0.0	0.1	0.3	0.1	0.0	0.2	0.0	0.1	1.6	0.6	0.2	0.0	0.0	0.0	0.2	0.1	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.6	0.0	0.0	0.0	1.3	0.1	5.7	2.6
Rajasthan	26	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.4	0.0	0.7	1.8	0.2	0.1	0.1	0.1	0.0	1.7	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.1	0.0	0.0	1.5	0.2	8.1	5.0
Sikkim	27	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
TamilNadu	28	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.8	1.3	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.1	0.0	0.0	0.0	0.0	0.1	4.5	7.2
Tripura	29	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3
Uttar Pradesh	30	0.0	0.1	0.0	0.2	2.0	0.1	0.0	1.2	0.0	0.2	1.0	0.2	0.1	0.1	0.0	0.0	2.0	0.7	0.0	0.0	0.0	0.0	0.1	0.0	0.8	1.3	0.0	0.1	0.2	0.0	0.8	11.5	16.5
West Bengal	31	0.0	0.2	0.1	0.5	5.4	0.0	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.2	0.0	0.1	0.1	1.4	0.0	9.5	8.2
	Total	0.2	4.9	0.2	1.8	10.1	0.4	0.0	3.2	0.4	4.4	6.9	1.7	1.0	6.1	3.0	0.0	7.1	6.9	0.2	0.3	0.1	0.3	1.8	0.5	5.1	7.8	0.2	5.1	0.5	15.8	4.0	100	100
State Size (NSS sample)		0.0	8.2	0.1	2.2	10.1	0.1	0.0	0.9	0.2	4.7	2.0	0.6	0.9	5.4	3.4	0.0	8.0	8.9	0.2	0.2	0.1	0.0	4.0	0.1	2.6	5.0	0.0	7.2	0.3	16.5	8.2	100	

Note: Each cell of the table shows the proportion of all migrants households that moved from a particular origin state (the columns) to a particular destination state (the rows). The State Size row and column reports the proportion of all households currently living in each state at the time of the surveys. All proportions use the household weights provided by the NSS.

Table III: Differences between migrant and non-migrant households

	(1) Mean (full sample)	(2) Migrant difference (full sample)	(3) Migrant difference (wife moved for marriage sample)
Log caloric intake per person per day	7.6286	0.0133***	0.0034
	(0.3712)	(0.0049)	(0.0080)
Log household per capita expenditure	4.8272	0.0624***	0.0423***
(Rupees, 30 days)	(0.5816)	(0.0069)	(0.0108)
Log food expenditure	4.4205	0.0460***	0.0272***
(Rupees, 30 days)	(0.5045)	(0.0059)	(0.0094)
Log household size	1.7522	-0.0314***	0.0001
	(0.4848)	(0.0068)	(0.0104)
Proportion males 0-4	0.0677	0.0025	0.0034
	(0.1043)	(0.0015)	(0.0026)
Proportion females 0-4	0.0634	0.0021	0.0044*
	(0.1020)	(0.0014)	(0.0024)
Proportion males 5-9	0.0722	-0.0007	-0.0023
	(0.1055)	(0.0014)	(0.0024)
Proportion females 5-9	0.0656	-0.0001	-0.0002
	(0.1004)	(0.0015)	(0.0025)
Proportion males 10-14	0.0666	-0.0007	-0.0030
	(0.1052)	(0.0014)	(0.0024)
Proportion females 10-14	0.0575	-0.0006	-0.0027
	(0.0962)	(0.0013)	(0.0022)
Proportion males 15-55	0.2701	0.0113***	0.0023
	(0.1596)	(0.0023)	(0.0030)
Proportion females 15-55	0.2625	-0.0089***	-0.0016
	(0.1339)	(0.0017)	(0.0026)
Proportion males over 55	0.0375	-0.0013	-0.0007
	(0.0875)	(0.0011)	(0.0017)
Proportion females over 55	0.0370	-0.0036***	0.0005
	(0.0917)	(0.0010)	(0.0016)
Rural self-employed in non-agriculture	0.0972	0.0036	0.0041
	(0.2963)	(0.0038)	(0.0068)
Urban self-employed	0.0853	0.0212***	0.0158***
	(0.2793)	(0.0043)	(0.0060)
Rural agricultural labor	0.2151	-0.0110**	-0.0160*
	(0.4109)	(0.0044)	(0.0085)
Urban wage earner	0.0513	-0.0054*	-0.0022
	(0.2207)	(0.0031)	(0.0042)
Rural other labor	0.0553	0.0069**	-0.0006
	(0.2286)	(0.0028)	(0.0046)
Urban casual labor	0.0143	-0.0037**	-0.0010
	(0.1188)	(0.0018)	(0.0023)
Rural self-employed in agriculture	0.3464	-0.0075	-0.0016
	(0.4758)	(0.0049)	(0.0097)
Rural other	0.0643	0.0080**	0.0141***
	(0.2453)	(0.0031)	(0.0049)
Urban other	0.0707	-0.0120***	-0.0126***
	(0.2564)	(0.0033)	(0.0046)
Log price paid (no controls)	0.6458	0.00341***	-0.00033
(Rupees per 1000 calories)	(1.0874)	(0.00115)	(0.00204)
Log price paid (food expenditure controls)	0.6451	0.00154	-0.00137
(Rupees per 1000 calories)	(1.0876)	(0.00113)	(0.00202)
Log price paid (controls)	0.6447	0.00102	-0.00184
(Rupees per 1000 calories)	(1.0874)	(0.00112)	(0.00202)

Note: Column 1 shows the mean of each household-level variable in the row title. Column 2 shows the coefficient on a migrant dummy when the variable is regressed on a village-round fixed effect and a migrant-status dummy. Column 3 shows the coefficient for the same regression but restricting attention to households in the "*wife moved for marriage sample*" described in the text. The last three rows show the coefficient on a migrant dummy from a regression of log unit values for every product purchased on product-village-round fixed effects and a migrant-status dummy. The "controls" row includes the same vector of controls for log per-capita food expenditure and household characteristics used in all later regressions, while the "food expenditure controls" row just uses a cubic in log household food expenditure per capita. Regressions are weighted using household weights except the last three rows that use budget shares interacted with household weights. Regressions clustered at the level of the fixed effects. * signifies significance at the 10 percent level, ** at the 5 percent level and *** at the 1 percent level.

Table IV: Comparing the Caloric Intake of Migrants and Non-Migrants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Dependent variable: Log caloric intake (per person per day)							Dependent variable: Log calories per Rupee of food expenditure	
Specification:	Baseline Specification	Wife moved for marriage sample	Total expenditure controls	Real food expenditure controls	State-specific food expenditure controls	Food expenditure instrumented	PDS foods (rice, wheat and sugar) excluded	Baseline Specification	Priced at village median prices
migrant $_i$	-0.0159***	-0.0121***	-0.0136***	-0.0117***	-0.0168***	-0.0118***	-0.0205***	-0.0145***	-0.0149***
	(0.00326)	(0.00356)	(0.00442)	(0.00418)	(0.00316)	(0.00332)	(0.00505)	(0.00322)	(0.00351)
Observations	235,126	91,406	235,104	235,122	235,126	234,961	237,328	235,104	235,104
Within R-squared	0.732	0.718	0.520	0.676	0.723	0.711	0.729	0.325	0.306
Food Expenditure Controls	Yes	Yes	No	No	No	Yes	Yes	No	No
Total Expenditure Controls	No	No	Yes	No	No	No	No	Yes	Yes
Real Food Expenditure Controls	No	No	No	Yes	No	No	No	No	No
State-Specific Food Expenditure Controls	No	No	No	No	Yes	No	No	No	No
Demographics/Household Type Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Dependent variable for columns 1-7 is log caloric intake per person per day. Dependent variable for column 8 is log caloric intake per Rupee of actual food expenditure and for column 9 is log caloric intake per Rupee of food expenditure where calories are priced at the village median price for each food. Independent variable $migrant_i$ is a dummy for whether the household head or their spouse is an interstate migrant. Column 2 restricts attention to households in which the wife of the household head moved village at the time of marriage and compares wives who moved interstate (migrants) to those who moved intrastate (non-migrants). All specifications include village-round fixed effects and flexible survey-round-specific controls for household size, demographics and type as well as a third-order polynomial in log per-capita food expenditure (bar columns 3 to 5 and columns 8 to 9 which use alternate expenditure controls). Column 6 instruments the polynomial in total food expenditure with a polynomial in total non-food expenditure. Column 7 excludes the foods commonly sold through the Public Distribution System. All regressions are weighted using household weights and the standard errors are two-way clustered at the village-round and origin-state level. * signifies significance at the 10 percent level, ** at the 5 percent level and *** at the 1 percent level.

Table V: Comparing the Caloric Intake of Migrants and Non-Migrants: Subpopulations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable: Log caloric intake (per person per day)							
Sample:	Baseline Specification	Consuming < 1850 calories per day	Consuming < 2000 calories per day	Consuming < 2400 (rural) 2100 (urban)	Below median per capita expenditure	Bottom quartile per capita expenditure	Below median per capita food expend.	Bottom quartile per capita food expend.
$migrant_i$	-0.0159***	-0.0120**	-0.0145***	-0.0158***	-0.0107**	-0.0141	-0.0147**	-0.00981
	(0.00326)	(0.00495)	(0.00440)	(0.00414)	(0.00471)	(0.00865)	(0.00581)	(0.00703)
Observations	235,126	67,116	89,886	131,149	87,010	36,436	86,531	36,921
Within R-squared	0.732	0.625	0.626	0.640	0.674	0.689	0.643	0.633
Food Expenditure Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics/Household Type Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Dependent variable: Log caloric intake (per person per day)							
Sample:	Children under 5 in household	Children under 5 and below median food expend.	Children under 16 in household	Children under 16 and below median food expend.	Migrated 5 or more years ago	Migrated 10 or more years ago	Migrated 20 or more years ago	Literate
$migrant_i$	-0.0132***	-0.0141**	-0.0169***	-0.0140***	-0.0114***	-0.00650**	0.00183	-0.0171***
	(0.00461)	(0.00715)	(0.00330)	(0.00533)	(0.00335)	(0.00323)	(0.00421)	(0.00401)
Observations	103,652	47,881	178,456	75,644	231,268	228,442	223,489	128,706
Within R-squared	0.700	0.630	0.713	0.637	0.732	0.732	0.732	0.732
Food Expenditure Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics/Household Type Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Dependent variable is log caloric intake per person per day. Independent variable $migrant_i$ is a dummy for whether the household head or their spouse is an interstate migrant. All specifications include village-round fixed effects and flexible survey-round-specific controls for household size, demographics and type as well as a third-order polynomial in log per-capita food expenditure. The column headings denote the various subpopulations on which the regressions are run. All regressions are weighted using household weights and the standard errors are two-way clustered at the village-round and origin-state level. * signifies significance at the 10 percent level, ** at the 5 percent level and *** at the 1 percent level.

Table VI: Comparing Bundles of Migrants and Non-Migrants

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Rice expenditure share $rice_i/(rice_i + wheat_i)$		Correlation ρ_i^d of household i budget shares with budget shares of current state d	Correlation $\rho_i^{o_v}$ of household i budget shares with the average budget shares of the origin states o_v of migrants in village (169 foods)		
	Migrant Households Only	All Households		Baseline Specification	Wife moved for marriage sample	Total expenditure controls
$rice_i^o/(rice_i^o + wheat_i^o)$	0.189***	0.123***				
	(0.0204)	(0.0146)				
$migrant_i$			-0.0111***			
			(0.00199)			
$migrant_i^{o_v}$				0.0226***	0.0103***	0.0225***
				(0.00290)	(0.00284)	(0.00293)
Observations	14,156	226,472	235,126	108,743	23,397	108,726
Within R-squared	0.065	0.007	0.080	0.098	0.062	0.074
Food Expenditure Controls	Yes	Yes	Yes	Yes	Yes	No
Alternative Expnediture Controls	No	No	No	No	No	Yes
Demographics/Household Type	Yes	Yes	Yes	Yes	Yes	Yes
Village-Round FE	Yes	Yes	Yes	No	No	No
Village-o_v-Round FE	No	No	No	Yes	Yes	Yes

	(7)	(8)	(9)	(10)	(11)	(12)
Dependent variable:	Correlation $\rho_i^{o_v}$ of household i budget shares with the average budget shares of the origin states o_v					Correlation of household caloric shares with caloric shares of origin states o_v
	Real food expenditure controls	State-specific food expenditure controls	PDS foods (rice, wheat and sugar) excluded	Reference basket by income quartile	Interactions with ln$K$ difference	
$migrant_i^{o_v}$	0.0215***	0.0223***	0.0171***	0.0217***	0.0272***	0.0263***
	(0.00299)	(0.00296)	(0.00287)	(0.00245)	(0.00382)	(0.00334)
$migrant_i^{o_v}$ × $\mathbf{1}[\ln K(\boldsymbol{bshare}_i^o, \boldsymbol{P}_i^v) < \ln K(\boldsymbol{bshare}_i^v, \boldsymbol{P}_i^v)]$					-0.0147**	
					(0.00618)	
Observations	108,737	108,743	108,696	108,638	105,810	108,743
Within R-squared	0.135	0.066	0.104	0.092	0.096	0.060
Food Expenditure Controls	No	No	Yes	Yes	Yes	Yes
Alternative Expnediture Controls	Yes	Yes	No	No	No	No
Demographics/Household Type	Yes	Yes	Yes	Yes	Yes	Yes
Village-o_v-Round FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Columns 1 and 2 regress the share of household rice and wheat expenditures spent on rice on the average rice share of the household's origin state (with column 1 restricting attention only to migrants). Column 3 regresses ρ_i^d, the correlation between household i's vector of food budget shares and the vector of mean budget shares of non-migrant households in household i's current state d, on a migrant dummy. Dependent variable in columns 4 to 11 is the correlation between household i's vector of food budget shares and the vector of mean budget shares for non-migrant households in state o_v, where state o_v is the origin state of a migrant in the household's village. The independent variable is $migrant_i^{o_v}$ (a dummy for whether the household head or their spouse is an interstate migrant from o_v). All specifications include either village-round or village-o_v-round fixed effects and flexible survey-round-specific controls for household size, demographics and type as well as a third-order polynomial in log per-capita food expenditure (bar columns 6 to 8 which use alternate expenditure controls). Column 5 restricts attention to households with wives who moved at the time of marriage, column 9 excludes the foods commonly sold through the Public Distribution System, column 10 matches households to a reference basket calculated separately by income-quartile, column 11 includes an additional interaction with an indicator variable that takes the value of 1 if the migrant's origin-state bundle provide fewer calories than the local bundle, and column 12 uses caloric shares in lieu of budget shares for the dependent variable. All regressions are weighted using household weights and the standard errors are two-way clustered at the village-round and origin-state level. * signifies significance at the 10 percent level, ** at the 5 percent level and *** at the 1 percent level.

Table VII: Comparing Bundles of Migrants and Non-Migrants: Subpopulations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable: Correlation $\rho_i^{o_v}$ of household i budget shares with the average budget shares of the origin states o_v of migrants in village (169 foods)							
Sample:	Baseline Specification	Consuming < 1850 calories per day	Consuming < 2000 calories per day	Consuming < 2400 (rural) 2100 (urban)	Below median per capita expenditure	Bottom quartile per capita expenditure	Below median per capita food expend.	Bottom quartile per capita food expend.
$migrant_i^{o_v}$	0.0226***	0.0196***	0.0213***	0.0213***	0.00646**	0.00142	0.00900**	0.00312
	(0.00290)	(0.00367)	(0.00348)	(0.00352)	(0.00314)	(0.00694)	(0.00392)	(0.00688)
Observations	108,743	28,560	38,666	50,925	21,179	7,326	23,165	8,445
Within R-squared	0.098	0.121	0.115	0.105	0.090	0.100	0.102	0.104
Food Expenditure Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics/Household Type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-o_v-Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Dependent variable: Correlation $\rho_i^{o_v}$ of household i budget shares with the average budget shares of the origin states o_v of migrants in village (169 foods)							
Sample:	Children under 5 in household	Children under 5 and below median food expend.	Children under 16 in household	Children under 16 and below median food expend.	Migrated 5 or more years ago	Migrated 10 or more years ago	Migrated 20 or more years ago	Literate
$migrant_i^{o_v}$	0.0183***	0.0114**	0.0191***	0.00875**	0.0216***	0.0209***	0.0187***	0.0257***
	(0.00359)	(0.00469)	(0.00273)	(0.00407)	(0.00275)	(0.00283)	(0.00385)	(0.00320)
Observations	41,718	12,628	76,074	20,407	99,383	93,185	82,590	76,684
Within R-squared	0.073	0.113	0.066	0.093	0.093	0.093	0.092	0.125
Food Expenditure Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics/Household Type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-o_v-Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Dependent variable is the correlation between household i's vector of food budget shares and the vector of mean budget shares for non-migrant households in state o_v, where state o_v is the origin-state of a migrant in the household's village. The independent variable is $migrant_i^{o_v}$ (a dummy for whether the household head or their spouse is an interstate migrant from o_v). All specifications include village-o_v-round fixed effects and flexible survey-round-specific controls for household size, demographics and type as well as a third-order polynomial in log per-capita food expenditure. The column headings denote the various subpopulations on which the regressions are run. All regressions are weighted using household weights and the standard errors are two-way clustered at the village-round and origin-state level. * signifies significance at the 10 percent level, ** at the 5 percent level and *** at the 1 percent level.

Table VIII: Results broken-down by within-household migrant structure

Full sample of households		Variable interacted with migrant-structure dummies	Only head or spouse is migrant	Both migrants	No spouse	Observations	Within R-squared
(1)	Dependent variable: Correlation $\rho_i^{o_v}$	$migrant_i^{o_v}$	0.00786***	0.0416***	0.0331***		
			(0.00251)	(0.00453)	(0.00819)	108,743	0.100
(2)	Dependent variable: Log caloric intake	$migrant_i$	-0.0125***	-0.0228***	-0.0139**		
			(0.00292)	(0.00570)	(0.00618)	235,126	0.732
(3)	Dependent variable: Log caloric intake	$migrant_i$	-0.00351	-0.00886	0.00246		
			(0.00374)	(0.00675)	(0.00675)		
		$migrant_i$	-0.0188***	-0.0628***	-0.0487***		
		$\times \mathbf{1}[\ln K(\boldsymbol{bshare}_i^o, \boldsymbol{P}_i^v) < \ln K(\boldsymbol{bshare}_i^v, \boldsymbol{P}_i^v)]$	(0.00521)	(0.0144)	(0.0124)	234,155	0.732
Subsample of households consuming less than 2000 calories per person day		Variable interacted with migrant-structure dummies	Only head or spouse is migrant	Both migrants	No spouse	Observations	Within R-squared
(4)	Dependent variable: Correlation $\rho_i^{o_v}$	$migrant_i^{o_v}$	0.00778*	0.0375***	0.0348***		
			(0.00398)	(0.00610)	(0.00876)	38,666	0.117
(5)	Dependent variable: Log caloric intake	$migrant_i$	-0.0123***	-0.0151**	-0.0286***		
			(0.00470)	(0.00716)	(0.00865)	89,886	0.626
(6)	Dependent variable: Log caloric intake	$migrant_i$	-0.00252	-0.00447	-0.0101		
			(0.00687)	(0.00721)	(0.0101)		
		$migrant_i$	-0.0247**	-0.0494***	-0.0517**		
		$\times \mathbf{1}[\ln K(\boldsymbol{bshare}_i^o, \boldsymbol{P}_i^v) < \ln K(\boldsymbol{bshare}_i^v, \boldsymbol{P}_i^v)]$	(0.0101)	(0.0170)	(0.0228)	89,596	0.626

Note: Panels 1-6 repeat the regressions shown in Tables IV, VI and IX but interacting every instance of a migrant dummy variable with indicator variables for three mutually-exclusive categories of within-household migrant structure: only one of head or spouse is a migrant, both head and spouse are migrants, and there is no spouse. Each panel comprises one regression. Panels 4-6 restrict attention to households consuming fewer than 2000 calories per person per day. As in Table VI, Panels 1 and 4 include village-o_v-round fixed effects. As in Tables IV and IX, Panels 2-3 and 5-6 include village-round fixed effects. All panels include flexible survey-round-specific controls for household size, demographics and type as well as a third-order polynomial in log per-capita food expenditure and a dummy for households where the head has no spouse. All regressions are weighted using household weights and the standard errors are two-way clustered at the village-round and origin-state level. * signifies significance at the 10 percent level, ** at the 5 percent level and *** at the 1 percent level.

Table IX: Comparing the Caloric Intake of Migrants and Non-Migrants across Migration Routes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	Log caloric intake (per person per day)					Log calories per Rupee of food expenditure		
Specification:	Baseline Specification	Quartiles of lnK difference	lnK difference interaction	PDS foods excluded	PDS foods excluded	Baseline Specification	Prices at Village Median Prices	Prices at Village Median Prices
$migrant_i$	-0.00498			0.00611		-0.00369	-0.00010	
	(0.00433)			(0.00446)		(0.00440)	(0.00417)	
$migrant_i$	-0.0283***			-0.0823***		-0.0280***	-0.0376***	
$\times \mathbf{1}[\ln K(\boldsymbol{bshare}^o_i, \boldsymbol{P}^v_i) < \ln K(\boldsymbol{bshare}^v_i, \boldsymbol{P}^v_i)]$	(0.00607)			(0.0125)		(0.00645)	(0.00647)	
$migrant_i$		-0.0393***			-0.0820***			-0.0435***
$\times$ *1st quartile* $[\ln K(.,.)\text{-}\ln K(.,.)]$		(0.00555)			(0.0117)			(0.00766)
$migrant_i$		-0.0191***			-0.0303***			-0.0208***
$\times$ *2nd quartile* $[\ln K(.,.)\text{-}\ln K(.,.)]$		(0.00483)			(0.0105)			(0.00443)
$migrant_i$		-0.0136***			0.00708			-0.00736
$\times$ *3rd quartile* $[\ln K(.,.)\text{-}\ln K(.,.)]$		(0.00516)			(0.00662)			(0.00539)
$migrant_i$		0.0227***			0.0366***			0.0303***
$\times$ *4th quartile* $[\ln K(.,.)\text{-}\ln K(.,.)]$		(0.00575)			(0.00656)			(0.00574)
$migrant_i$			0.0629***					
$\times (\ln K(\boldsymbol{bshare}^o_i, \boldsymbol{P}^v_i) - \ln K(\boldsymbol{bshare}^v_i, \boldsymbol{P}^v_i))$			(0.0115)					
Observations	234,157	234,157	234,157	236,323	236,323	234,135	234,135	234,135
Within R-squared	0.731	0.732	0.731	0.729	0.729	0.325	0.307	0.307
Food Expenditure Controls	Yes	Yes	Yes	Yes	Yes	No	No	No
Alternative Expnediture Controls	No	No	No	No	No	Yes	Yes	Yes
Demographics/Household Type Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Dependent variable:	Log caloric intake (per person per day)							
Specification:	Wife moved for marriage sample	Total expenditure controls	Real food expenditure controls	State-specific food expenditure controls	Food expenditure instrumented	Reference basket by income quartile	Log migrant travel distance interactions	Reference basket by destination state
$migrant_i$	-0.00117	-0.00260	-0.00303	-0.00572	-0.00242	-0.00746	0.0171	-0.00879*
	(0.00514)	(0.00510)	(0.00579)	(0.00436)	(0.00437)	(0.00475)	(0.0218)	(0.00533)
$migrant_i$	-0.0220***	-0.0285***	-0.0222***	-0.0289***	-0.0241***	-0.0246***	-0.0267***	-0.0136**
$\times \mathbf{1}[\ln K(\boldsymbol{bshare}^o_i, \boldsymbol{P}^v_i) < \ln K(\boldsymbol{bshare}^v_i, \boldsymbol{P}^v_i)]$	(0.00763)	(0.00643)	(0.00680)	(0.00622)	(0.00583)	(0.00658)	(0.00629)	(0.00643)
$migrant_i$							-0.00413	
$\times \ln distance_{ov}$							(0.00382)	
Observations	91,158	234,135	234,153	234,157	233,993	231,774	213,444	235,004
Within R-squared	0.718	0.521	0.676	0.723	0.711	0.731	0.732	0.732
Food Expenditure Controls	Yes	No	No	No	Yes	Yes	Yes	Yes
Alternative Expnediture Controls	No	Yes	Yes	Yes	No	No	No	No
Demographics/Household Type Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Dependent variable is log caloric intake per person per day (except columns 6 to 8 where it is log caloric intake per Rupee of food expenditure, with calories priced at village median prices in the latter two cases). Independent variables are $migrant_i$, a migrant household dummy, and $migrant_i$ interacted with with an indicator variable that equals 1 if a 1 Rupee reference bundle from the migrant's origin state provides fewer calories than a 1 Rupee reference bundle from the migrant's destination village (both priced at destination-village prices). All specifications include village-round fixed effects and flexible survey-round-specific controls for household size, demographics and type as well as a third-order polynomial in log per-capita food expenditure (except columns 6 to 8 and columns 10 to 12 which use alternate expenditure controls). Column 1 reports the baseline specification. Columns 2 and 3 use alternative functions of these calorie differences. Columns 4 and 5 exclude the foods commonly sold through the Public Distribution System. Column 9 restricts attention to households in which the wife moved village at the time of marriage and compares interstate to intrastate movers. Column 13 instruments total food expenditure with total non-food expenditure. Column 14 matches households to a reference basket calculated separately by income-quartile. Column 15 includes an interaction with the distance between a migrant's destination region and their origin state. Column 16 uses a state-specific reference bundle for the local bundle. All regressions are weighted using household weights and the standard errors are two-way clustered at the village-round and origin-state level. * signifies significance at the 10 percent level, ** at the 5 percent level and *** at the 1 percent level.

Table X: Comparing the Caloric Intake of Migrants and Non-Migrants across Migration Routes: Subpopulations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable: Log caloric intake (per person per day)							
Sample:	Baseline Specification	Consuming < 1850 calories per day	Consuming < 2000 calories per day	Consuming < 2400 (rural) 2100 (urban)	Below median per capita expenditure	Bottom quartile per capita expenditure	Below median per capita food expend.	Bottom quartile per capita food expend.
$migrant_i$	-0.00498	-0.000922	-0.00412	-0.00560	0.00347	0.0281	-0.000500	0.0242
	(0.00433)	(0.00676)	(0.00578)	(0.00547)	(0.00754)	(0.0194)	(0.00910)	(0.0175)
$migrant_i$	-0.0283***	-0.0329***	-0.0304***	-0.0249***	-0.0222**	-0.0547***	-0.0220**	-0.0449**
$\times \mathbf{1}[\ln K(\boldsymbol{bshare}_i^o, \boldsymbol{P}_i^v) < \ln K(\boldsymbol{bshare}_i^v, \boldsymbol{P}_i^v)]$	(0.00607)	(0.0103)	(0.00796)	(0.00696)	(0.00908)	(0.0189)	(0.00868)	(0.0188)
Observations	234,157	66,899	89,596	130,770	86,877	36,390	86,378	36,869
Within R-squared	0.731	0.624	0.626	0.640	0.674	0.689	0.643	0.633
Food Expenditure Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics/Household Type Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Dependent variable: Log caloric intake (per person per day)							
Sample:	Children under 5 in household	Children under 5 and below median food expend.	Children under 16 in household	Children under 16 and below median food expend.	Migrated 5 or more years ago	Migrated 10 or more years ago	Migrated 20 or more years ago	Literate
$migrant_i$	0.00290	0.00296	-0.00588	0.000205	-0.00100	0.00454	0.0149***	-0.0108**
	(0.00597)	(0.0116)	(0.00403)	(0.00829)	(0.00412)	(0.00395)	(0.00458)	(0.00526)
$migrant_i$	-0.0364***	-0.0266**	-0.0270***	-0.0217**	-0.0264***	-0.0276***	-0.0328***	-0.0223***
$\times \mathbf{1}[\ln K(\boldsymbol{bshare}_i^o, \boldsymbol{P}_i^v) < \ln K(\boldsymbol{bshare}_i^v, \boldsymbol{P}_i^v)]$	(0.00811)	(0.0130)	(0.00589)	(0.00889)	(0.00639)	(0.00644)	(0.00882)	(0.00848)
Observations	103,325	47,784	177,897	75,507	230,547	227,894	223,220	128,026
Within R-squared	0.699	0.629	0.713	0.637	0.732	0.732	0.732	0.731
Food Expenditure Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics/Household Type Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village-Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Dependent variable for all columns is log caloric intake per person per day. Independent variables are $migrant_i$, a dummy for whether the household head or their spouse is an interstate migrant, and $migrant_i$ interacted with with an indicator variable that takes the value of 1 if a 1 Rupee reference bundle from the migrant's origin-state provides fewer calories than a 1 Rupee reference bundle from the migrant's destination village both priced at destination-village prices. All specifications include village-round fixed effects and flexible survey-round-specific controls for household size, demographics and type as well as a third-order polynomial in log per-capita food expenditure. The column headings denote the various subpopulations on which the regressions are run. All regressions are weighted using household weights and the standard errors are two-way clustered at the village-round and origin-state level. * signifies significance at the 10 percent level, ** at the 5 percent level and *** at the 1 percent level.

Table XI: Results broken-down by whether husband or wife is a migrant

Full sample of households		Variable interacted with migrant-structure dummies	Only head is migrant	Only spouse is migrant	Both migrants	No spouse	Observations	Within R-squared
(1)	Dependent variable: Correlation $\rho_i^{o_v}$	$migrant_i^{o_v}$	0.00982**	0.00715***	0.0416***	0.0331***		
			(0.00459)	(0.00267)	(0.00454)	(0.00820)	108,743	0.100
(2)	Dependent variable: Log caloric intake	$migrant_i$	-0.0186***	-0.0103***	-0.0230***	-0.0141**		
			(0.00479)	(0.00323)	(0.00570)	(0.00619)	235,126	0.732
(3)	Dependent variable: Log caloric intake	$migrant_i$	-0.00753	-0.00189	-0.00901	0.00233		
			(0.00599)	(0.00427)	(0.00677)	(0.00677)		
		$migrant_i$	-0.0253***	-0.0173***	-0.0628***	-0.0487***		
		$\times \mathbf{1}[\ln K(\boldsymbol{bshare}_i^o, \boldsymbol{P}_i^v) < \ln K(\boldsymbol{bshare}_i^v, \boldsymbol{P}_i^v)]$	(0.00920)	(0.00608)	(0.0144)	(0.0124)	234,155	0.732
Full sample of households		Variable interacted with migrant-structure dummies	Only head is migrant and no mother in house	Only spouse is migrant	Both migrants or migrant head with mother in house	No spouse	Observations	Within R-squared
(4)	Dependent variable: Correlation $\rho_i^{o_v}$	$migrant_i^{o_v}$	0.0123**	0.00713***	0.0388***	0.0330***		
			(0.00483)	(0.00267)	(0.00486)	(0.00819)	108,743	0.100
(5)	Dependent variable: Log caloric intake	$migrant_i$	-0.0141***	-0.0103***	-0.0244***	-0.0142**		
			(0.00476)	(0.00323)	(0.00549)	(0.00617)	235,126	0.732
(6)	Dependent variable: Log caloric intake	$migrant_i$	-0.00340	-0.00190	-0.0101	0.00232		
			(0.00596)	(0.00427)	(0.00665)	(0.00676)		
		$migrant_i$	-0.0236***	-0.0172***	-0.0613***	-0.0486***		
		$\times \mathbf{1}[\ln K(\boldsymbol{bshare}_i^o, \boldsymbol{P}_i^v) < \ln K(\boldsymbol{bshare}_i^v, \boldsymbol{P}_i^v)]$	(0.00826)	(0.00608)	(0.0135)	(0.0124)	234,155	0.732

Note: Panels 1-3 repeat the regressions shown in Tables IV, VI and IX but interacting every instance of a migrant dummy variable with indicator variables for four mutually-exclusive categories of within-household migrant structure: only the head is a migrant, only the spouse is a migrant, both head and spouse are migrants, and there is no spouse. Each panel comprises one regression. Panel 1 includes village-o_v-round fixed effects, and Panel 2 and 3 include village-round fixed effects. Panels 4-6 repeat the same exercise but replace the only-head indicator with an indicator for households where only the head is a migrant and his mother does not live in the house (with households where the mother is present placed in the both category). All panels include flexible survey-round-specific controls for household size, demographics and type as well as a third-order polynomial in log per-capita food expenditure and a dummy for households where the head has no spouse. All regressions are weighted using household weights and the standard errors are two-way clustered at the village-round and origin-state level. * signifies significance at the 10 percent level, ** at the 5 percent level and *** at the 1 percent level.