

MIT Open Access Articles

The Caloric Costs of Culture: Evidence from Indian Migrants

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Atkin, David. "The Caloric Costs of Culture: Evidence from Indian Migrants." *American Economic Review* 106, no. 4 (April 2016): 1144–1181. © 2016 American Economic Association

As Published: <http://dx.doi.org/10.1257/aer.20140297>

Publisher: American Economic Association

Persistent URL: <http://hdl.handle.net/1721.1/102451>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



The Caloric Costs of Culture: Evidence from Indian Migrants[†]

By DAVID ATKIN*

Anthropologists have documented substantial and persistent differences in food preferences across social groups. My paper asks whether such food cultures can constrain caloric intake? I first document that interstate migrants within India consume fewer calories per rupee of food expenditure compared to their neighbors. Second, I show that migrants bring their origin-state food preferences with them. Third, I link these findings by showing that the gap in caloric intake between locals and migrants depends on the suitability and intensity of the migrants' origin-state preferences. The most affected migrants would consume seven percent more calories if they possessed their neighbors' preferences. (JEL D12, I12, O15, R23, Z12, Z13)

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 AD. Consistent with the cultural economics literature, I will describe differences in food preferences across groups as different *food cultures*.¹ One natural question to ask is whether such food cultures can 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 Famine Inquiry Commission. The commission was established in the aftermath of the 1943 Bengal Famine in which between 1.5 and

*Department of Economics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, and NBER (e-mail: atkin@mit.edu). Thanks to Keith Chen, Angus Deaton, Adrian de Froment, 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. The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

[†]Go to <http://dx.doi.org/10.1257/aer.20140297> to visit the article page for additional materials and author disclosure statement.

¹Since the types of food a group of people traditionally consumes embody the preferences, beliefs, and social attitudes of the group, this definition accords the existing definitions of culture in the economics literature (see Section IC for further details).

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. 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 nonmigrant neighbors who face the same prices but possess different cultural food preferences. This methodology allows me to broadly quantify the "costs" that culture can impose.

The interpretation of this cost depends on whether households are knowingly trading-off calories for the utility gained from cultural practice. If they are, then this cost is a measure of the strength of their cultural preferences. I provide support for this interpretation by showing that migrants moderate their consumption choices in contexts where consuming their favored foods is most disadvantageous. Alternatively, if households are unaware or undervalue the (often later-life) health consequences of low caloric intake, this cost may correspond to a welfare cost of culture. Of course, both interpretations require that caloric costs correspond to nutritional costs, an issue I discuss in depth and provide direct evidence for.

The key observation for my empirical strategy is that migrants face the same prices as their neighbors but bring their cultural food preferences with them when they migrate. To use that observation to analyze the caloric costs of culture, I require a dataset that has: (i) information on household consumption across many food products; (ii) matched to the migrant status of household members including their origin locations; (iii) with a large sample of migrants and nonmigrants across many finely-specified locations; and (iv) in a country with both malnourishment and diverse regional food preferences.

The 1983 and 1987–1988 Indian National Sample Surveys (NSS) satisfy all four of these criteria. The surveys record household purchases of 169 different food products alongside migration particulars. Over 240,000 households are surveyed in groups of 10 drawn from blocks of no more than 180 neighboring households within a village or city: ensuring that migrant and nonmigrant households face very similar prices, an assumption I test explicitly. The surveys cover all states of India, a country with many different food cultures across religious, caste, and ethnolinguistic groups. (In contrast, later Indian survey rounds either ask households about consumption or migrations but not both, while the equivalent surveys in equally large and diverse

China are neither publicly available nor do they record the required migration particulars.) And finally, I show that at the time of the surveys, 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 setting, it is important to understand why.

The analysis proceeds in three stages. Finding that migrants consume fewer calories than otherwise-similar locals provides the first hint that households are willing to forgo calories to accommodate their cultural preferences. The first stage of the analysis presents this finding: migrant households consume fewer calories per person compared to nonmigrant 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 villages 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). I also find no evidence that the caloric tax is restricted to well-nourished households for whom reductions in caloric intake may have no nutritional consequences.

Of course, migrants may consume fewer calories per rupee of food expenditure for reasons unrelated to their food cultures. And the small magnitudes do not imply that cultural preferences for food can only have limited impacts: the caloric tax averages over two types of migrant; those whose preferences are well-suited to the local price-vector and enjoy a caloric dividend, and those whose preferences are unsuited and pay a substantial caloric tax. The second stage of the analysis addresses both these issues.

In the second stage, I form a chain of evidence showing that migrants are making calorically 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 or even other migrant households in the same village hailing from different states, 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 7.2 percent. The caloric tax for this group remains a substantial 5.3 percent when I restrict my analysis to undernourished households. These are substantial magnitudes. For example, extrapolating from Schofield's (2014) estimates of the productivity decline among Indian farmers due to caloric reductions among Muslims during Ramadan, the caloric tax for these migrant groups would reduce agricultural productivity by between 4 and 10 percent.

Finally, the third 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 without migrants having different cultural preferences. Under these alternative explanations, the caloric tax should not persist many years after migration or be present if other household members are familiar with the local foods. I find no evidence for either hypothesis. 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 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 nonmigrants 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.² And, of course, the magnitudes I find are specific to the context of migrants within India (and potentially a lower bound if migrants are more adaptable).

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); Fernández (2011); and Alesina and Giuliano (2015). In using the behavior of migrants to examine the influence of culture on household decisions, it is particularly closely related to Carroll, Rhee, and Rhee's (1994) study of savings behavior; Fernández, Fogli, and Olivetti's (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.³

²For example, in the context of the United States where over-consumption of calories is the problem, recent immigrants have lower levels of obesity than locals (Goel et al. 2004). The same is not true for US-born children of immigrants, in part because they adopt less-healthy American foods to fit in (Guendelman, Cheryan, and Monin 2011).

³In this sense, the paper has similarities with Almond and Mazumder (2011) and Schofield (2014) which quantify the health and productivity impacts of fasting for the Muslim holiday of Ramadan.

Second, I add to the literature on the persistence of food preferences initiated by Staehle (1934), with recent contributions by Logan and Rhode (2010) and Bronnenberg, Dube, and Gentzkow (2012). Although these papers document that migrants bring their food preferences with them (and in the latter case, the consequences for brands' market shares) none of these papers explore the nutritional consequences and hence the costs of such preferences. Finally, this study is related to Nunn and Qian (2011). Their finding, that over hundreds of years the Old World adopted New World crops with consequent nutritional improvements, suggests that the persistent food culture I find may weaken over many generations.

Atkin (2013) provides theoretical and empirical evidence for the existence of regional food preferences in India. The two papers differ in that Atkin (2013) lays out a model in which the combination of agroclimatic endowments and habits generate regional food tastes that favor the locally-abundant foods and explores the implications of this correlation between preferences and endowments for the size of the gains generated by trade. In contrast, this paper takes India's regional food preferences as given, interprets these as cultural phenomena, and furthers our understanding of the importance of culture by quantifying the calories households are willing to forgo to accommodate their cultural preferences.

I lay out the paper as follows. Section I introduces the data and provides background on nutrition and cultural food preferences in India. Section II explains how I use migrants to identify the caloric costs of culture. Sections III and IV show that migrant households consume fewer calories than comparable nonmigrant households and that this finding is driven by migrants making calorically suboptimal food choices in order to accommodate their origin-state food preferences. Section V rules out the two noncultural explanations. Finally, Section VI discusses policy implications and concludes.

I. Data and Background on Malnutrition and Cultural Preferences for Food in India

A. Data Description

My analysis draws on two cross sections of the NSS collected by the National Sample Survey Organization (NSSO): the thirty-eighth round (1983) and the forty-third round (1987–1988). As mentioned in the introduction, these are the only 2 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.

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). 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). I stack the two cross-sections and create a combined dataset containing 240,040 households. Throughout the paper, I use the provided survey weights to make the sample nationally representative.

The surveys record household expenditures and quantities for each food item consumed in the last 30 days (with homegrown foods and gifts both valued at the

prevailing local prices).⁴ 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 (using the calorie per unit quantity estimates contained in the NSSO reports that accompany each survey round) by the quantity consumed over the previous 30 days. I use this number to calculate the daily caloric intake per household member.⁵ The surveys also provide information on expenditures on non-food items as well as household demographics and characteristics.

Turning to migration particulars, 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. (Hence, for households that migrated multiple times, the state of their last residence may not coincide with their regional preferences, a measurement issue which will tend to attenuate my results.) I define interstate 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 head and 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 interchangeably, and similarly the terms spouse and wife.

Table 1 provides summary statistics for the dataset. 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. In terms of household structure, in 41.3 percent of cases only the wife is a migrant (consistent with Indian "patrilocality" norms that I exploit later whereby, particularly in North India, wives move in with their husband's family upon marriage). 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.

B. *Malnutrition in India*

Indian households in 1983 and 1987–1988 consume a small number of calories. The mean caloric intake is 2,224 per person per day across the 2 samples. To get a sense of magnitudes, recent figures (Food and Agriculture Organization of the United Nations 2008) were 3,750 calories per person per day for the United States, 2,990 for China, and 2,360 for India.

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 2,400 calories per person

⁴The survey records expenditures of consumed foods and so includes any foods or meals purchased nonlocally if away from home or foods purchased in other periods with expenditures costed at the actual purchase price.

⁵These numbers likely overestimate actual caloric intake. Some of these calories are fed to servants, pets, and guests; or are wasted (due to spoilage or simply thrown away). Online Appendix Table C2 shows meals served and meals eaten out are similar for migrant and nonmigrant households. If wastage rates are higher for migrants as they are less familiar with local foods, my later estimates underestimate the difference in caloric intake between migrants and nonmigrants. Actual caloric absorption will be even lower if household members have ailments such as diarrhea that prevent the body from fully absorbing the calories.

TABLE 1—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 | 240,040 | 1.000 | 0.778 | 0.0608 | 0.459 |
| Wife moved for marriage sample | 93,818 | 0.487 | 0.847 | 0.0606 | 0.576 |
| | No spouse | Only spouse is migrant | Only head is migrant | Both migrants (same state) | Both migrants (different state) |
| <i>Migrant breakdown by within-household migrant structure</i> | | | | | |
| 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 |
| | 0 to 4 years | 5 to 9 years | 10 to 20 years | 20 or more years | |
| <i>Migrant breakdown by years since migration (most recent migrant in household)</i> | | | | | |
| Full sample | 0.152 | 0.141 | 0.293 | 0.413 | |
| Wife moved for marriage sample | 0.072 | 0.137 | 0.320 | 0.470 | |
| | Consuming below 2,400 calories rural/2,100 urban | Consuming below 2,150 calories rural/1,850 urban | Consuming below 1,900 calories rural/1,600 urban | Below median per capita expenditure | Below 25th percentile of per capita expenditure |
| <i>Migrant breakdown by per capita expenditure and nutrition</i> | | | | | |
| Full sample | 0.649 | 0.503 | 0.341 | 0.502 | 0.251 |
| Wife moved for marriage sample | 0.652 | 0.503 | 0.337 | 0.529 | 0.267 |

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

per day for rural India and 2,100 calories per person per day for urban India. 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 50.3 percent of households more than 250 calories per person per day below these norms and 34.1 percent more than 500 calories below these norms.⁶

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.

⁶Table 1 provides further details and online Appendix Figure A2 shows the full distribution of caloric intake.

C. Cultural Preferences for Food 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. For example, unlike Muslims and low-caste Hindus, high-caste Hindus are typically vegetarian but adherence varies dramatically by region. In online Appendix A, I review this literature and provide examples of these different food cultures.

The regional taste differences these food cultures generate fit squarely within the definitions of culture used in the economics literature. Fernández (2011, p. 484) 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 fits this definition since state boundaries were drawn primarily along major ethnolinguistic divisions. Furthermore, India’s religious minorities are often concentrated in particular states. Guiso, Sapienza, and Zingales (2006, p. 23) 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.⁷ Accordingly, this paper treats regional and cultural food preferences as synonymous.

Online Appendix A also explores one particularly important example—regional preferences for rice and wheat. Despite these two cereals providing a similar number of calories and micro-nutrients per rupee, there is dramatic regional variation in consumption. For example, the relative price of rice and wheat is similar in the states of Kerala and Punjab, yet Keralites consumed 13 times more rice than wheat and Punjabis 10 times more wheat than rice. Atkin (2013) shows how agroclimatic endowments coupled with habit formation can generate these different food cultures. 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 rice-loving states switched the quantity of rice and wheat that they were consuming (and vice versa for wheat-loving states). However, since every village has a different price vector, complex substitution patterns may rationalize this consumption behavior, necessitating the methodology I propose below.

II. Empirical Methodology: Examining the Behavior of Migrants

The behavior of interstate migrants provides more convincing evidence that culture can constrain caloric intake. I break the analysis into two sections. In Section III, I show that the average migrant consumes fewer calories per rupee than the nonmigrants around them. Then in Section IV, I link the size of this caloric tax to cultural preferences for the traditional foods of the migrants’ origin states. For the first result to be interpreted as preference-driven, I require that migrants

⁷Birch (1999) surveys the evidence from the psychology and nutrition literature which finds that adult food preferences form in part through consumption in childhood.

and nonmigrants 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). In the following paragraphs, I argue that these assumptions are likely to be satisfied. For the second result, I am essentially comparing migrants to each other (by showing that the caloric tax only appears for migrants whose origin-state bundle is unsuited to the local price vector), and in fact replicate the analysis focusing only on migrants. Therefore, even if migrants and nonmigrants do differ on unobservables, in order to explain both results any differences would have to be systematically correlated with the composition of the average food bundle consumed in the migrants' origin state.

The first reason why the assumptions above are tenable is that the NSS data allow me to make this comparison at an extremely disaggregated geographic level. In each survey round the NSSO draws a sample of around 8,000 rural villages and 4,500 urban blocks and surveys 10 households in each village/block. Any village/block with more than 1,200 inhabitants (approximately 180 households) is subdivided into smaller geographical subgroups, and only 1 subgroup is surveyed. Therefore, I can compare migrant and nonmigrant households that live in very close proximity and so are likely to shop at the same stores, particularly in rural areas where there may only be one store in the village, and face similar external environments.

Second, Indian migration patterns are not concentrated along only a few specific routes. If this were the case, the assumption that migrants and nonmigrants have the same desire for nutrition and dietary variety may be violated. For example, suppose most migrants come from Kerala and Keralites particularly value nutrition. In this scenario, if migrants (i.e., Keralites) consume fewer calories than locals (i.e., non-Keralites) I cannot infer 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 nonmigrants place an equal value on nutrition or variety). Online Appendix Table C1 displays the proportion of migrants that moved between particular origin and destination states. Unsurprisingly, the larger states in India are either the source or destination of most migrant flows. However, routes are dispersed with migrants moving from many different states and often in both directions, mitigating the above concern.

Third, migrants do face the same prices as nonmigrants, at least after controlling for observable characteristics. Table 2 compares both household characteristics and prices paid across migrant and nonmigrant households within the same village. The characteristics I focus on are the set of controls used by Subramanian and Deaton (1996) to estimate Indian caloric elasticities. 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.⁸

⁸Migrant households are also slightly smaller, contain a larger proportion of prime-age males, are less likely to be categorized as agricultural labor, and are more likely to be categorized as urban self-employed.

TABLE 2—DIFFERENCES BETWEEN MIGRANT AND NONMIGRANT HOUSEHOLDS

| | Mean (full sample) (1) | Migrant difference (full sample) (2) | Migrant difference (wife moved for marriage sample) (3) |
|---|------------------------------|--|--|
| log caloric intake per person per day | 7.6286 (0.3712) | 0.0133*** (0.0049) | 0.0034 (0.0080) |
| log household per capita expenditure (rupees, 30 days) | 4.8272 (0.5816) | 0.0624*** (0.0069) | 0.0423*** (0.0108) |
| log food expenditure (rupees, 30 days) | 4.4205 (0.5045) | 0.0460*** (0.0059) | 0.0272*** (0.0094) |
| log household size | 1.7522 (0.4848) | -0.0314*** (0.0068) | 0.0001 (0.0104) |
| Proportion males 0-4 | 0.0677 (0.1043) | 0.0025 (0.0015) | 0.0034 (0.0026) |
| Proportion females 0-4 | 0.0634 (0.1020) | 0.0021 (0.0014) | 0.0044* (0.0024) |
| Proportion males 5-9 | 0.0722 (0.1055) | -0.0007 (0.0014) | -0.0023 (0.0024) |
| Proportion females 5-9 | 0.0656 (0.1004) | -0.0001 (0.0015) | -0.0002 (0.0025) |
| Proportion males 10-14 | 0.0666 (0.1052) | -0.0007 (0.0014) | -0.0030 (0.0024) |
| Proportion females 10-14 | 0.0575 (0.0962) | -0.0006 (0.0013) | -0.0027 (0.0022) |
| Proportion males 15-55 | 0.2701 (0.1596) | 0.0113*** (0.0023) | 0.0023 (0.0030) |
| Proportion females 15-55 | 0.2625 (0.1339) | -0.0089*** (0.0017) | -0.0016 (0.0026) |
| Proportion males over 55 | 0.0375 (0.0875) | -0.0013 (0.0011) | -0.0007 (0.0017) |
| Proportion females over 55 | 0.0370 (0.0917) | -0.0036*** (0.0010) | 0.0005 (0.0016) |
| Rural self-employed in non-agriculture | 0.0972 (0.2963) | 0.0036 (0.0038) | 0.0041 (0.0068) |
| Urban self-employed | 0.0853 (0.2793) | 0.0212*** (0.0043) | 0.0158*** (0.0060) |
| Rural agricultural labor | 0.2151 (0.4109) | -0.0110** (0.0044) | -0.0160* (0.0085) |
| Urban wage earner | 0.0513 (0.2207) | -0.0054* (0.0031) | -0.0022 (0.0042) |

(Continued)

The last three rows test the assumption that migrants face the same prices as nonmigrants. I calculate household-level prices by dividing household expenditure on a food by the calories purchased. The first row shows the migrant-dummy coefficient when the log price per calorie is regressed on a product-village fixed effect and a migrant dummy. The second row shows the same coefficient but the regression now includes expenditure controls in the shape of a round-specific cubic in log household food expenditure per capita. The third row also includes household characteristic controls. Migrants pay 0.3 percent more than nonmigrant households in the same village buying the same product. However, this difference is due to migrants being wealthier than nonmigrants—and presumably buying higher quality

TABLE 2—DIFFERENCES BETWEEN MIGRANT AND NONMIGRANT HOUSEHOLDS (*Continued*)

| | Mean (full sample) (1) | Migrant difference (full sample) (2) | Migrant difference (wife moved for marriage sample) (3) |
|---|------------------------------|--|--|
| Rural other labor | 0.0553 (0.2286) | 0.0069** (0.0028) | −0.0006 (0.0046) |
| Urban casual labor | 0.0143 (0.1188) | −0.0037** (0.0018) | −0.0010 (0.0023) |
| Rural self-employed in agriculture | 0.3464 (0.4758) | −0.0075 (0.0049) | −0.0016 (0.0097) |
| Rural other | 0.0643 (0.2453) | 0.0080** (0.0031) | 0.0141*** (0.0049) |
| Urban other | 0.0707 (0.2564) | −0.0120*** (0.0033) | −0.0126*** (0.0046) |
| log price paid (no controls) (rupees per 1,000 calories) | 0.6458 (1.0874) | 0.00341*** (0.00115) | −0.00033 (0.00204) |
| log price paid (food expenditure controls) (rupees per 1,000 calories) | 0.6451 (1.0876) | 0.00154 (0.00113) | −0.00137 (0.00202) |
| log price paid (controls) (rupees per 1,000 calories) | 0.6447 (1.0874) | 0.00102 (0.00112) | −0.00184 (0.00202) |

Notes: Column 1 shows mean of each household-level variable in the row title. Column 2 shows coefficient on a migrant dummy when the variable is regressed on a village-round fixed effect and a migrant-status dummy. Column 3 shows coefficient for same regression but restricting attention to households in the wife moved for marriage sample described in text. Last three rows show 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 vector of controls for log per capita food expenditure and household characteristics used in later regressions, the food expenditure controls row just uses a cubic in log food expenditure per capita. Regressions weighted by household weights except the last three rows that use food-budget shares interacted with household weights to ensure that important foods are more influential and that the sum of weights for each household equals the NSS household weight. Regressions clustered at level of the fixed effects.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

products—since it becomes small and insignificant once I control for household food expenditure in rows two and three. Therefore, in order to ensure that migrants and nonmigrants in the same village are paying the same prices, all my specifications include a full set of controls for expenditure and household characteristics.

After controlling for observables, migrant and nonmigrant households may still differ on unobservables. As an important robustness check, I draw on an alternative sample in which migrant and nonmigrant 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 selection problems that arise when household heads are choosing to migrate for better employment opportunities or because they are particularly adaptable to different cultures.⁹

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

⁹To produce my findings, migrants need to consume higher price per calorie foods than nonmigrants with similar incomes for reasons unrelated to their origin-state tastes. The bias works in the other direction if migrants are more likely to be manual laborers consuming diets heavy in cheap carbohydrates or have unusually adaptable tastes.

(Srinivas 1980). This norm of “patrilocality” is so prevalent that 57 percent of wives report both that their current village is not their last usual residence and list the reason for leaving that location as “on marriage.”¹⁰ 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 only on households in which the wife moved for marriage, and comparing those in which the wife moved interstate (migrant households) to those in which the wife moved intrastate (nonmigrant households). In the spirit of the exercise, I also exclude households in which both husband and wife moved at the same time since these households may be moving for work opportunities.

Although a similar proportion of households are classified as migrants in both this *wife moved for marriage* sample and the main sample, Table 1 shows that migrants 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 more likely to be long-term migrants. Table 2 confirms the conjecture that migrant and nonmigrant households appear more similar in the *wife moved for marriage* sample. Although migrant households still have higher expenditures than nonmigrants, the difference in expenditures 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 nonmigrants for the same product, and these differences are minuscule 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 nonmigrants.

Inevitably, the methodology outlined in this section 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.

III. Migrants Consume Fewer Calories per Rupee than Nonmigrants

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 nonmigrant households living around them.

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

¹⁰In contrast, under 1 percent of male household heads moved location for the purpose of marriage. Among male movers, 48 percent cite employment reasons and 18 percent cite marriage (these figures are 2 and 85 percent for wives).

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):

$$(1) \quad \ln \text{calories}_i = \beta_1 \text{migrant}_i + d_{vt} + \Pi_t \mathbf{X}_i + \varepsilon_i.$$

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 nonmigrants in their village, conditioning on food expenditure and other household-level controls.

The demographic and characteristic variables in \mathbf{X}_i 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. The controls for household demographics include log household size as well as the proportion of household members that fall into 5 sex-specific age brackets; 0–4, 5–9, 10–14, 15–55, and over 55. The controls for household characteristics are indicator variables for the household's primary activity.¹¹ I allow the coefficients on all my controls to differ by survey round. Subramanian and Deaton (1996) also include indicators for religion and caste. Since religious affiliation and caste membership may be cultural determinants of food preferences, I exclude these 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 3 shows the results of this regression. I reject the null hypothesis, that migrants consume an equal or greater number of calories per rupee than nonmigrants, (e.g., $\beta_1 \geq 0$) at the 1 percent level: interstate migrant households are consuming 1.59 percent fewer calories than their nonmigrant neighbors, controlling for food expenditure (with a 95 percent confidence interval between 0.95 and 2.23 percent). 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.

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

¹¹ The categories are: 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.

TABLE 3—COMPARING THE CALORIC INTAKE OF MIGRANTS AND NONMIGRANTS

| Specification: | log caloric intake (per person per day) | | | | | |
|--|---|---|--------------------------------|----------------------------|--------------------------------|--|
| | Baseline specification | Caloric intake from rice and wheat only | Wife moved for marriage sample | Total expenditure controls | Real food expenditure controls | State-specific food expenditure controls |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| <i>migrant_i</i> | -0.0159*** (0.00326) | -0.00962*** (0.00214) | -0.0121*** (0.00356) | -0.0136*** (0.00442) | -0.0117*** (0.00418) | -0.0168*** (0.00316) |
| Observations | 235,126 | 226,584 | 91,406 | 235,104 | 235,122 | 235,126 |
| Within <i>R</i> ² | 0.732 | 0.905 | 0.718 | 0.520 | 0.676 | 0.723 |
| Food expenditure controls | Yes | Yes | Yes | No | No | No |
| Total expenditure controls | No | No | No | Yes | No | No |
| Real food expenditure controls | No | No | No | No | Yes | No |
| State-specific food expenditure controls | No | No | No | No | No | Yes |
| Demographics/household type controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Village-round FE | Yes | Yes | Yes | Yes | Yes | Yes |

| Specification: | log caloric intake (per person per day) | | | | log calories per rupee of food expenditure | |
|--|--|-------------------------------------|-------------------------------|---|--|---------------------------------|
| | Occupation, industry, meals served, and meals out controls | Meals served and meals out controls | Food expenditure instrumented | PDS foods (rice, wheat, and sugar) excluded | Baseline specification | Priced at village median prices |
| | (7) | (8) | (9) | (10) | (11) | (12) |
| <i>migrant_i</i> | -0.0135*** (0.00343) | -0.0150*** (0.00346) | -0.0118*** (0.00332) | -0.0209*** (0.00483) | -0.0145*** (0.00322) | -0.0149*** (0.00351) |
| Observations | 230,027 | 231,281 | 234,961 | 235,068 | 235,104 | 235,104 |
| Within <i>R</i> ² | 0.831 | 0.730 | 0.711 | 0.729 | 0.325 | 0.306 |
| Food expenditure controls | Yes | Yes | Yes | Yes | No | No |
| Total expenditure controls | No | No | No | No | Yes | Yes |
| Real food expenditure controls | No | No | No | No | No | No |
| State-specific food expenditure controls | No | No | No | No | No | No |
| Demographics/household type controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Village-round FE | Yes | Yes | Yes | Yes | Yes | Yes |

Notes: Dependent variable for columns 1–10 is log caloric intake per person per day. Dependent variable for column 11 is log caloric intake per rupee of actual food expenditure and for column 12 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. 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 4 to 6 and columns 11 to 12 which use alternate expenditure controls detailed in column headers, where the food expenditure controls in column 5 are deflated by a state-specific Stone price index: the sum of log prices weighted by state budget shares). Column 2 restricts attention to calories from rice and wheat products only. Column 3 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 (nonmigrants). Columns 7 and 8 include round dummies interacted with controls for the occupation of the household head (643 categories), industry of head (526 categories), the ratio of meals served to guests and employees to those consumed by household members at home, and the ratio of meals consumed by household members outside the house to those consumed inside the house. Column 9 instruments the polynomial in total food expenditure with a polynomial in total non-food expenditure. Column 10 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.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

along a multitude of routes and facing positive and negative caloric taxes.¹² 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 Section IVC and IVD, I explore both these dimensions of heterogeneity and find substantially higher caloric taxes for the most adversely affected migrant groups.

A. Robustness Checks

The remaining columns of Table 3 report a variety of robustness checks. Reductions in caloric intake may not lower nutrition if households substitute calorie-rich foods for protein- or vitamin-rich foods. As discussed in Section II, the diversity of migration patterns give us no reason to think migrants substitute between the components of a nutritious diet differently than observationally-equivalent nonmigrants facing the same prices. Thus, the smaller number of calories per rupee that migrants consume likely implies lower levels of nutrition. To buttress this claim, column 2 restricts attention to caloric intake from rice and wheat products that together account for 65.5 percent of total caloric consumption. The migrant caloric tax falls to 0.96 percent but remains significant. As rice and wheat have very similar nutritional profiles, fewer calories from rice and wheat almost certainly correspond to less nutrition from the two foods.¹³ Furthermore, Section IIIB shows similar reductions in protein as well as other nutrition metrics.

Column 3 runs the regression on the *wife moved for marriage* sample for which unobservable differences between migrant and nonmigrant households were less of a concern. I compare households where wives moved intrastate at the time of marriage (a nonmigrant) with those where the wife moved interstate (a migrant). The caloric tax on migrants is still significantly negative for this sample but is attenuated by 24 percent.¹⁴ The decline in the coefficient size is not surprising. Wives typically move into their husband's household (often containing extended family members such as the husband's parents). Any cultural preferences brought by the wife are likely to have smaller impacts on household spending decisions compared to the scenario where both husband and wife are migrants (a hypothesis I test in Section IV).

Columns 4 to 6 use alternative expenditure controls in place of the polynomial in log per capita food expenditure. Column 4 uses a third-order polynomial in log per capita expenditure on all goods. I find a caloric tax of 1.36 percent. Since the coefficients are of similar magnitude when I control for either food expenditure or total expenditure, migrants are not simply substituting from non-food to food to accommodate their food preferences. Results are similar in column 5, which uses a

¹²The 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 (2013).

¹³Of course migrants may compensate by purchasing more nutritious foods in other categories or spending more on rice and wheat. However, both of these solutions come at a cost which this paper seeks to quantify.

¹⁴I would also find $\beta_1 < 0$ if wives from further away are more valued and fed higher quality foods; or if wives choose cheaper calorie sources than others in the household and wives from further away control a smaller share of household budgets.

polynomial in log per capita real food expenditure, and column 6, which allows the food expenditure elasticities to vary by state.

My measure of caloric intake may be misleading if migrants work in jobs that require less energy, eat out more, or serve a greater share of food purchases to non-household members. In column 7, I include finer controls for the occupation (643 categories) and industry (526 categories) of the household head, as well as for the number of meals consumed by household members away from home and served to employees and guests (both as ratios of the number of meals consumed at home by household members). The caloric tax declines slightly to 1.35 percent, predominantly due to the thousands of industry-occupation controls (column 8 includes only meal controls and finds a tax of 1.50 percent and online Appendix Table C2 shows that migrants do not differ in terms of meals out or meals served to others).

Column 9 of Table 3 instruments the food expenditure polynomial with a polynomial in non-food expenditure to mitigate correlated measurement error concerns (since both calories and food expenditure are calculated from the same raw data).¹⁵ Food expenditure may also be endogenous. A shock that increases the demand for calories, such as changing work patterns, will also increase food expenditure, biasing the coefficients on food expenditure upwards. However, there will be a negative or no correlation with non-food expenditure so the true value of β_1 will very likely be bounded between the instrumented and uninstrumented estimates. (Online Appendix B formalizes this claim.) Since the estimated β_1 is only attenuated by a quarter and is still significantly negative in the instrumental variable specification, neither measurement error nor the endogeneity of food expenditure seem to be problematic in this context.

Finally, I return to the concern that migrants and nonmigrants pay different prices (recall from Table 2 that, although insignificant, migrants paid slightly higher prices). This concern is most severe for the three foods (rice, wheat, and sugar) that are commonly sold through the subsidized Public Distribution System (PDS).¹⁶ Although the system was not restricted to poor households until the 1990s, migrants may still have had worse access to this system and hence paid higher prices. Column 10 reproduces my main specification excluding these three food groups. Reassuringly, I find a larger caloric tax of 2.09 percent when I exclude these PDS foods. Columns 11 and 12 take a different approach. Column 11 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 12 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 nonmigrant 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

¹⁵The higher-order non-food expenditure terms are weak predictors of higher-order food expenditure and so there are weak instruments concerns. If I use just first order terms instead, the Angrist-Pischke F -statistics are 1,821 for round 38 and 606 for round 43, and the coefficient on the migrant dummy is essentially unchanged (rising slightly from 0.118 to 0.120).

¹⁶If I reproduce the price regressions just for foods in these three groups, migrants pay (an insignificant) 0.24 percent more for these foods than nonmigrants in the same village after conditioning on the full set of controls.

consume fewer calories than locals through purchasing different consumption bundles rather than through paying different prices.

Online Appendix Table C3 allows the caloric tax to differ by survey round, by whether the migrant's origin and destination were rural or urban, and by whether the migrant's origin and destination states were rice or wheat dominated. Results are similar across rounds, but the migrant caloric tax is substantially smaller for migrants moving from rural areas or from wheat to rice states. In Section IVC, I return to this heterogeneity when I explore how the caloric tax varies by migration route.

B. 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 4 repeats the basic specification for various sub-populations that are poor and under-nourished, and also explores several alternative nutrition metrics.

Column 1 repeats the baseline specification. Columns 2 to 4 restrict attention to undernourished households (those consuming fewer calories than the 2,400 rural/2,100 urban calorie norms, or those either 250 or 500 calories per person per day below the norms). Columns 5 to 8 restrict attention to poorer households (those spending less than either the median or twenty-fifth percentile of either per capita expenditure or per capita food expenditure in that 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.¹⁷

Adequate nutrition is particularly important for households with young children as shortfalls at young ages have substantial scarring effects on productivity, earnings, and health in adulthood (Almond and Currie 2010). Columns 9 to 12 restrict attention to families with children below the age of either 5 or 16, and then further restricting attention to households spending less than the median level of per capita food expenditure. Even in these subpopulations, I find a significant caloric tax, 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.

Columns 17 to 32 of Table 4 replace log caloric intake with other nutritional metrics. To explore where in the distribution the migrant caloric tax hits I consider the "caloric shortfall," the number of calories a household consumes below the 2,400 rural/2,100 urban calorie norms. As shown in columns 17 to 20, the shortfall is significantly greater for migrant households, and migrants are 1.4 percent more likely to be 250 calories and 1.5 percent more likely to be 500 calories (per person per day) below the calorie norms. Columns 21 to 28 repeat this exercise for households with children under 5 and under 16, and draw similar conclusions (although when I halve the sample by only considering households with children under 5, the binary

¹⁷The coefficients are all significantly negative except when considering the bottom expenditure quartiles where the sample size is much smaller and the standard errors much higher (rather than the effect size being much lower).

TABLE 4—COMPARING THE CALORIC INTAKE OF MIGRANTS AND NONMIGRANTS: SUBPOPULATIONS AND ALTERNATIVE NUTRITION MEASURES

| Sample: | log caloric intake (per person per day) | | | | | | | |
|--------------------------------------|---|--------------------------|-------------------------------|-------------------------------|--|---|---|--|
| | Baseline specification (1) | Caloric shortfall (2) | Caloric shortfall >250 (3) | Caloric shortfall >500 (4) | Below median per capita expenditure (5) | Bottom quartile per capita expenditure (6) | Below median per capita food expenditure (7) | Bottom quartile per capita food expenditure (8) |
| <i>migrant_i</i> | -0.0159*** (0.00326) | -0.0158*** (0.00414) | -0.0122*** (0.00448) | -0.0134** (0.00577) | -0.0107** (0.00471) | -0.0141 (0.00865) | -0.0147** (0.00581) | -0.00981 (0.00703) |
| Observations | 235,126 | 131,149 | 96,169 | 59,588 | 87,010 | 36,436 | 86,531 | 36,921 |
| Within <i>R</i> ² | 0.732 | 0.640 | 0.627 | 0.622 | 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 |

| Sample: | log caloric intake (per person per day) | | | | | | | |
|--------------------------------------|---|--|--|---|--------------------------------------|---------------------------------------|---------------------------------------|-------------------------|
| | Children under 5 in household (9) | Children under 5 and below median food expend. (10) | Children under 16 in household (11) | Children under 16 and below median food expend. (12) | Migrated 5 or more years ago (13) | Migrated 10 or more years ago (14) | Migrated 20 or more years ago (15) | Literate (16) |
| <i>migrant_i</i> | -0.0132*** (0.00461) | -0.0141** (0.00715) | -0.0169*** (0.00330) | -0.0140*** (0.00533) | -0.0114*** (0.00335) | -0.00650** (0.00323) | 0.00183 (0.00421) | -0.0171*** (0.00401) |
| Observations | 103,652 | 47,881 | 178,456 | 75,644 | 231,268 | 228,442 | 223,489 | 128,706 |
| Within <i>R</i> ² | 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 |

| Sample: | All households | | | | Children under 5 in household | | | |
|--------------------------------------|---------------------------|------------------------------------|---|---|-------------------------------|------------------------------------|---|---|
| | Caloric shortfall (17) | Caloric shortfall (binary) (18) | Caloric shortfall >250 (binary) (19) | Caloric shortfall >500 (binary) (20) | Caloric shortfall (21) | Caloric shortfall (binary) (22) | Caloric shortfall >250 (binary) (23) | Caloric shortfall >500 (binary) (24) |
| <i>migrant_i</i> | 14.21*** (4.515) | 0.00330 (0.00417) | 0.0137** (0.00606) | 0.0149** (0.00608) | 13.40** (6.209) | -0.00109 (0.00651) | 0.0122 (0.00808) | 0.0115 (0.00792) |
| Observations | 235,126 | 235,126 | 235,126 | 235,126 | 103,652 | 103,652 | 103,652 | 103,652 |
| Within <i>R</i> ² | 0.633 | 0.397 | 0.401 | 0.377 | 0.626 | 0.346 | 0.363 | 0.356 |
| 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 |

(Continued)

measures become marginally insignificant). As above, it is not only well nourished households who pay the caloric tax.

I also consider several alternative nutrition metrics. Columns 29 and 30 consider protein and fat intake, respectively. Migrant households consume significantly less

TABLE 4—COMPARING THE CALORIC INTAKE OF MIGRANTS AND NONMIGRANTS: SUBPOPULATIONS AND ALTERNATIVE NUTRITION MEASURES (Continued)

| Sample: | Children under 16 in household | | | | All households | | | 38th round only |
|--------------------------------------|--------------------------------|---------------------------------|--------------------------------------|--------------------------------------|--|--|---|---|
| | Caloric shortfall (25) | Caloric shortfall (binary) (26) | Caloric shortfall >250 (binary) (27) | Caloric shortfall >500 (binary) (28) | log protein intake (gm. per person per day) (29) | log fat intake (gm. per person per day) (30) | log meals per household member per day (31) | Obtain two square meals a day (binary) (32) |
| <i>migrant_i</i> | 14.49*** (4.502) | 0.00461 (0.00537) | 0.0126* (0.00667) | 0.0135** (0.00615) | -0.0160*** (0.00365) | -0.000544 (0.00352) | -0.000915 (0.00208) | 0.000227 (0.00431) |
| Observations | 178,456 | 178,456 | 178,456 | 178,456 | 235,124 | 235,121 | 234,838 | 109,672 |
| Within R^2 | 0.628 | 0.367 | 0.379 | 0.365 | 0.682 | 0.746 | 0.043 | 0.162 |
| 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 |

Notes: Dependent variable for columns 1–16 is log caloric intake per person per day. Dependent variable for columns 17–32 is detailed in column heading and described in Section IIIB. 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 on columns 1–16 and the subheadings in columns 17–32 denote the various subpopulations on which the regressions are run. Caloric shortfall defined relative to the 2,400 calorie urban/2,100 rural calorie norms. All regressions are weighted using household weights and the standard errors are two-way clustered at the village-round and origin-state level.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

protein, an essential component of nutrition, than nonmigrant households but there is little difference in fat consumption. Column 31 explores the number of meals consumed per day by each household member, while column 32 explores the binary response to the survey question “Do all members of your household get two square meals a day?” asked in round 38. As discussed by Deaton and Drèze (2009), caloric shortfalls and skipping meals are not the same thing and are uncorrelated in the Indian NSS data. Thus, it is perhaps not surprising that I do not find significant differences between migrant and nonmigrant households along these dimensions (although, as discussed in Section IVC, significant and borderline significant relationships appear for migrants moving to locations where relative prices are unsuited to their origin-state preferences).

Online Appendix Table C4 presents supportive results using another dataset, the 2005 and 2011 rounds of the India Human Development Survey (Desai et al. 2005, 2011), that contains anthropometric data. Households are asked if they migrated from another state, although not which state that was, which precludes using these data in the later analysis. I repeat the basic specification but replacing log calories with either child weight-for-age or child height-for-age Z-scores (both calculated using the 2006 WHO child growth standards for children aged 0 to 5). I find that migrant children weigh 0.220 standard deviations less and are 0.222 standard deviations smaller than nonmigrant children in their same village (significant at the 5 and 10 percent levels, respectively, with further details and specifications for older ages provided in the table).

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

IV. Why Do Migrants Consume Fewer Calories than Nonmigrants?

In this section, I form a chain of evidence in support of an explanation for the previous findings based on culture: that migrants make calorically suboptimal food choices due to strong preferences for the favored foods of their origin states. First, I focus on the preferences themselves. In Section IVA and IVB, I show that migrants bring their origin-state food preferences with them when they migrate and that the intensity of these preferences depends on whether both husband and wife are migrants. Second, I combine these preference results with my caloric tax findings. Sections IVC and IVD show that the heterogeneity in the size of the migrant caloric tax is related to the suitability and intensity of the migrant household’s 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. This chain of evidence suggests that Indian migrants consume fewer calories than nonmigrants because they prefer to purchase the favored products from their origin state even when these products are relatively expensive compared to local alternatives.

A. Migrants Bring Their Food Preferences with Them

In this subsection, 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 rice and wheat. I test whether a migrant’s rice consumption relates to rice consumption in their origin state. As in the economics of culture literature, I can just focus on migrants and test whether migrants 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 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 on their origin state (where the origin state o average is calculated using nonmigrant households interviewed in the same survey round as household i),

$$(2) \quad \frac{rice_i}{rice_i + wheat_i} = \alpha_1 \frac{rice_i^o}{rice_i^o + wheat_i^o} + d_{vt} + \Pi_t \mathbf{X}_i + \varepsilon_i.$$

For nonmigrants, the origin-state rice share is the average rice share of their current state. The regression also includes village fixed effects, d_{vt} , and the vector \mathbf{X}_i of household-level controls used in Section III.

TABLE 5—COMPARING BUNDLES OF MIGRANTS AND NONMIGRANTS

| Dependent variable: | Rice expenditure share $rice_i/(rice_i + wheat_i)$ | | Correlation $corr_{id}$ of household i budget shares with budget shares of current state d | Correlation $corr_{io_v}$ 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 (1) | All households (2) | | Baseline specifica- tion (all households) (4) | Migrant households only (5) | Wife moved for marriage sample (6) |
| $rice_i^o/(rice_i^o + wheat_i^o)$ | 0.189*** (0.0204) | 0.123*** (0.0146) | | | | |
| $migrant_i$ | | | -0.0111*** (0.00199) | | | |
| $migrant_i^{o_v}$ | | | | 0.0226*** (0.00290) | 0.0355*** (0.00285) | 0.0103*** (0.00284) |
| Observations | 14,156 | 226,472 | 235,126 | 108,743 | 36,635 | 23,397 |
| Within R^2 | 0.065 | 0.007 | 0.080 | 0.098 | 0.142 | 0.062 |
| Food expenditure controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Alternative expenditure controls | No | No | No | No | No | No |
| Demographics/house- hold 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 |

(Continued)

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 in their village consume a larger share of rice than other migrants (and vice versa for migrants from more wheat-loving states). In column 1 of Table 5, I find support for the hypothesis with a positive and highly significant estimate of α_1 equal to 0.189. 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 5 reports this regression. Once more, I find a positive and highly significant estimate of α_1 (here equal to 0.123).¹⁸

Although informative, such an exercise only considers two types of food. If I wish to consider all 169 foods, I could repeat the exercise above for each food and then aggregate the coefficients. 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.¹⁹

¹⁸If 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.

¹⁹If I regress food-budget shares for each food on the average food-budget share for that food in the household's origin state and the controls in equation (2), 145 of the 180 coefficients are positive and 72 are significant at the 5 percent level.

TABLE 5—COMPARING BUNDLES OF MIGRANTS AND NONMIGRANTS (Continued)

| Dependent variable: | Correlation $corr_{io_v}$ of household i budget shares with the average budget shares of the origin states o_v of migrants in village | | | | | | Correlation of household caloric shares with caloric shares of origin states o_v |
|---|---|------------------------------------|--|--|---|---|--|
| | Total expenditure controls (7) | Real food expenditure controls (8) | State-specific food expenditure controls (9) | PDS foods (rice, wheat, and sugar) excluded (10) | Reference basket by expenditure quartile (11) | Interactions with $\ln K$ difference (12) | |
| $migrant_i^{o_v}$ | 0.0225*** (0.00293) | 0.0215*** (0.00299) | 0.0223*** (0.00296) | 0.0171*** (0.00287) | 0.0217*** (0.00245) | 0.0270*** (0.00379) | 0.0263*** (0.00334) |
| $migrant_i^{o_v} \times$ $1[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v)$ $< \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$ | | | | | | -0.0142** (0.00617) | |
| Observations | 108,726 | 108,737 | 108,743 | 108,696 | 108,638 | 105,810 | 108,743 |
| Within R^2 | 0.074 | 0.135 | 0.066 | 0.104 | 0.092 | 0.096 | 0.060 |
| Food expenditure controls | No | No | No | Yes | Yes | Yes | Yes |
| Alternative expenditure controls | Yes | Yes | Yes | No | No | No | No |
| Demographics/household type | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Village- o_v -round FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Notes: Columns 1 and 2 regress rice’s share of household rice and wheat expenditures on the average rice share of the household’s origin state, with column 1 restricting attention only to migrants (where rice and wheat expenditure includes all 12 of the rice and wheat-based products in the surveys). Column 3 regresses $corr_{id}$, the correlation between household i ’s vector of food budget shares and the vector of mean budget shares of nonmigrant households in household i ’s current state d , on a migrant dummy. Dependent variable in columns 4 to 12 is the correlation between household i ’s vector of food budget shares and the vector of mean budget shares for nonmigrant 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 7 to 9 which use alternate expenditure controls). Column 5 restricts attention to migrant households in villages with multiple migrant origin states. Column 6 restricts attention to households with wives who moved at the time of marriage, column 10 excludes the foods commonly sold through the Public Distribution System, column 11 matches households to a reference basket calculated separately by expenditure-quartile, column 12 includes an additional interaction with an indicator variable that takes the value of 1 if the migrant’s origin-state bundle provides fewer calories than the local bundle, and column 13 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.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

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 $corr_{is} \equiv corr(\mathbf{bshare}_i, \mathbf{bshare}_s^i)$, 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}_s^i). As with the rice-wheat specification, the state-averages are calculated using only nonmigrant households interviewed in the same survey round as household i (hence the need for an i superscript on \mathbf{bshare}_s^i). 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.²⁰

²⁰To see this, note that the correlation between vectors \mathbf{x} and \mathbf{y} is equal to $\sum(x_j y_j - \bar{x}\bar{y}) / (n - 1)s_x s_y$ 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 $1/169$.

These $corr_{is}$ correlations measure the similarity between household i 's preferences and the average preferences of nonmigrants in state s .²¹ I test whether migrant and nonmigrants possess the same preferences by comparing the size of these correlations 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 $corr_{id}$ 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,

$$(3) \quad corr_{id} = \beta_1 migrant_i + d_{vt} + \Pi_t \mathbf{X}_i + \varepsilon_{id}.$$

Once again, I include village fixed effects, d_{vt} , and household-level controls, \mathbf{X}_i . A negative value of β_1 indicates that migrant households consume bundles that are less similar to the current-state bundle (in comparison to nonmigrant households in the village). As shown in column 3 of Table 5, I find support for this sign prediction with an estimated coefficient of -0.0111 . I can reject the null hypothesis, that migrants do not differ from nonmigrants in the manner described above ($\beta_1 \geq 0$), at the 1 percent level.²²

The finding that migrants possess different preferences than nonmigrants 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 nonmigrants in the village to the migrant's origin-state bundle. To do this, I switch correlation measures to the correlation $corr_{io_v}$ 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 ,

$$(4) \quad corr_{io_v} = \gamma_1 migrant_i^{o_v} + d_{vo_v t} + \Pi_t \mathbf{X}_i + \varepsilon_{io_v}.$$

Villages may have multiple origin states o_v if there are migrants from more than 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

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.

²¹If households have Cobb-Douglas preferences, $u = \prod_{g=1}^{169} c_g^{\theta_g}$ 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 nonmigrants in state s . Atkin (2013) proposes an alternative approach that allows for non-homotheticities and budget shares that respond to prices. However, estimating migrant preferences in this manner is infeasible in this context since there are very few migrants on any particular route.

²²The negative coefficient on $migrant_i$ is partly mechanical since average budget shares of state d were calculated using only nonmigrant households. Although this bias is likely to be small (the average state-round sample contains 3,700 nonmigrant households), I can reproduce the regression using average bundles calculated using all households where such a bias is absent. The β_1 coefficient remains significantly negative at the 1 percent level, rising only slightly to -0.0106 .

effect for each origin state in each village, $d_{v,o,t}$, in addition to the set of household-level controls used in the previous specifications.²³ 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 in column 4 of Table 5, 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. 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.

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 nonmigrant 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).

Since there are villages containing migrants from multiple states, I can rerun the regression in equation (4) only on migrant households. Such an approach mitigates concerns that my results are driven by migrant and nonmigrant households systematically differing on unobservables. In this case, the interpretation of a positive γ_1 coefficient is that, compared to migrants from other states living in the same village, a migrant from o_v 's bundle is more similar to state o_v 's average bundle. As shown in column 5 of Table 5, I find a significant and larger positive coefficient of 0.0355 when I restrict attention only to migrants.

Columns 6 through 10 of Table 5 run the regression specified in equation (4) for the *wife moved for marriage* sample, for the alternative expenditure specifications, 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 11 presents one such correction. I recalculate the reference bundles using nonmigrant households in the same national expenditure 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 expenditure 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 13 calculates the correlations using vectors of caloric shares instead (where a caloric share is a food item's share of household caloric

²³ As previously, I two-way cluster at the village-round and origin state o of the household. 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.

consumption). Results are similar across all the robustness specifications, with γ_1 significantly positive in every regression.

Finally, online Appendix Table C5 reports these correlation results for the various subsamples of poor and undernourished households detailed in Section IIIB. Mirroring the caloric tax results, I find positive (and significant) coefficients across the various subsamples.²⁴

In summary, migrant households bring their origin-state food preferences with them when they migrate.²⁵ These food preferences may be related to observable origin-state characteristics such as prices, temperature, and income. However, I interpret them as cultural food preferences as they persist after migration when migrants face a new set of prices and external conditions, and they are purged of differences in current expenditure and other household-level characteristics through a rich set of controls.

B. 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 with non-local preferences in 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.

Table 6 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 $migrant_i^{ov}$ with dummies for the migrant status of the household head and their spouse: (i) only one of either the head or spouse is a migrant (*onlyone_i*); (ii) both the head and spouse are migrants (*both_i*); or (iii) there is no spouse (*nospouse_i*). 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

$$(5) \text{corr}_{io_v} = \gamma_1 \text{migrant}_i^{ov} \times \text{onlyone}_i + \gamma_2 \text{migrant}_i^{ov} \times \text{both}_i \\ + \gamma_3 \text{migrant}_i^{ov} \times \text{nospouse}_i + d_{vo_vt} + \pi_1 \text{nospouse}_i + \Pi_t \mathbf{X}_i + \varepsilon_{io_v}$$

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 nonmigrant 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 row 1 of Table 6.

²⁴The two exceptions are the bottom-quartile columns 6 and 8 where the sample size shrinks by more than 90 percent.

²⁵Online Appendix Table C6 provides a related piece of evidence that migrants are more likely to grow the crops that are commonly grown in their origin state (using the forty-third round survey that asked whether a household cultivates each of 8 major crops).

TABLE 6—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 ² |
|--|---|--------------------------------|-------------------------|------------------------|--------------|-----------------------|
| (1) Dependent variable: correlation $corr_{io_v}$ | $migrant_i^{ov}$ | 0.00786*** (0.00251) | 0.0416*** (0.00453) | 0.0331*** (0.00819) | 108,743 | 0.100 |
| (2) Dependent variable: log caloric intake | $migrant_i$ | -0.0125*** (0.00292) | -0.0228*** (0.00570) | -0.0139** (0.00618) | 235,126 | 0.732 |
| (3) Dependent variable: log caloric intake | $migrant_i$ | -0.00397 (0.00376) | -0.00913 (0.00672) | -0.000555 (0.00662) | | |
| | $migrant_i \times 1[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$ | -0.0179*** (0.00476) | -0.0629*** (0.0146) | -0.0359*** (0.0116) | 234,155 | 0.732 |
| <i>Subsample of households whose primary occupation is agriculture</i> | | | | | | |
| (4) Dependent variable: log caloric intake | $migrant_i$ | 0.0105 (0.00781) | 0.0354** (0.0146) | 0.0343* (0.0200) | | |
| | $migrant_i \times 1[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$ | -0.0296*** (0.00851) | -0.0854*** (0.0239) | -0.0671** (0.0274) | 108,719 | 0.745 |
| <i>Subsample of households with caloric shortfalls of 250 calories or more</i> | | | | | | |
| (5) Dependent variable: log caloric intake | $migrant_i$ | 0.00641 (0.00821) | -0.000969 (0.00773) | -0.00713 (0.0115) | | |
| | $migrant_i \times 1[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$ | -0.0332*** (0.0103) | -0.0512** (0.0199) | -0.0568*** (0.0206) | 95,910 | 0.627 |
| <i>Subsample of households with caloric shortfalls of 500 calories or more</i> | | | | | | |
| (6) Dependent variable: log caloric intake | $migrant_i$ | 0.00869 (0.0109) | -0.0125 (0.0115) | -0.0178 (0.0174) | | |
| | $migrant_i \times 1[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$ | -0.0311** (0.0144) | -0.0405* (0.0223) | -0.0610* (0.0339) | 59,457 | 0.622 |

Notes: Rows 1–6 repeat the regressions shown in Tables 3, 5, and 7 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. Rows 4–6 restrict attention to subsamples described in italicized heads above rows. Caloric shortfall defined relative to the 2,400 calorie urban/2,100 rural calorie norms. If both head and spouse are migrants from different origin states, I replace $migrant_i^{ov}$ in panel 1 with the value of one half for each of the two origin states, and $1[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$ with its average value across the two origin states in rows 3–6. Results are essentially unchanged if these households are dropped. As in Table 5, panel 1 includes village- α_v -round fixed effects. As in Tables 3 and 7, rows 2–6 include village-round fixed effects. All rows 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.

- *** Significant at the 1 percent level.
- ** Significant at the 5 percent level.
- * Significant at the 10 percent level.

When 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 $migrant_i^{ov} \times onlyone_i$ and $migrant_i^{ov} \times both_i$ 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.²⁶

²⁶The 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

C. 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. 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 nonmigrants in their origin state. I then calculate $\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^o)$, 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 .²⁷ 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 nonmigrant 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, $1[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^o) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$, that takes the value of 1 for negative values of the log difference described above,²⁸

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

The values of the log differences range from -0.55 for the first percentile of migrants to 0.95 for the ninety-ninth percentile, with negative values for a quarter of migrant households. As before, I include village-round fixed effects, d_{vt} , and the vector \mathbf{X}_i of household controls described in Section III.

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 term is no longer significant.

²⁷I 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. I exclude migrant households from these median calculations to avoid issues arising from migrants paying different prices. If none of the resulting village sample purchase a good, I use the median price at an incrementally higher level of aggregation.

²⁸Note that no variable used to calculate this indicator variable uses budget share or price data from migrant household i .

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 7 presents the results of this regression. I can reject the null hypothesis at the 1 percent level. The estimated β_2 coefficient is significantly negative and equal to -0.0267 . 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 preferences are badly suited to the local price vector consume 3.23 percent fewer calories than comparable nonmigrant households (with a 95 percent confidence interval between 2.41 and 4.06 percent).

Columns 2 and 3 of Table 7 present alternative specifications for the $\ln K(\cdot, \cdot)$ interaction. Column 2 allows the caloric tax to vary for migrants in each quartile of the $\ln K(\cdot, \cdot)$ 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.²⁹ 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 2.11 percent more calories per rupee than their nonmigrant neighbors. Column 3 interacts the migrant dummy with $\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^o) - \ln K(\mathbf{bshare}_i^l, \mathbf{P}_i^l)$, a continuous measure of a migrant's caloric advantage over locals. I find a positive and significant coefficient of 0.06: the size of the caloric tax increases with the caloric disadvantage of a migrant's origin-state preferences.³⁰

As in Section IVA, I rerun these specifications restricting attention only to migrant households. Results are reported in columns 4 and 5 of Table 7. As with the full sample, I find that within villages that contain migrants from multiple states, migrant households from states with more advantageous origin-state preferences consume a greater number of calories per rupee of food expenditure than those from states with less advantageous preferences.³¹

I present many robustness checks and results for poor and undernourished subpopulations. Table 7 reports my findings using the *wife moved for marriage* subsample; only rice and wheat calories; alternative expenditure controls; occupation, industry, meals served, and meals out controls; only non-PDS foods; expenditure-quartile

²⁹ A negative tax for small positive values of the $\ln K(\cdot, \cdot)$ difference (i.e., quartiles two and three) is not surprising. First, the origin-state bundle is an imperfect measure of preferences unless the utility function takes the form in footnote 21, is permanent and is identical for everyone in the state. If richer households obtain fewer calories per rupee due to non-homotheticities and income rises with migration, the true $\ln K(\cdot, \cdot)$ difference is more negative. Second, if migrants pay higher prices than nonmigrants, the caloric advantage of the origin-state bundle is actually a small disadvantage when priced at the prices migrants pay rather than at village (nonmigrant) median prices. For example, if I just price the bundle at village median prices that include rather than exclude migrants, I find similar results, but here the second quartile corresponds to negative $\ln K(\cdot, \cdot)$ differences.

³⁰ The fact that the coefficient is substantially smaller than one implies that either migrants adapt their preferences after migrating, not all household members possess such preferences, or preferences are not of the Cobb-Douglas/identical-within-state form. I provide some explicit evidence for the first two of these explanations in later sections.

³¹ Since I am restricting attention to migrants, in column 4 the fourth quartile dummy is omitted and the results on the other quartiles are relative to the omitted quartile. Using an indicator variable for whether migrant preferences are more or less advantageous than local preferences is not appropriate here since I am not comparing migrants to locals.

TABLE 7—COMPARING THE CALORIC INTAKE OF MIGRANTS AND NONMIGRANTS ACROSS MIGRATION ROUTES

| Dependent variable: | log caloric intake (per person per day) | | | | | | |
|--|---|---------------------------------------|---------------------------------|-------------------------------------|---|---|------------------------------------|
| | Baseline specification (1) | Quartiles of ln K difference (2) | ln K difference interaction (3) | Migrants only (quartiles) (4) | Migrants only (interaction) (5) | PDS foods excluded (6) | PDS foods excluded (7) |
| <i>migrant_i</i> | -0.00564 (0.00423) | | | | | 0.00568 (0.00422) | |
| <i>migrant_i</i> × 1 [$\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v)$ < $\ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)$] | -0.0267*** (0.00574) | | | | | -0.0823*** (0.0123) | |
| <i>migrant_i</i> × 1st quartile [$\ln K(\cdot, \cdot) - \ln K(\cdot, \cdot)$] | | -0.0393*** (0.00547) | | -0.0162 (0.0152) | | | -0.0827*** (0.0120) |
| <i>migrant_i</i> × 2nd quartile [$\ln K(\cdot, \cdot) - \ln K(\cdot, \cdot)$] | | -0.0186*** (0.00446) | | -0.0224* (0.0134) | | | -0.0311*** (0.00980) |
| <i>migrant_i</i> × 3rd quartile [$\ln K(\cdot, \cdot) - \ln K(\cdot, \cdot)$] | | -0.0130** (0.00572) | | -0.0179** (0.00869) | | | 0.00694 (0.00571) |
| <i>migrant_i</i> × 4th quartile [$\ln K(\cdot, \cdot) - \ln K(\cdot, \cdot)$] | | 0.0211*** (0.00518) | | | | | 0.0371*** (0.00675) |
| <i>migrant_i</i> × ($\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v)$ - $\ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)$) | | | 0.0600*** (0.0112) | | 0.0395* (0.0226) | | |
| Observations | 234,157 | 234,157 | 234,157 | 14,186 | 14,186 | 234,099 | 234,099 |
| Within R^2 | 0.731 | 0.732 | 0.731 | 0.721 | 0.721 | 0.729 | 0.729 |
| Food expenditure controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Alternative expenditure controls | No | No | No | No | No | No | No |
| Demographics/household type controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Village-round FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Dependent variable: | log caloric intake (per person per day) | | | | | | |
| Specification: | Wife moved for marriage sample (8) | Calories from rice and wheat only (9) | Total expenditure controls (10) | Real food expenditure controls (11) | State-specific food expenditure controls (12) | Occupation industry and meals controls (13) | Food expenditure instrumented (14) |
| | | | | | | | |
| <i>migrant_i</i> | -0.00178 (0.00514) | 0.00672** (0.00337) | -0.00282 (0.00506) | -0.00318 (0.00578) | -0.00633 (0.00431) | -0.00314 (0.00447) | -0.00301 (0.00426) |
| <i>migrant_i</i> × 1 [$\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^v)$ < $\ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)$] | -0.0208*** (0.00727) | -0.0286*** (0.00410) | -0.0280*** (0.00608) | -0.0219*** (0.00687) | -0.0274*** (0.00597) | -0.0263*** (0.00577) | -0.0227*** (0.00541) |
| Observations | 91,158 | 225,587 | 234,135 | 234,153 | 234,157 | 229,120 | 233,993 |
| Within R^2 | 0.718 | 0.905 | 0.521 | 0.676 | 0.723 | 0.830 | 0.711 |
| Food expenditure controls | Yes | Yes | No | No | No | Yes | Yes |
| Alternative expenditure controls | No | No | Yes | Yes | Yes | No | No |
| Demographics/household type controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Village-round FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

(Continued)

adjusted reference baskets; instrumented food expenditure; and replacing calories with calories per rupee calculated using either actual or village median prices. Online Appendix Table C7 presents results for poor and undernourished subpopulations. In all the specifications, I find a significantly negative β_2 coefficient. The magnitudes for migrants with disadvantageous preferences range from a caloric tax of 1.92 percent for households below median expenditure, to a caloric tax of 3.35 percent for households 250 or more calories below the per person per day caloric norms.

TABLE 7—COMPARING THE CALORIC INTAKE OF MIGRANTS AND NONMIGRANTS ACROSS MIGRATION ROUTES (Continued)

| Dependent variable: | log calories per rupee of food expenditure | | log caloric intake (per person per day) | | |
|---|---|--|---|--|--|
| | Baseline specification (15) | Prices at village median prices (16) | Reference basket by expenditure quartile (17) | log migrant travel distance interactions (18) | Reference basket by destination state (19) |
| $migrant_i$ | -0.00454 (0.00424) | -0.00125 (0.00415) | -0.00727 (0.00467) | 0.0157 (0.0220) | -0.0103** (0.00504) |
| $migrant_i \times 1[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^o) < \ln K(\mathbf{bshare}_i^d, \mathbf{P}_i^d)]$ | -0.0259*** (0.00621) | -0.0348*** (0.00680) | -0.0254*** (0.00637) | -0.0251*** (0.00601) | -0.0108* (0.00578) |
| $migrant_i \times \ln distance_{ov}$ | | | | -0.00399 (0.00386) | |
| Observations | 234,135 | 234,135 | 231,774 | 213,444 | 235,004 |
| Within R^2 | 0.325 | 0.307 | 0.731 | 0.732 | 0.732 |
| Food expenditure controls | No | No | Yes | Yes | Yes |
| Alternative expenditure controls | Yes | Yes | No | No | No |
| Demographics/household type controls | Yes | Yes | Yes | Yes | Yes |
| Village-round FE | Yes | Yes | Yes | Yes | Yes |

Notes: Dependent variable is log caloric intake per person per day (except columns 15 and 16 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 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 migrants 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 10 to 12 and 15 to 16 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 report these alternative functions restricting attention only to migrant households in villages with multiple migrant origin states. Columns 6 and 7 exclude the foods commonly sold through the Public Distribution System. Column 8 restricts attention to households in which the wife moved village at the time of marriage and compares interstate to intrastate movers. Column 9 uses only calories from rice and wheat products. Column 13 includes round dummies interacted with controls for the occupation of the household head (643 categories), industry of head (526 categories), the ratio of meals served to guests and employees to those consumed by household members at home, and the ratio of meals consumed by household members outside the house to those consumed inside the house. Column 14 instruments total food expenditure with total non-food expenditure. Column 17 matches households to a reference basket calculated separately by expenditure-quartile. Column 18 includes an interaction with the distance between a migrant's destination region and their origin state. Column 19 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.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Online Appendix Table C7 also contains results using other nutrition measures as the dependent variable. Migrants with disadvantageous preferences are 1.4 percent more likely than local non migrants to consume fewer calories than the recommended calorie norms (a t -statistic of 2.0) and 2.6 percent more likely to consume 500 fewer calories than recommended (a t -statistic of 3.4), and these findings hold even for households with children. This set of migrants also consume 0.6 percent fewer meals (a t -statistic of 1.8) and are 0.6 percent less likely to obtain two square meals a day (but here the t -statistic is an insignificant 0.9, in part due to the smaller sample size as this question is asked only in round 38).

I perform two additional robustness tests that address particular concerns with this exercise. One concern is that the $\ln K(\cdot, \cdot)$ difference is related to the distance migrants have traveled and 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 18 of Table 7). 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 nonmigrant households (recall that $\ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)$ is the average calories per rupee for nonmigrants in the village). The average state-round sample contains 3,700 nonmigrant households and so any bias due to measurement error should be small at higher levels of disaggregation. Accordingly, column 19 calculates the $\ln K(\cdot, \cdot)$ 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 a migrant's preferences and the caloric tax, with the size of the tax related to the suitability of a migrant's origin-state preferences to the local price vector.

D. The Size of the Caloric Tax Depends on the Intensity of the Migrant Preferences

In Section IVB, I showed that household preferences for origin-state foods are more intense if both husband and wife are migrants. In this subsection, I present the corresponding result for caloric intake: that the migrant caloric tax is larger if both husband and wife are migrants as opposed to just one.

I interact the migrant terms in the caloric tax specification, equation (1), with the same set of migrant-structure dummies I used in Section IVB:

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

The β coefficients provide separate estimates for the caloric tax faced by migrants for each of these three structures. Row 2 of Table 6 reports these regression coefficients. The results mirror the findings of Section IVB. When 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 .³²

Row 3 of Table 6 performs a similar breakdown for equation (6). For each of the three migrant structures, I find that the size of the caloric tax is larger when the migrant's origin-state reference bundle provides fewer calories per rupee than the local bundle (both priced at the local price vector). As above, the tax is also

³²I 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 percent.

significantly smaller if only one of the head and spouse are migrants compared to if both are migrants.³³ 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 7.20 percent (with a 95 percent confidence interval between 5.01 and 9.39 percent).

The magnitude of this caloric tax is substantial. The median caloric intake for this migrant subgroup is 2,156 calories per person per day with 57 percent of households consuming less than the recommended calorie norms (2,400 calories rural, 2,100 urban). If these migrants had the same preferences as locals, the median would rise to 2,317 calories and the percentage of households below the norms would fall to 46 percent. To provide a mapping to economic outcomes, I draw on Schofield's (2014) estimates of the productivity decline among Indian farmers due to caloric reductions among Muslims during Ramadan. She estimates that the approximate 700 calorie decline due to Ramadan (also measured using NSS surveys) lowered agricultural productivity by between 20 and 40 percent. Crudely extrapolating this effect size, the 7.2 percent caloric tax would reduce agricultural productivity by between 4.9 and 9.8 percent. Row 4 of Table 6 reruns the regression above on agricultural households, and finds a caloric tax of 5.0 percent for households in the most affected subgroup (i.e., where both husbands and wives migrated and origin-state bundles are relatively expensive). Applying Schofield's (2014) estimates on this potentially more appropriate sample suggests a reduction in agricultural productivity of between 3.6 and 7.2 percent.

As earlier, these effects are not limited to better-nourished households. Rows 5 and 6 of Table 6 rerun the specification on the subsamples of malnourished households consuming either 250 or 500 calories less than the caloric norms. The size of the caloric tax for these households in which both husband and wife are migrants with unsuitable preferences remains a sizeable 5.2 and 5.3 percent, respectively.

In summary, I find strong evidence that culture can constrain caloric intake. I show that migrants bring 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 possessing preferences unsuited to the local price vector.

V. Alternative Explanations

I have shown that migrants bring their food preferences with them and that the size of the caloric tax is related to the suitability of a migrant's origin-state preferences to local prices. These findings are inconsistent with a story in which migrant preferences differ from those of nonmigrants 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.

³³ I can reject the null that the caloric tax for migrants with unsuited preferences is equal for these two types of households at the 1 percent level.

A. An Information Story

The first alternative explanation is that migrants have poor information about local prices or the availability and nutritional properties of local alternatives to their origin-state foods. Under these scenarios, migrants would consume fewer calories per rupee than nonmigrants as they are unaware of cheaper alternatives. Migrants may also consume bundles that more closely resemble the origin-state bundle they are more familiar with. I provide five pieces of evidence that contradict this information-based explanation.

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 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 4 present 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 IV 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, 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.³⁴ 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 a migrant caloric tax when only one of the husband or wife are migrants (row 1 of Table 6), and even when wives are moving to their husband's village (column 2 of Table 3). 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 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 in panels 1 to 3 of online Appendix Table C8 which presents similar specifications to Table 6 but breaks the "only one" category into two: only the head is a migrant and only the spouse is a migrant. I find no support for the information story. Across each of the three regressions, the ordering of the coefficients accords with the preference-driven story above (i.e., I find larger correlations and caloric taxes if the husband is a migrant as opposed to the wife). For example, the size of the caloric tax for migrants moving to locations where their origin-state bundle provides fewer calories than the local bundle is 1.9 percent

³⁴These results are shown in columns 13 to 15 of online Appendix Tables C5 and C7. For migrants who left 20 or more years ago the size of the tax is 1.66 percent (significant at the 5 percent level).

when only the spouse is a migrant and rises to 3.3 percent when only the husband is a migrant (a difference of 1.4 percent).³⁵ I also allow the size of the tax to differ by the migrant's region of origin in online Appendix Table C9. Consistent with the stylized fact that there is a north-south gradient in both patrilocality and attitudes toward women, the difference between the migrant-husband and migrant-spouse caloric tax is largest in the north (a difference of 2.8 percent), smallest in the south (a difference of -0.2 percent), and of intermediate value in the center of India (a difference of 1.4 percent).

One possible explanation for these findings is that the husband's mother is in charge of household food purchases and preparation and that husbands bring their mothers 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 results are essentially unchanged when these households are excluded.³⁶

Fourth, if the explanation is that migrants have poor information, the caloric 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 Table 4 restricts 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 (i.e., 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 $[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^o) - \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$. I report this regression in column 12 of Table 5. I find a significantly negative coefficient on the $migrant_i^{ov} \times 1[\ln K(\mathbf{bshare}_i^o, \mathbf{P}_i^o) < \ln K(\mathbf{bshare}_i^v, \mathbf{P}_i^v)]$ interaction term 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 as these migrants still consume bundles that more closely resemble their origin-state bundle (i.e., the sum of the coefficient on the interaction and the coefficient on $migrant_i^{ov}$ is still positive and significantly different from zero at the 1 percent level).

B. A Technology Story

The second alternative explanation is that migrants do not possess the technologies to make high-quality meals using 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, if a family migrated

³⁵I can reject the null that the only spouse caloric tax is equal or greater to the only head caloric tax at the 5 percent level.

³⁶Panels 4 to 6 of online Appendix Table C8 replaces the "only head is migrant" dummy with "only head is migrant and no mother in house."

from a wheat to rice area, they may continue to consume wheat-based meals as they enjoy well-prepared meals over badly-prepared meals (rather than wheat over rice).

Once more the evidence contradicts a story in which technology is the sole explanation for my findings. If technology was responsible, the caloric 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 moving into a nonmigrant household since these households should have appropriate kitchen equipment 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 only the wife. 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.

VI. Conclusion and Policy Implications

This paper sets out to answer a simple question: can food cultures constrain caloric intake and contribute to malnutrition? I address this question by exploiting the fact that migrants and nonmigrants face the same relative prices, yet possess very different preferences. Drawing on detailed household survey data from India, I document that interstate migrants consume fewer calories per rupee of food expenditure compared to their nonmigrant neighbors, 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 calorically 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 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.

Before discussing policy implications, it is important to address the fact that the paper uses data from the 1980s and so the findings may no longer be relevant for policymakers today. (Recall more recent data do not contain detailed consumption records matched to migration particulars.) There are several reasons to think the findings are still relevant. The fact that similar observations about India were made by the Famine Inquiry Commission in 1945 suggests these forces are persistent. Indeed, regional consumption patterns change little across NSS survey rounds. For example, I calculate the coefficient of variation of state average food-budget shares for each of the 135 foods that are common across surveys. The mean of these 135

coefficients of variation is 0.75 in 1983, 0.67 in 1987–1988, and 0.61 in 2009–2010.³⁷ In fact, given the increase in interstate migration in India (the percentage of the population born in a different state rose from 3.3 to 4.1 between 1991 and 2001, the last available census), the caloric tax culture imposes may be becoming more important not less. Finally, there is also contemporary evidence for another country, the United States, that migrants bring their preferences with them (Bronnenberg, Dube, and Gentzkow 2012).

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: providing food aid or price subsidies to consumers; reducing tariffs or using other trade policies to increase food imports; and developing biofortified or high-yielding variety (HYV) crops. In all three cases, 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.³⁸ However, much food aid to Africa comes in the form of imported yellow maize, and vitamin-A biofortification involves the addition of carotenes which turn maize yellow-orange. Programs that provide cheap yellow maize to hungry communities, or try to reduce vitamin A deficiency through biofortified 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 biofortification of traditional foods may prove more successful in such cases. Similarly, the introduction of HYV 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 (see Paarlberg 2010), adoption of these HYV crops was held back by strong local preferences for sub-Saharan staples such as sweet potato, cassava, sorghum, and white maize.

Another potential remedy, and one mooted by the 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 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. ... Alternative cereals could be used for school meals. ... 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

³⁷I define states using 1983 state borders. For unit values, the equivalent numbers are 0.43, 0.44, and 0.49, respectively.

³⁸See McCann (2005) for the historical origins and De Groote and Kimenju (2008) for empirical evidence of this preference.

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.

REFERENCES

- Alesina, Alberto, and Paola Giuliano.** 2015. "Culture and Institutions." *Journal of Economic Literature* 53 (4): 898–944.
- Almond, Douglas, and Janet Currie.** 2010. "Human Capital Development Before Age Five." National Bureau of Economic Research Working Paper 15827.
- Almond, Douglas, and Bhashkar Mazumder.** 2011. "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy." *American Economic Journal: Applied Economics* 3 (4): 56–85.
- Atkin, David.** 2013. "Trade, Tastes, and Nutrition in India." *American Economic Review* 103 (5): 1629–63.
- Atkin, David.** 2016. "The Caloric Costs of Culture: Evidence from Indian Migrants: Dataset." *American Economic Review*. <http://dx.doi.org/10.1257/aer.20140297>.
- Birch, Leann L.** 1999. "Development of Food Preferences." *Annual Review of Nutrition* 19 (1): 41–62.
- Bronnenberg, Bart J., Jean-Pierre H. Dube, and Matthew Gentzkow.** 2012. "The Evolution of Brand Preferences: Evidence from Consumer Migration." *American Economic Review* 102 (6): 2472–2508.
- Carroll, Christopher D., Byung-Kun Rhee, and Changyong Rhee.** 1994. "Are There Cultural Effects on Saving? Some Cross-Sectional Evidence." *Quarterly Journal of Economics* 109 (3): 685–99.
- Deaton, Angus, and Jean Drèze.** 2009. "Food and Nutrition in India: Facts and Interpretations." *Economic & Political Weekly* 44 (7): 42–65.
- De Groot, 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): 362–70.
- Desai, Sonalde, Amaresh Dubey, B. L. Joshi, Mitali Sen, Abusaleh Shariff, and Reeve Vanneman.** 2005. *India Human Development Survey*. <http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/22626> (accessed August 28, 2015).
- Desai, Sonalde, Amaresh Dubey, B. L. Joshi, Mitali Sen, Abusaleh Shariff, and Reeve Vanneman.** 2011. *India Human Development Survey*. <http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/36151> (accessed August 28, 2015).
- Famine Inquiry Commission.** 1945. *Report on Bengal*, 176–78. New Delhi: Government of India. <http://www.bowbrick.org.uk/Single%20page%20fic/FIC175-end.pdf> (accessed February 8, 2016).
- Fernández, Raquel.** 2011. "Does Culture Matter?" *Handbook of Social Economics*, Vol. 1, edited by Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, 481–510. Amsterdam: North-Holland.
- Fernández, Raquel, and Alessandra Fogli.** 2009. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility." *American Economic Journal: Macroeconomics* 1 (1): 146–77.
- Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti.** 2004. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." *Quarterly Journal of Economics* 119 (4): 1249–99.
- Food and Agriculture Organization of the United Nations.** 2008. "Food Consumption Statistics 2006–2008." http://www.fao.org/fileadmin/templates/ess/documents/food_security_statistics/FoodConsumptionNutrients_en.xls (accessed February 8, 2016).
- Giuliano, Paola.** 2007. "Living Arrangements in Western Europe: Does Cultural Origin Matter?" *Journal of the European Economic Association* 5 (5): 927–52.
- Goel, Mita Sanghavi, Ellen P. McCarthy, Russell S. Phillips, and Christina C. Wee.** 2004. "Obesity Among US Immigrant Subgroups by Duration of Residence." *Journal of the American Medical Association* 292 (23): 2860–67.
- Guendelman, Maya D., Sapna Cheryan, and Benoît Monin.** 2011. "Fitting In but Getting Fat: Identity Threat and Dietary Choices Among U.S. Immigrant Groups." *Psychological Science* 22 (7): 959–67.
- 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*. New York: Simon and Schuster.

- International Institute for Population Sciences.** 1995. "National Family Health Survey, India 1992–93." <http://rchiips.org/nfhs/>.
- Logan, Trevon D., and Paul W. Rhode.** 2010. "Moveable Feasts: A New Approach to Endogenizing Tastes." http://www.econ.ohio-state.edu/trevon/pdf/TF_08-05-10.pdf (accessed February 8, 2016).
- McCann, James.** 2005. *Maize and Grace: Africa's Encounter with a New World Crop, 1500–2000*. Cambridge, MA: Harvard University Press.
- National Sample Survey Organization.** 1983–1988. *Indian National Sample Surveys*, Rounds 38 and 43. <http://mospi.nic.in/>.
- Nunn, Nathan, and Nancy Qian.** 2011. "The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment." *Quarterly Journal of Economics* 126 (2): 593–650.
- Paarlberg, Robert.** 2010. *Food Politics: What Everyone Needs to Know*. Oxford, UK: Oxford University Press.
- Prakash, Om.** 1961. *Food and Drinks in Ancient India: from Earliest Times to c. 1200 AD*. Delhi: Munshi Ram Manohar Lal.
- Schofield, Heather.** 2014. "The Economic Costs of Low Caloric Intake: Evidence from India." <http://www.cgdev.org/sites/default/files/Schofield%2C%20Calories%20and%20Productivity%2C%202014.01.27%28new%29.pdf> (accessed February 8, 2016).
- Srinivas, M. N.** 1980. *India: Social Structure*. New Brunswick, NJ: Transaction Publishers.
- Staehle, Hans.** 1934. "The Reaction of Consumers to Changes in Prices and Income: A Quantitative Study in Immigrants' Behavior." *Econometrica* 2: 59–72.
- Subramanian, Shankar, and Angus Deaton.** 1996. "The Demand for Food and Calories." *Journal of Political Economy* 104 (1): 133–62.