Trade, Tastes and Nutrition in India*

David Atkin†

November 2009

Abstract

This paper introduces habit formation into an otherwise standard model of international trade. Household tastes evolve over time to favor foods consumed as a child. The opening of trade causes preferred goods to rise in price, as these were relatively inexpensive in autarky. Neglecting the correlation between tastes and agro-climatic endowments overstates the short-run nutritional gains from agricultural trade liberalization and masks potential caloric losses for laborers. I examine the predictions of this model of trade with habit formation using household survey data from India, both by looking across Indian regions and by examining the consumption patterns of inter-state migrants.

*Thanks to Richard Chiburis, Angus Deaton, Dave Donaldson, Adrian de Froment, Scott Fulford, Penny Goldberg, Marco Gonzalez-Navarro, Gene Grossman, Gordon Hanson, Giovanni Maggi, Chris Paxson, Jesse Rothstein, Sam Schulhofer-Wohl and seminar participants at Princeton, Duke, UCSD, Kellogg M&S, Yale, MIT, LSE, Stanford, Chicago Booth, Brown, UCLA Anderson, CREI, Munich, ULB, Oxford, the 2009 SITE at Stanford, Columbia and the Federal Reserve Bank of Minneapolis. Financial aid from the Fellowship of Woodrow Wilson Scholars and the International Economics Section at Princeton is gratefully acknowledged. The Center for Economic Policy Studies at Princeton is thanked for purchasing the Indian Harvest Data.

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

Abstract

This paper introduces habit formation into an otherwise standard model of international trade. Household tastes evolve over time to favor foods consumed as a child. The opening of trade causes preferred goods to rise in price, as these were relatively inexpensive in autarky. Neglecting the correlation between tastes and agro-climatic endowments overstates the short-run nutritional gains from agricultural trade liberalization and masks potential caloric losses for laborers. I examine the predictions of this model of trade with habit formation using household survey data from India, both by looking across Indian regions and by examining the consumption patterns of inter-state migrants.

1 Introduction

The impacts of agricultural trade on the developing world are of great interest to economists and policymakers alike. For example, the Doha round of global trade negotiations aims to both substantially liberalize agricultural trade and ensure that trade brings greater benefits to developing countries. At the same time, surging demand from emerging markets has pushed up the traded prices of many food crops. Even in food-exporting countries, policymakers have raised concern about the potentially adverse effects of these price changes on certain consumers. In this paper, I explore a new channel which provides a more complete understanding of the nutritional impacts of agricultural trade on the poor through examining the role of tastes in international trade, a subject that has been neglected in the literature to date.

Standard international trade theory typically focuses on the production side, rather than the consumption side, and assumes that preferences are identical across regions and independent of resource endowments.[1] This paper explores the consequences of relaxing this assumption in a very natural way. I incorporate habit formation into an overlapping generations model of trade and demonstrate that this leads to regional food tastes that favor crops relatively well-suited to local agro-climatic endowments. This correlation between tastes and

[1] International trade textbooks mention that taste differences can be one source of comparative advantage, however the literature dealing with cross-country taste differences only explores the consequences of non-homothetic preferences, for example Hunter and Markusen (1987) or more recently Fieler (2009).

endowments systematically reduces the short-run nutritional gains from trade compared to models without habit formation. Using household consumption data from regions within India, I provide two sets of empirical evidence that support the model's predictions and I quantify the extent to which regional taste differences will alter the nutritional impact of trade if India were to liberalize its highly restricted internal agricultural markets.

In particular, I define tastes for food as the component of the food budget share that cannot be explained by the vector of prices or total food expenditure in a demand system. Based on extensive evidence in the nutrition literature, surveyed in Birch (1999), I assume that adult tastes favor the foods consumed as a child and term this process habit formation. While large literatures explore the implications of habit formation for demand, asset pricing, monetary policy, growth and addiction, this paper is, as far as I am aware, the first exploration of how habit formation in consumption alters standard models of international trade.

The first generation of adults, who value only calories and dietary variety, purchase large quantities of a region's relatively abundant (comparative advantage) agricultural goods, as these foods are relatively cheap under autarky. Their children are fed these locally abundant foods, and develop particular tastes for them in adulthood. Over many generations, a home bias in household consumption emerges endogenously through habit formation. The same affinities for local foods emerge if recipes and preparation techniques improve faster for commonly consumed foods. Households will choose to purchase the familiar local foods that they know how to transform into high-quality meals.

At the time of trade liberalization, the preferred foods rise in price in each region as these foods were relatively inexpensive in autarky, and trade equalizes prices across regions. Consumers spend a large portion of their incomes on these favored foods and are reluctant to substitute out of them and into less familiar imports, which reduces the consumption gains from trade compared to a model without habit formation. Only decades after trade liberalization can consumers realize the full caloric gains from trade, as food tastes gradually adapt, eventually resulting in even larger caloric gains than a model without habit formation would predict.

The aggregate production gains from trade also shrink with habit formation because habits reduce the autarky relative price differences between regions.[2] As in a standard trade model, these reduced production gains will be distributed unevenly. If labor is mobile and combined with crop-specific land to produce a food, the specific factors model implies that labor's nominal wage gains from trade will be strictly smaller than the price rise in the locally abundant food. Therefore, although there are still aggregate caloric gains from trade in this model—albeit smaller than the gains in a standard trade model—trade can spell short-run nutritional losses for landless laborers. This is because habit formation results in this group spending a large portion of their budget on the local staple whose price rises more than their income.[3]

The recent surge in world food prices provides supportive evidence of this link between tastes, trade and nutrition. In March 2006, Argentina banned all exports of beef in order to lower the domestic price for beef-loving Argentinians. In 2008, Vietnam, Cambodia and Egypt banned rice exports; Russia, Kazakhstan and Ukraine banned wheat exports; while Zambia and Malawi banned maize exports. In all of these cases, important agricultural exports are the preferred calorie sources of low-income households. Trade-induced price rises increased hunger among poor consumers, and governments explicitly tried to help this group by restricting trade and bringing prices of favored foods back towards autarky levels. In contrast, a standard trade model without habit formation would recommend transferring income from exporters, and allowing poor consumers to substitute into relatively cheap foods.

I test my theory more directly by using detailed household survey data from India. India contains many agro-climatic zones and extremely varied diets. At the same time, India maintains extensive internal food-trade barriers in addition to a poor transport infrastructure. These barriers include tariffs at state borders, numerous licensing requirements for traders and physical movement restrictions. Despite much publicized economic reforms in the early

[2] Habit formation has allowed each region to exploit some of the gains from specialization in previous generations, thereby reducing the production gains possible at the moment of trade liberalization.

[3] The model may not be relevant to households on the edge of starvation, who will presumably maximize caloric intake. However, the average rural Indian household consumes 300 calories less than the recommended daily intake of 2,400 calories per day (2004/5 NSS data). As 2,400 calories could be purchased for forty percent of per-capita daily food expenditure yet half the children in rural India are underweight (2006 NFHS data), many malnourished households seem not to be maximizing nutrition alone.

1990's, agriculture continues to be subject to enormous state intervention in the name of food self-sufficiency and agricultural markets are not integrated. Accordingly, my empirical work treats Indian regions like many small partially closed economies, and provides an excellent opportunity to test the autarky predictions of my model.

In order to investigate the relationship between tastes and local resource endowments across India, I estimate the tastes defined as above with the further restriction that the price and income effects are common across India. I regress household demand for agricultural products on a set of regional dummies and a flexible set of common price and expenditure terms. The regional component of the unexplained variation in budget shares spent on each food then provides my main taste measure.

It is impossible to observe the impacts of trade liberalization in societies both with and without habit formation. Therefore, I use rural household survey data from 77 agro-climatic regions within India to provide empirical evidence for the mechanisms in my model that reduce the caloric gains from trade on the consumption side. In the first stage, I show that regions have stronger tastes for the foods that their agro-climatic endowments are relatively well-suited to producing, and that these foods are inexpensive compared to other regions. To highlight the role of habits, I confirm that the ordering of tastes within a region responds to relative price changes in the past. Therefore, if India were to liberalize internal trade, each region's more favored foods will be expected to rise in relative price as regional prices converge to a uniform national price. In the second stage, I verify that these expected relative price rises in more favored foods will negatively impact nutrition by showing that between 1987 and 2005, caloric intake declined more in regions where (non trade-induced) price rises were more concentrated in locally favored foods, controlling for changes in food expenditure.

I confirm my findings from looking across Indian regions using a second approach. Interstate migrants mimic small economies opening to trade, since migrants bring their labor endowment and preferences but face a new set of prices. Therefore, moving should affect migrants similarly to how my model predicts trade liberalization affects consumers. I show that inter-state migrants in India do carry their food tastes with them, consuming food bundles less similar to those consumed in their destination state and more similar to those consumed in their origin state. Migrant households consume fewer calories for a given level of food expenditure, because they continue to buy favored products from their origin state that are now relatively expensive. This effect dissipates with time, disappearing only four generations after migration. Finally, mirroring the effects of temporal price variation within regions, I find that for the 484 observed migration routes, the caloric intake from a given level of food expenditure declines more where the relative price rises faced by a migrant are more concentrated in that migrant's preferred origin-state foods. These results even hold when I restrict attention to households in which the household head's wife moved for marriage, and compare households where the wife moved inter-state to those where the wife moved intra-state.

With these two sets of evidence in place, I quantify the caloric impact of Indian agricultural trade liberalization, and the error associated with ignoring the correlation between tastes and endowments that comes about through habit formation. If India were to liberalize its internal agricultural trade (when food prices converge to the Indian median price), trade will have to generate income gains of at least 3.3 percent for the average Indian household to maintain their pre-liberalization caloric intake. No such income gains would be necessary if tastes were identical across India. Poorer regions, which consume larger shares of the local staple foods predicted to rise in price, will require even larger income gains to maintain caloric intake. Household incomes are likely to increase with liberalization through increased specialization in food production.[4] However, the 3.3 percent increase required to avoid caloric losses is larger than existing estimates of the static income gains for typical rural households from other agricultural trade liberalization scenarios.[5]

In this paper, I focus on the nutritional impacts of trade, measured through caloric intake,[6]

[4]For poor Indians, these gains will primarily come through nominal wage gains or falls in the average food price, since most rural households derive the majority of their income from labor.

[5]The most direct comparison is China, where reductions in caloric intake over the reform period (Du et al., 2006) were accompanied by the dismantling of extensive barriers to internal agricultural trade.

[6]Caloric intake is not the same as nutrition, with vitamin and protein intake also important. However, as there is no single agreed-upon nutritional metric, in this paper I focus only on calories.

although my theoretical results apply to welfare in the isomorphic model where preparation techniques improve with past consumption. In few countries is malnutrition a more important issue than in India, which has a higher prevalence of undernutrition than Sub-Saharan Africa (Deaton and Dreze, 2008). Food consumption itself has regularly been used as a poverty measure. For example, poverty lines for countries as diverse as the US and India initially derive from the amount of money required to meet basic caloric needs.

There are several reasons why economists should be directly concerned about poor nutrition. Sen (1999) has argued that improving the health of the poor and increasing their capabilities should be a goal of development in itself. Low caloric intake directly reduces productivity by reducing energy levels, health capital and the ability of the immune system to fight off infectious disease. These effects exert externalities on other members of society. Malnourished populations allow contagious diseases to spread more readily, and Fogel (1994) has argued that improved nutrition and its synergies with technological advance can account for much of the economic growth seen in the West since the Industrial Revolution.

Policymakers often cite explicitly paternalistic concerns. Many of the gains from proper nutrition come through good health later in life, which uninformed consumers may undervalue. Barker (1992) and others have demonstrated the substantial scarring effects of nutritional shortfalls at young ages on productivity, earnings and health in adulthood. Accordingly, even the short-term nutritional declines that can occur during an episode of trade liberalization are of serious concern, because an entire generation malnourished as children will continue to suffer irremediable consequences for the rest of their lives.

In section 2, I provide a diagrammatic discussion of the theory, with the formal proofs relegated to appendix C. Section 3 introduces the data and my taste estimates. In section 4, I investigate variations in tastes, prices, endowments and caloric intake across 77 regions of India. Section 5 uses data on inter-state migrants within India to confirm the regional results. Section 6 discusses India's internal trade restrictions and predicts the caloric impact of relaxing these restrictions. Finally, section 7 concludes and discusses the policy implications.

2 Theoretical Background

2.1 Modelling Habit Formation in Food Consumption

I start by describing consumer preferences in an overlapping generations model that features habit formation in food tastes. I then justify my assumptions about the form of the utility function by reviewing the literature on the development of food preferences. Finally, I analyze this simple economy in autarky.

Identical individuals in a small region live for two periods, childhood and adulthood. In the second period, individuals obtain factors of production, spend their full income from these factors and have a single child. Adults in generation t choose their relative consumption of the two goods in the economy, rice, r, and wheat, w, both measured in calories. The child and the parent share the parent's preferred consumption bundle, and form a single household.[7]

I model household demand as follows. The budget share spent on rice is a function $h_r(.)$ of relative prices, p_r/p_w, total (food) expenditure, $food$, and a rice taste shifter, $tastes_r$:[8]

$$bshare_{rt} = tastes_{rt} + h_r(\frac{p_{rt}}{p_{wt}}, food_t). \tag{1}$$

In the next period the child grows up and the bundle that he or she consumed as a child influences his or her adult preferences. I will call this habit formation, with an adult developing tastes for the foods of which he or she consumed relatively more as a child.[9] Specifically,

$$tastes_{rt} = g(\frac{r_{t-1}}{w_{t-1}}), \text{ with } \frac{\partial tastes_{rt}}{\partial(\frac{r_{t-1}}{w_{t-1}})} > 0.$$

Therefore, the adult's utility function depends on both past and present consumption, as is common in the habit formation literature starting with Stone (1956) and including Becker and Murphy (1988). Additionally, I assume that parents are myopic, and so do not antici-

[7] Specifically, parents gain equal utility from their own and their child's consumption. Parents are not altruistic, or are uninformed about the importance of childhood nutrition, as they choose their child's diet based solely on their own preferences.

[8] The budget share spent on rice increases with the tastes for rice and decreases with the relative rice price. The Cobb-Douglas utility function generates household demands of this type, with $h_r(.)$ equal to zero. The main results from my paper carry through with a more general demand specification, $bshare_{rt} = \widetilde{h}_r(\frac{p_{rt}}{p_{wt}}, food_t, tastes_{rt})$, as long as $\partial bshare_{rt}/\partial tastes_{rt} > 0$.

[9] Habit formation may occur more quickly than this, which can be accommodated by including several stages of adult life, with past consumption influencing current preferences. For example, preferences for eating raw fish in sushi have developed rapidly in the West over a single generation.

pate the effect of their consumption choices on the future tastes of their child. Accordingly, household demand in period t does not depend on expected future prices or incomes.[10]

Strong and enduring taste patterns characterize food consumption. Ample evidence in the psychology and nutrition literatures indicates that certain food preferences form in childhood. Children have a predisposition to fear new foods, which is only overcome through repeated opportunities to consume a food (Birch, 1999). The literature hypothesizes that this response serves a protective function, so that foragers can learn what foods are safe to eat. This is common across omnivores, and has been shown in controlled experiments among both humans and rats. More directly, a mother's diet during pregnancy and lactation affects her child's preferences for flavors and foods in later life (Mennella et al., 2001).

Social factors also play an important role in forming preferences. There is abundant evidence, again from controlled experiments involving humans and other mammals, that the young are more likely to accept new or disliked foods if they observe their mothers or other role models consuming them (Birch, 1999). This effect works through two channels; role models both induce children to try a food for the first time which overcomes the initial neophobia, and provide a pleasant context in which the food is eaten. As an example of the latter channel, second generation US immigrants from India may enjoy eating curry more than their non-Indian peers in part because of the positive association that has formed between consuming Indian food and dining at home with family members.

Crucially for my assumptions about habit formation, preferences gained in childhood persist in the available longitudinal data. Data from the Minnesota Heart Health Program (Kelder, 1994) show that food preference rankings remain unchanged over 6 years.[11] Therefore, this extensive set of evidence supports my assumption that food preferences are positively related to the consumption patterns of the previous generation.

In order to model the production side of the economy, I use a specific factors model that

[10]Even if this assumption is violated, parents will only partially adjust their child's diet if there is uncertainty over the prevailing relative prices when their child reaches adulthood, or if there are costs to preparing two separate meals for both themselves and their child.

[11]All of the sample faced the same relative prices, yet there were substantial and persistent variations in food rankings which can not be explained by the common price effects.

matches the realities of food production well. A region has a fixed endowment of laborers, L, and two additional factors: land suitable for rice cultivation, T_r, and land suitable for wheat cultivation, T_w. Both rice and wheat are produced with constant returns to scale technology using labor and the single specific land factor, with diminishing returns to an increase in any one factor.[12] To grow wheat, a farmer requires well-drained soil, low humidity and moderate temperatures. Rice grows most easily in coastal plains, lowland deltas and tidal plains where paddies are submerged in water. These agro-climatic conditions are fixed over time and can explain why the arid plains of Rajasthan produce mainly wheat and West Bengal in the Ganges Delta produces mainly rice.

There is no migration[13] and factor endowments are fixed over generations. High transport costs and trade restrictions imply an initial equilibrium with autarky. In addition, there is no storage technology, so that all food is consumed in the period that it is produced.

Figure 1 describes the autarkic equilibrium. I first plot the production possibilities frontier (PPF), the locus of the maximum feasible combinations of wheat and rice that can be produced using a region's endowment and a technology that is identical across all regions. The figure shows a bowed out PPF[14] for the home region which has a relatively larger endowment of rice land than wheat land. I represent preferences by an aggregate indifference curve for the whole economy. I assume that in the first generation, adult consumers have "neutral" tastes only for calories and dietary variety, corresponding to $tastes_{r1} = 1/2$ and a indifference curve that is perpendicular to the 45 degree line at any half-rice half-wheat bundle.[15]

Consumption occurs at point A_1, with relatively more rice than wheat consumed. Since rice land is abundant in the economy, labor flows into rice production to equalize the wage across sectors, increasing the relative production of rice. This higher relative production induces a drop in the rice price to equilibrate supply and demand and so $p_{r1}/p_{w1} < 1$.[16] The

[12] All of the main results carry through to the Heckscher-Ohlin model where there are only two factors, T_r and T_w, which are required to produce both goods, but rice is relatively more intensive in factor T_r.

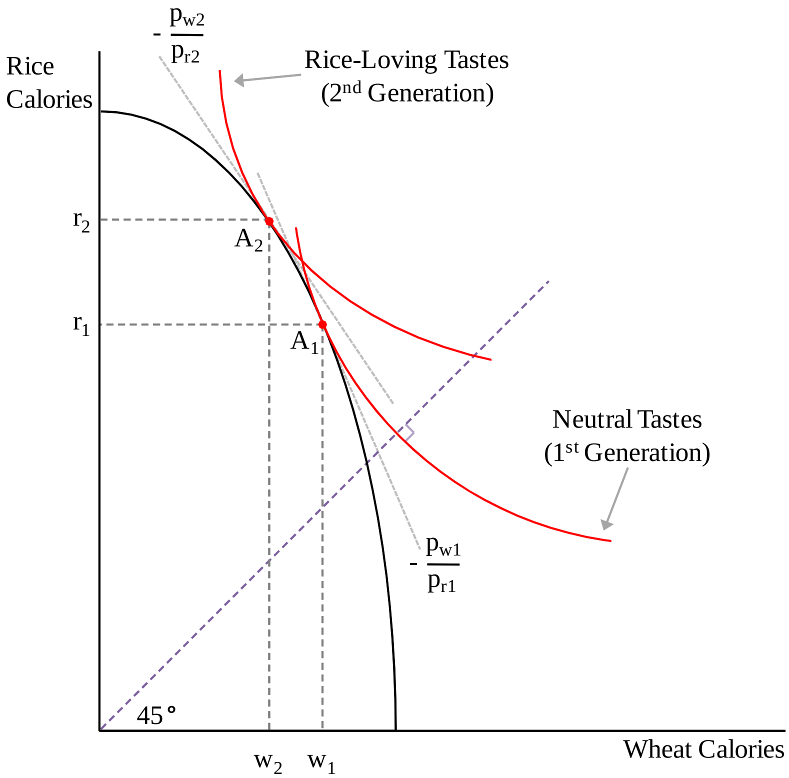
[13] This is a reasonable assumption in India, where there is very little migration (Munshi and Rosenzweig, 2009). In section 5, I study the small population of inter-state migrants.

[14] The PPF bows out due to diminishing returns to each factor in the rice and wheat production functions.

[15] In the representation shown, consumers have a taste for variety and not simply for calories.

[16] The relative price is equal to $-dw/dr$ at the point where the indifference curve is tangential to the PPF.

Figure 1: Habit Formation in a Two-Good Economy



full equilibrium is described in appendix C.1.

Tastes develop through habit formation, and so when the children of the first generation reach adulthood they have stronger tastes for rice than their parents as they consumed relatively more rice than wheat in their youth ($tastes_{r2} > tastes_{r1}$). I label the preferences of the second generation "rice-loving tastes" in figure 1. This generation has an increased relative demand for rice. As the endowment is fixed, the rice price must rise in order to induce an increase in the supply of rice and bring the market back into equilibrium ($p_{r2}/p_{w2} > p_{r1}/p_{w1}$). The second generation adults consume at A_2, with relative rice consumption even higher than in the previous generation ($r_2/w_2 > r_1/w_1$).

Tastes for rice will continue to increase with each generation until generation s, when the price rise induced by the increased demand for rice is sufficiently large to leave relative rice consumption unchanged, $d(r_s/w_s)/d(r_{s-1}/w_{s-1}) = 0$. An interior steady state may not exist if habit formation is so strong that the price response is never large enough to negate fully the consumption increase. In appendix C.5, I explore the conditions for the existence of an interior steady state in the case of Cobb-Douglas consumption and production functions. Such a steady state exists and is stable, with rice remaining cheaper than wheat ($p_{rs}/p_{ws} < 1$), as long as there is a sufficient love of variety (complementarity in consumption) so that tastes do not respond excessively to consumption changes.

It is an empirical question whether habit formation is sufficiently strong to make the good produced using the abundant factor relatively more expensive under autarky ($p_{rs}/p_{ws} > 1$). I will provide supportive empirical evidence for India that the price of a particular food is indeed cheaper in regions where resource endowments are relatively well-suited to growing that food crop. The same is true for the many countries listed in the introduction which imposed export bans. Accordingly, in the exposition that follows I will assume that an interior autarky steady state exists and rice remains relatively cheaper than wheat if rice land is more abundant in the region, $T_r > T_w$, and term this Assumption 1.[17]

Hypothesis 1: *A region develops tastes inversely related to the relative prices it faces. Therefore, tastes will become positively correlated with a region's relative resource endowments.* Proof under the assumption of homothetic preferences in appendix C.1.

Habit formation leads to something similar to the much used Armington (1968) home-biased preferences,[18] where preferences favor locally-produced varieties of a given good by assumption. Here, biased preferences for local goods are not ad hoc but emerge endogenously from endowments and are far more plausible for non-differentiated goods, where the region of origin cannot be inferred from the good itself.

Nutrition improves in the second generation. My nutrition metric is caloric intake, which is simply $r + w$. Accordingly, the isocalorie line is perpendicular to the 45 degree line. The second generation more readily consumes the abundant calorie source, rice. Therefore, as long as rice remains cheaper than wheat, the second generation ends up better nourished.[19] In

[17]Precise conditions appear in appendix C.2. If rice does not remain relatively cheap, then habit formation has changed which goods each region has a comparative advantage in, and more favored goods will fall in price upon liberalization.

[18]Armington home-biased preferences are used in the empirical trade literature to explain home-bias effects found in trade data as well as in most modern Computable General Equilibrium trade models.

[19]In the model, calories are not a production input. Changing tastes may lower net nutrition by inducing

each subsequent generation, caloric intake will further increase until a steady state is reached.

Hypothesis 2: *The next generation will consume a larger total quantity of calories if tastes develop to favor the relatively cheap calorie source, as long as that calorie source remains relatively cheap.* Proof in appendix C.3.

The fact that caloric consumption increases with habit formation provides an evolutionary justification for a utility function that depends on past relative consumption. Habit formation would evolve endogenously from a simple game-theoretic evolutionary model. Societies exhibiting such traits will be better nourished and hence fitter in an evolutionary sense, making them able to outcompete other groups.

2.2 Opening the Economy to Trade

Comparing changes in welfare upon trade liberalization between two societies, one with and one without habit formation, is not possible if the preferences of the two societies differ. Accordingly, I restrict my focus to analyzing caloric intake rather than welfare. Section 2.5 details an isomorphic model where equivalent welfare statements can be made as consumers have fixed preferences for food quality, and the quality level of a particular food depends on recipes and preparation techniques that develop alongside past consumption.

What happens when this small region liberalizes trade after many generations under autarky? Trade liberalization generally takes place in waves over several years, but tastes change only across generations. Therefore, I evaluate the short-run impact of a trade liberalization at time T, as shown in the timeline below, with tastes held fixed at their pre-liberalization values. I compare the case where tastes in the region favor the relatively abundant good, rice, against the case in which tastes are still neutral and independent of endowments. As I have shown in the previous section, this positive correlation between preferences and endowments arises naturally from a model of habit formation, although the following trade implications hold whenever such a positive correlation exists.

a movement of labor into more physically demanding sectors. Although jobs in India have become less physically demanding (Deaton and Dreze, 2008), it is unlikely that the resource comparative advantage sectors are systematically more physically intensive, as would be required to reverse my results.

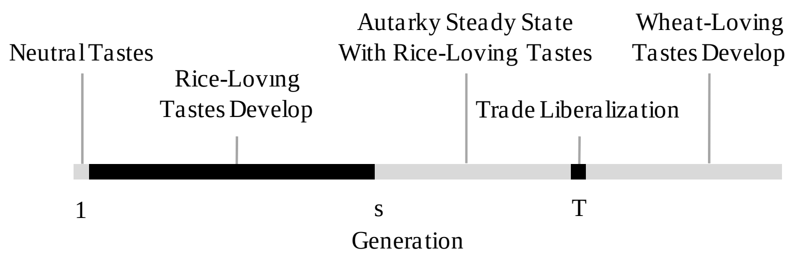
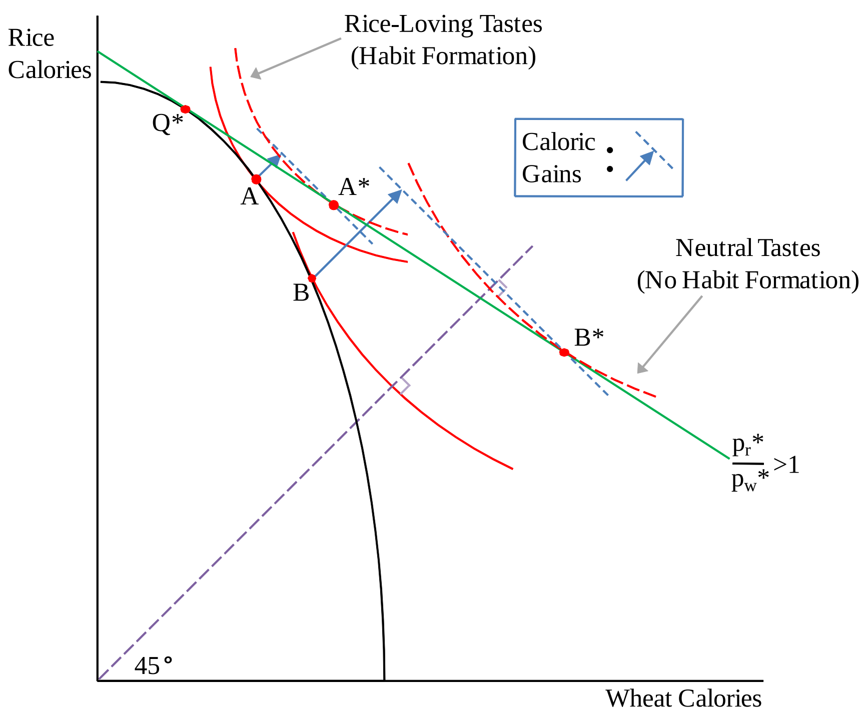


Figure 2 shows the new equilibrium for the region after opening to trade with a world where wheat is relatively cheap (the world price line has a gradient less than -1). Post trade values are denoted by asterisks. With trade, the region can separate its consumption and production decisions and produces at point Q*, and accordingly there are standard gains from trade through specialization in rice production.

Figure 2: Trade Liberalization With and Without Habit Formation



Without habit formation, when tastes for rice and wheat are neutral, the region consumes at point B under autarky and B* after liberalization. Caloric intake increases by the vector length of the arrow from point B to the isocalorie line passing through B*. With habit formation, tastes in the region favor rice and consumption moves from A to A*. The aggregate caloric gains from trade are much smaller (the vector length of the arrow originating at

A). Aggregate caloric consumption with rice-loving tastes is smaller post trade liberalization compared with neutral tastes, but was larger pre trade (hypothesis 2), implying that the short-run caloric gains from trade shrink with habit formation.

Hypothesis 3: *Habit formation reduces the short-run aggregate caloric elasticities with respect to trade liberalization.* Proof under assumption 1 for $p_r^*/p_w^* \geq 1$ in appendix C.4.

In the generations following trade liberalization, the taste for rice will decline as relative wheat consumption rises. This process produces further caloric gains for future generations, as they spend an increasingly large share of their budget on wheat, the relatively cheaper calorie source post trade. After many generations wheat loving preferences develop and the aggregate caloric intake will actually exceed that with neutral tastes.[20] However, the effects on the current generation are of primary importance to elected policymakers, and accordingly I focus on these initial impacts in the paper.

The reduction in the aggregate caloric gains from trade derives from both the consumption and production sides of the economy. On the production side, the correlation between preferences and endowments brought about by habit formation brings autarky prices closer together by bidding up the price of the region's relatively abundant food, thereby reducing the gains from specialization at the moment of trade liberalization. However, it is the consumption side that will be the focus of my empirical work, and in the next section I will highlight the consumption effects by looking at individual consumption, holding income fixed.

2.3 Consumption Side Impacts of Trade

I can illustrate more clearly the consumption side effects of trade liberalization in a society with habit formation by analyzing the effect of price changes on individuals, as opposed to the aggregate effects described previously in a general equilibrium setting. My empirical data contain consumers with a range of factor endowments, whose income gains from trade will vary. Therefore, to motivate my empirical strategy, I analyze how tastes alter caloric gains from trade on the consumption side in partial equilibrium by holding incomes constant.

[20]This economy is small. For two large symmetric economies opening to trade, the post trade price will be 1, and long run caloric intake will be identical with and without habit formation.

Figure 3: The Effect of Price Changes on Individual Calorie Consumption

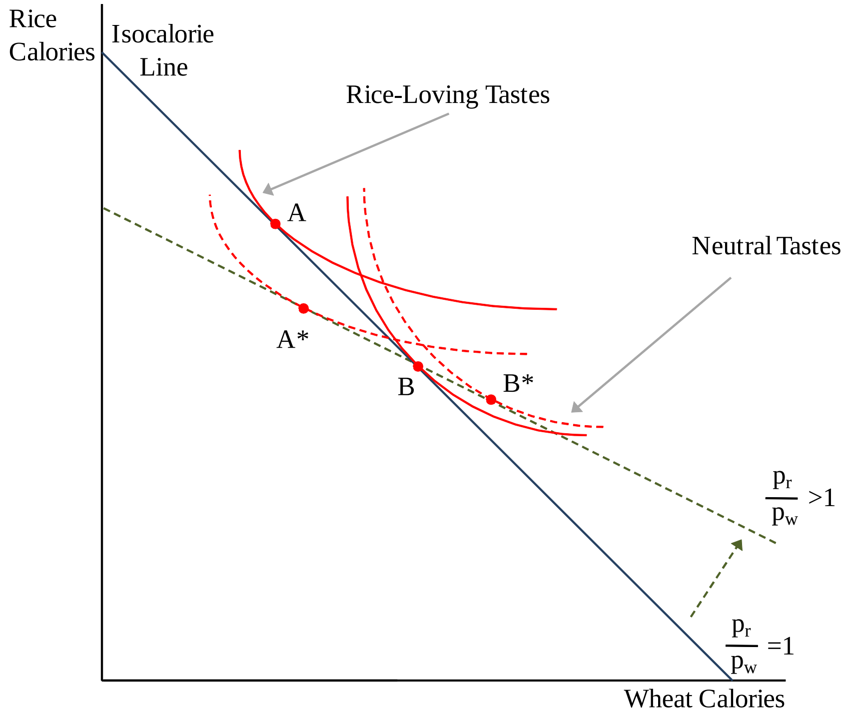


Figure 3 shows the consumption impact of an exogenous rise in the relative price of rice for an individual factor owner, holding income constant. For simplicity, I assume that one calorie of either wheat or rice can be purchased for the same price prior to the price change. At this price the consumer will have the same caloric intake regardless of his or her relative tastes for rice and wheat.[21] I display the individual utility functions for two sets of preferences, rice-loving and neutral tastes.

To explore the effect of a rise in the price of rice and an equally sized fall in the price of wheat, I rotate the budget set counterclockwise around bundle B, thereby keeping income fixed. These are the type of price changes that would be seen in regions abundant in rice land as regional prices move towards the integrated equilibrium prices. The solid lines show the pre-price-change and the dashed lines the post-price-change situation. In the case of rice-loving tastes, caloric intake declines, with the new consumption bundle A* on a lower isocalorie line, while caloric intake increases for the case of neutral tastes in the move to B*.

Caloric intake declines for rice-loving consumers through the combination of wealth and

[21] I assume local non-satiation. The budget set edge now coincides with the isocalorie line.

substitution channels. Rice-loving consumers initially spend a large portion of their budget on rice, and so a larger increase in expenditure would be required to afford their original consumption bundle after the relative price change. Because of their strong tastes for rice, these consumers are reluctant to substitute into the cheapening calorie source, wheat, which would allow them to avoid a decline in caloric intake. Therefore, consumers who have developed tastes that favor rice require larger increases in nominal food expenditure to maintain their caloric intake when rice becomes relatively more expensive.

The simple exposition in figure 3 holds nominal food expenditure constant. Trade changes real factor incomes and so the budget set may also shift. My empirical work estimates how large the increase in food expenditure will have to be in order for the average rural Indian household to maintain its caloric intake upon internal trade liberalization. I then compare this number to the required food expenditure increase if tastes were identical across India.

In the example above, the two goods had an equal price per calorie prior to the price change. Even if rice was a cheaper calorie source prior to liberalization and remained so, the net effect of tastes on caloric intake will generally still be negative as rice-loving tastes increase the budget share spent on rice before the price change, reducing caloric intake through the wealth channel.[22] The price per calorie will also vary for reasons unrelated to endowments. For example, to produce one edible calorie, meat requires greater energy inputs than rice. A rise in the price of rice will induce a standard substitution effect, with relative meat consumption rising, lowering caloric intake. These effects will be asymmetric, with meat producing regions gaining calories through this channel and rice producing regions losing. However, my habit formation mechanism will be occurring for all regions within broad food groups. Consumers are reluctant to substitute from local to foreign cereals, and at the same time from local to foreign meats. Therefore, I control for the initial price per calorie explicitly in the empirical work as this substitution between food groups of different caloric intensity

[22] If rice is a Giffen good for some consumers, rice consumption will rise after the price change, and consumption of the other good will fall considerably. However, this will generally still result in caloric intake decreasing unless rice remains a substantially cheaper calorie source after the price change. Jensen and Miller (2008) provide evidence of Giffen behavior in extremely poor households in Hunan, China.

is not my focus of attention.

The effect of price changes on caloric intake shown in figure 3 can be derived for G goods subject to small equilibrium price deviations. The total calories consumed by an individual are equal to the sum of the quantities of each food consumed, $calories = \sum_{g=1}^{G} \frac{food \times bshare_g}{p_g}$. I log-linearize $calories$ around the equilibrium price and apply the envelope theorem:[23]

$$\Delta \ln calories \simeq \Delta \ln food - \sum_{g=1}^{G} [tastes_g + h_g(P, food)] \left[\frac{food}{calories} \Big/ p_g \right] \Delta \ln p_g. \qquad (2)$$

The log change in caloric intake, $\Delta \ln calories$, equals the log change in food expenditure, minus the summation of the interaction between log price changes and both the budget shares and the inverse of the relative price per calorie of each good g. Tastes affect caloric intake since $\Delta \ln calories$ decreases with the correlation between tastes and price changes.

Hypothesis 4: *For a given set of price changes, the greater the correlation between tastes and the price changes, the more caloric intake will decrease, conditional upon total food expenditure and the relative price per calorie.*

2.4 Distributional Issues and Absolute Caloric Losses from Trade

Up to this point, I have shown that the aggregate caloric gains from trade shrink in a model with habit formation but I have not discussed the distribution of these reduced gains. Liberalizing agricultural trade will lead to production gains through specialization, and the total world output of calories will increase. Accordingly, the small region will still generally see aggregate caloric gains from trade. As with the Heckscher-Ohlin model, the real rental income accruing to the owners of rice land will rise, and the real rental income accruing to the owners of wheat land will fall. Habit formation reduces the aggregate production gains from trade by bringing autarky prices closer together, and so these income changes for landowners will be muted. At the same time all factor owners, including the mobile factor, labor, will be negatively affected on consumption side as their more favored foods rise in price.

Most of the poor in the developing world have few productive assets other than their own

[23] In appendix D.2, I show the full linearization that relaxes the envelope theorem assumption that budget shares remain constant. P is the vector of G prices. Food expenditure, $food$, is assumed additively separable from other expenditure. If this is not the case $h_g(P, income)$ should replace $h_g(P, food)$.

labor, and so the likely impact on wages is particularly relevant for exploring the impact of agricultural trade on the poor.[24] Ruffin and Jones (1977) analyze the effect of trade liberalization on landless labor's real income in a specific factors model. The nominal income gain from trade for the mobile factor is a weighted average of the price changes. The authors term this fact the neoclassical ambiguity, since with a sufficiently large share of consumption in the high $\Delta \ln p_g$ goods, the real wage of labor will actually fall.[25] A similar result holds for caloric intake, as can be seen from equation 2. Landless labor can suffer absolute caloric losses from trade, with or without habit formation. However, habit formation produces a positive correlation between tastes and trade-induced price changes, and so larger increases in nominal food expenditure are required to avoid a decline in caloric intake.[26]

Absolute caloric losses for landless laborers are especially likely if this group has particularly strong tastes for the comparative advantage foods, yet does not see large rises in nominal income compared to other groups. This may be expected, since landless laborers are typically near the bottom of the income distribution, and the poor consume less diversified diets and disproportionately large quantities of the cheap local staples that rise in price with trade. Such differences in tastes arise naturally across income groups in the model above when preferences are non-homothetic, with richer households developing more diversified tastes as they chose to indulge in more expensive and varied diets in previous generations.[27]

In appendix C.7, I present a simple parametrization of the model using the same non-homothetic preferences that I use in my empirics. I demonstrate that, under reasonable parameter values, landless labor develops particularly strong tastes for the locally abundant

[24] In India 31 percent of rural households derive their livelihood primarily from agricultural labor and 33 percent from self-employment in agriculture (NSS 1987/88). However, much of the income for self-employed agricultural households with small landholdings comes from their labor rather than their implicit rent.

[25] Ruffin and Jones (1977) assume that the tastes of laborers are not biased towards the comparative advantage good relative to the population as a whole and are independent of comparative advantage. They conclude that there is a presumption that labor will benefit from trade.

[26] Habit formation may also alter the distribution of the production gains from trade. Appendix C.5.1 shows that for the mobile factor, labor, any redistributed gains accruing to labor do not compensate for the consumption loss, and habit formation reduces labor's caloric gains from trade in the Cobb-Douglas case.

[27] These different preferences may explain the fact that richer households are more in favor of trade liberalization than poorer ones. Richer households are better able to take advantage of price changes at the time of liberalization, as they have already developed tastes for the goods which trade makes relatively cheap.

foods. Accordingly, this group loses in caloric terms at the time of trade liberalization, while they would have gained from trade in a world without habit formation. In section 6, I present supporting evidence for India that poor landless households do indeed require relatively larger nominal income rises at the time of trade liberalization in order to avoid caloric losses.

2.5 Welfare Implications Using Quality Improvements

This paper focuses on the nutritional consequences of trade liberalization, but I can also interpret the model in utility terms with a small alteration. The model is isomorphic to one where individuals have fixed preferences and there are local transformation technologies for converting raw food products into palatable meals. The technologies encompass both recipes and preparation techniques. Improvements in the transformation technologies are functions of relative consumption.[28] Returning to the Indian example, a family in Rajasthan may be expert at transforming wheat into delicious roti (a flat bread), but may lack the ability to make a decent jhal-muri (a rice ball popular in West Bengal), because wheat-growing regions learn methods of transforming wheat into high quality meals faster than they learn how to improve the quality of rice dishes they rarely prepare. This reinterpretation allows for an evaluation of the welfare impacts of the model.

The utility gains from trade are muted since consumers continue to buy the local foods that they prepare well rather than the now cheaper imported foods that they are less familiar with. In appendix C.6, I formally show the isomorphism of the transformation technology and preference change models and prove that, for landless labor in the Cobb-Douglas case, the welfare elasticity with respect to trade liberalization declines if food transformation techniques develop proportionally to previous consumption.[29]

Hypothesis 3*: *Food transformation technologies that improve with relative consumption reduce the short-run aggregate welfare elasticities with respect to trade liberalization.*

[28] This has obvious similarities to the induced agricultural innovation of Hayami and Ruttan (1971). A similar bias evolves from a pure information story in which consumers know about the existence and nutritional content of local foods, and are less familiar with foods that trade makes relatively cheap.

[29] Similarly to the habit model, landless labor can lose in utility terms at the time of trade liberalization, but would avoid such a loss if technological improvements were independent of past relative consumption.

The historical example of the Columbian Exchange suggests that both tastes and transformation technologies for local foods reduce the initial gains from trade. Europe had imported both potatoes and tomatoes prior to 1544. However, tomatoes were initially used as table ornaments, with the first recipe for tomato sauce appearing only in 1839. Traditional staples continued to be preferred to potatoes even in times of famine well into the 1800's, at which time potato consumption increased with state intervention (Nunn and Qian, 2009) and tastier preparation methods (a fried potato slice recipe first appeared in 1795). It took both core preference changes and preparation improvements for European consumers to experience large caloric and utility gains from these new crops. In section 5.2, I provide evidence that local transformation technologies alone cannot account for the results I find in India.

3 Empirically Testing the Theory

It is impossible to observe the impacts of trade liberalization in societies both with and without habit formation. Therefore hypothesis 3, which compares the short-run caloric gains in a society with habit formation to one without, cannot be tested directly. However, the mechanisms in my model that reduce the caloric gains from trade on the consumption side can be tested. If a region has stronger tastes for the foods that its agro-climatic endowments are relatively well-suited to producing, and these foods are inexpensive compared to other regions (hypothesis 1 and assumption 1), price rises will systematically occur in more favored foods. Hypothesis 4, in which individual caloric intake declines when tastes correlate with price changes, then implies that trade needs to generate larger increases in total food expenditure for an individual to avoid a caloric loss at the time of liberalization. If these two hypotheses are supported, and if I can show that preferences do respond to past consumption, this is strong evidence that habit formation reduces the caloric gains from trade liberalization on the consumption side.[30]

I carry out the tests detailed above by looking at tastes and prices across 77 regions of India, and analyzing the impact of temporal price changes. I then provide complementary

[30] Habit formation should also reduce the aggregate production gains from trade, but this I will not be able to test without observing autarky prices in a society without habit formation.

evidence from the consumption patterns of inter-state migrants, and from the caloric impacts of the spatial price changes that these migrants' face when they move. Migrants provide a natural experiment that mimics a small economy opening to trade as they take the prices of their destination-state upon migration, yet maintain the preferences of the state in which they were born.[31] Finally, I quantify the caloric impact of the regional taste differences that I document if India were to liberalize its internal agricultural markets by performing the consumption side of the counterfactual exercise suggested by figure 3. I predict the size of the total food expenditure gains required to avoid a short-run caloric loss if prices were to equalize across regions and compare this figure to the predicted size of the total food expenditure gains required if tastes were identical across India.

3.1 Data

Both empirical approaches utilize household data from the Indian National Sample Survey Organization (NSSO). My main sample comes from the 43rd round (1987-88), which is the only comprehensive (thick) round available containing extensive migration data linked to a full consumption module. In order to explore the impact of temporal price changes, I also use household data from various other quinquennial thick rounds. Each round contains observations for 80,000 rural households and 45,000 urban households. The NSSO surveys record quantities purchased and expenditures for every food item consumed from a list of several hundred. From this data I calculate unit values which serve as my price data.[32] The surveys also provide many household characteristics, other expenditures and migration details.

I obtain calorie data for each household by multiplying each food's caloric content, estimated by the NSSO, by the quantity consumed over the previous 30 days. I use this number to calculate the daily caloric intake per household member. I aggregate up to 52 of the most common food products,[33] and define food expenditure as total expenditure on these 52

[31]The use of migrants to evaluate the impact of price changes dates back at least to Staehle (1934).

[32]For home production, consumption is valued at the prevailing local farm-gate price. The unit values correspond well to farm harvest prices and provide large amounts of price information at a fine geographic level. However, unit values are not actual prices since quality varies, an issue I discuss in section 4.1.

[33]They are: Rice, Wheat, Jowar, Bajra, Maize, Barley, Small Millets, Ragi, Gram, Cereal Substitutes, Arhar, Moong, Masur, Urd, Peas, Soyabean, Khesari, Other Pulses, Milk Products, Vanaspati/Margarine, Mustard Oil, Groundnut Oil, Coconut Oil, Other Oil, Meat, Chicken/Eggs, Fish, Potato, Onion, Sweet

foods.[34] Endowment data come from a variety of sources and are described in appendix B.

3.2 Estimating Tastes

Obtaining a measure of local tastes presents my most difficult empirical task. Since preferences change only over generations in my model, tastes will be fixed in the short run and can be identified using cross-sectional data. I use the data from the 1987-88 survey and regress individual budget shares on income, prices and household characteristics, and attribute the remaining regional variation to local tastes.

With a sufficiently flexible functional form, the tastes implicitly defined by $bshare_g = tastes_g + h_g(P, income)$ will be identified separately from the common price and income effects as long as there is price variation within regions which is not driven by taste differences. For example, if markets are not fully integrated within a region,[35] recent local supply shocks such as local rainfall variation or infrastructure improvements would cause prices to vary between villages, and allow the identification of regional tastes.

I estimate these residual tastes using the functional form for $h_g(P, income)$ suggested by the Linear Approximate Almost Ideal Demand System (LA/AIDS) of Deaton and Muellbauer (1980a). In the basic AIDS specification, the budget share spent on food g in region r is a function of a good-specific constant, log prices for every good and log real income. I allow this constant term to vary by region by including a full set of regional dummies, d_{gr}, and the coefficients on these dummies, $tastes_{gr}$, are my regional taste measure:[36]

$$bshare_{gr} = tastes_{gr} d_{gr} + \sum_{g'} \gamma_{gg'} \ln p_{g'r} + \beta_g \ln \frac{income}{P_r}, \tag{3}$$

where p_{gr} is the price per calorie of good g in region r, $income$ is total expenditure and P_r is the regional price index.[37] This specification derives from a "flexible functional form" cost

Potato, Cauliflower, Cabbage, Brinjal, Lady Finger, Tomato, Chillis, Other Vegetables, Coconuts, Other Nuts, Banana, Mango, Oranges, Lemon, Guava, Other Fruits, Sugar, Garlic, Ginger, Turmeric, Black Pepper, Other Spices and Pan/Supari.

[34] This aggregation omits processed foods and beverages that constitute less than 2 percent of caloric intake. For these goods, it is impossible to match the good to an agro-climatic endowment, or to obtain accurate quantity or caloric data. However, results are robust to including these goods and using recorded quantities and NSSO calorie approximations.

[35] Jha et al. (2005) show that there are rice markets within the same region that are not integrated.

[36] Blanciforti and Green (n.d.) first interpreted the AIDS constant as an evolving taste shifter.

[37] The price index P_r is defined by $\log P_r = \alpha_0 + \sum_g tastes_{gr} \ln p_{gr} + 1/2 \sum_g \sum_{g'} \gamma_{gg'} \ln p_{gr} \ln p_{g'r}$.

(or expenditure) function, $c(u,p)$, that can be regarded as a second order approximation to any arbitrary cost function. I allow the first-order price terms, $tastes_{gr}$, to vary by region:

$$\log c(u,p) = \alpha_0 + \sum_g tastes_{gr} \ln p_{gr} + \frac{1}{2} \sum_g \sum_{g'} \gamma^*_{gg'} \ln p_{gr} \ln p_{g'r} + u\beta_0 \prod_g p_{gr}^{\beta_g}.$$

Equation 3 is then obtained via Shephard's Lemma, where $\gamma_{gg'} = \gamma_{g'g} = 1/2(\gamma^*_{gg'} + \gamma^*_{g'g})$. Following Deaton and Muellbauer (1980a), I approximate the price index P_r by a Stone index, $\ln P^*_r = \sum_g \overline{bshare}_{gr} \ln p_{gr}$, making the system linear.[38] The Marshallian own and cross-price elasticities in this demand specification depend on the taste parameter $tastes_{gr}$, with increased tastes for a food reducing the cross-price elasticity of demand for substitutes.[39]

I assume weak separability between the consumption of my 52 food groups and other expenditures. This assumption allows me to estimate demand conditional upon total food expenditure by replacing budget shares with food expenditure shares and income with total expenditure on the 52 foods, $food_i$ (Deaton and Muellbauer, 1980b).[40] I estimate equation 4 separately for each good using OLS over all i households, where the within-region variation in prices, $food_i$ and Z_i allows the identification of β_g, Π, the $\gamma_{gg'}$'s and the residual $tastes_{gr}$:[41]

$$bshare_{gri} = tastes_{gr} d_{gr} + \sum_{g'} \gamma_{gg'} \ln p_{g'v} + \beta_g \ln \frac{food_i}{P^*_r} + \Pi Z_i + \varepsilon_{gri}. \tag{4}$$

I include additional demographic and seasonal controls Z_i and use survey weights.[42] I assume that there is a common price in each village, a reasonable assumption given that there is typically only one food market in an Indian village, and accordingly use median village prices, $p_{g'v}$, as the prices faced by households in village v.[43]

---

[38] $\overline{bshare}_{gr}$ is the average budget share of good g in region r.

[39] $\epsilon_{gg'r} = \frac{\gamma_{gg'} - \beta_g (tastes_{g'r} + \sum_{g''} \gamma_{g'g''} \ln p_{g''r})}{tastes_{gr} + \sum_{g'} \gamma_{gg'} \ln p_{g'r} + \beta_g \ln \frac{income}{P_r}} - \delta_{gg'}$, where $\begin{cases} \delta_{gg'}=1 \text{ if } g=g' \\ \delta_{gg'}=0 \text{ if } g\neq g' \end{cases}$. The partial derivatives with respect to tastes are as follows: $\partial\epsilon_{gg'r}/\partial tastes_{gr} = -\epsilon_{gg'r}/bshare_{gr} < 0$ if $g \neq g'$ and $\epsilon_{gg'r} > 0$). $\partial\epsilon_{ggr}/\partial tastes_{gr} = -(\epsilon_{ggr} + 1 - \beta_g)/bshare_{gr}$ if $g = g'$.

[40] The alternative solution involves making arbitrary assumptions about non-food elasticities since it is extremely difficult to obtain reliable unit values for non-food items which are often highly differentiated products, or services where it hard to record the quantity purchased.

[41] The AIDS should satisfy adding up, homogeneity and symmetry when every individual consumes every item. Since none of the 128,000 sample households purchase all 52 foods, Deaton (1997) interprets OLS equation 4 as a linear approximation to the conditional budget share averaging over zero and non-zero purchases.

[42] Demographic controls include household size, composition, religion, caste and primary activity.

[43] The use of individual prices imparts a bias (measurement errors in individual prices alter budget shares) and there are endogeneity concerns (if individual price paid is correlated with omitted variables). The median

Under the null hypothesis of no regional taste differences, all these estimated $tastes_{gr}$ should be identical. The null is rejected for all foods except sweet potato, with significant $tastes_{gr}$ differences across regions (the mean F-stat for the 52 Wald tests is 31.8). The coefficients on the price terms are also precisely estimated, suggesting that prices are not so poorly measured that the price coefficients are attenuated to zero.[44]

Reassuringly, the empirical results presented in sections 5.1 through 5.3, which are based on the consumption patterns of migrants, do not rely on the validity of these taste estimates. Appendix D addresses the robustness of these taste estimates directly. Although households take village prices as given, there are still endogeneity concerns if there is price variation across villages that originates from idiosyncratic village tastes that raise both local prices and demand.[45] Therefore, I instrument for local village prices with the prices in a nearby village, where supply conditions will be similar. However, this nearby village's taste shock will be uncorrelated with the individual's error term. The instrumentation procedure also provides consistent parameter estimates if village prices are poorly measured, but errors are uncorrelated across villages. I also show that the rank ordering of tastes for a given food across the 77 regions is unbiased if the foods are substitutes for each other and price deviations within regions are only weakly correlated across foods.

Appendix D contains several additional robustness checks: In case total food expenditure

village prices are robust to outliers and not contaminated by quality effects that would typically overstate the price response. In order to avoid losing all village observations when a single price is missing, if none of the village sample report purchasing a good, I use the median price at an incrementally higher level of aggregation until a purchase is reported. The $\gamma_{gg'}$'s will be biased if unrecorded products were available locally, but were unusually expensive and so not purchased by the sample. Therefore, in appendix D, I show results carry through when I add two types of ad valorem transport cost to these imputed prices: A 5 percent cost for each incremental level of aggregation, and a cost calculated from sugar price differences between the two areas.

[44] 47 of the foods have p-values less than 0.00005 from the F-tests on the price terms. If prices were poorly measured and there were no regional taste differences, $tastes_{gr}$ would absorb the true price effects. Therefore, I use other prices (mean, minimum, maximum, 25th and 75th percentile village price) instead of medians. Results are unchanged and shown in appendix D for the endowment and price relationships, and available on request for other regressions. Additionally, if the taste estimates simply reflect unmeasured regional price variation, they would not respond negatively to historic price changes, as I find in section 4.2.

[45] For example, the arrival of an immigrant who introduces a new food or recipe to the village. If, instead, village taste differences originate from agro-climatic endowments varying within regions, the analysis should ideally be carried out at a lower geographical level. When these village agro-climatic endowment deviations are uncorrelated across space, my instrumentation strategy provides unbiased estimates of average regional tastes as long as spatially-correlated recent local supply shocks provide additional price variation.

is endogenous to demand for a particular food, I instrument food expenditure with non-food expenditure, which allows me to bound this bias.[46] I allow for region specific $\gamma_{gg'}$ terms by drawing on additional data and price variation from the two adjacent NSSO thick rounds. Finally, I allow the own-price elasticities to depend on household characteristics by interacting the full set of control variables, Z_i, with the own price term, $\ln p_{gv}$. In all cases, the main results are robust to using these alternative tastes estimates.

4 Empirical Approach 1: Comparing Regions Across India

India contains 77 NSSO regions drawn along agro-climatic boundaries and within the borders of the 31 states.[47] The theory suggests that tastes are related to agro-climatic endowments, making this an appropriate unit. Inter-state tariffs, trade regulations and transportation costs, detailed in appendix A, mean that markets in these regions are not fully integrated, and so I think of these regions as small economies in autarky.

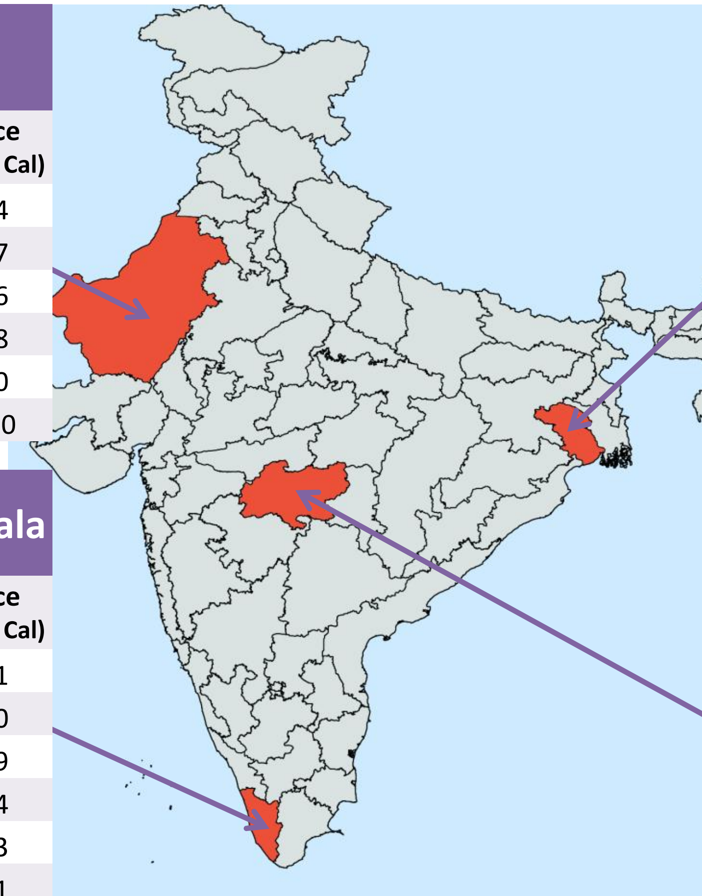
Figure 4 illustrates the variation in food expenditure shares and prices in 1987-88. In arid Western Rajasthan, wheat, bajra (pearl millet) and milk are the most important food sources. Households in Inland Eastern Maharashtra spend the largest portion of their food budgets on jowar (sorghum), while those in the Western Plains of West Bengal, by the Ganges Delta, devote a full 57 percent of their food budget to rice. In Southern Kerala, fish supplements rice as the major food source. In all these cases, prices are relatively cheaper in the regions where the corresponding foods are consumed most. However, this price variation is insufficient to fully explain the enormous variation in food expenditure shares, and these unexplained components form my taste estimates.

I restrict attention to rural households, which comprise around three-quarters of India's population. Rural households spend a larger portion of their income on food, compared to urban households, and their tastes are likely to be more closely related to local endowments because of traditional lifestyles, greater trade barriers with other parts of India, minimal

[46]This strategy also deals with correlated measurement error concerns (as $food_i$ appears in the denominator of the food share) if the measurement errors in food and non-food expenditure are independent.

[47]Only 76 regions have rural samples in 1987-88, since only urban strata were sampled in Nagaland. Four new states were formed in the 1990's, but regions did not change, bar Goa separating from Daman and Diu.

Figure 4: Price and Food Expenditure Share Variation Across Regions of India, 1987-88 (Percent of Total Food Expenditure Spent on Item and Rupee Median Price/1000 Calories)



Western Rajasthan

Item	Food Share	Price (1000 Cal)
Rice	0.4	1.4
Wheat	31.0	0.7
Jowar	2.7	0.6
Bajra	7.9	0.8
Milk	25.4	4.0
Fish	0.0	19.0

Western Plains West Bengal

Item	Food Share	Price (1000 Cal)
Rice	57.0	1.1
Wheat	2.1	0.9
Jowar	0.2	0.7
Bajra	0.0	-
Milk	4.1	4.0
Fish	4.6	14.3

Southern Kerala

Item	Food Share	Price (1000 Cal)
Rice	34.0	1.1
Wheat	1.9	1.0
Jowar	0.0	0.9
Bajra	0.0	1.4
Milk	8.0	4.3
Fish	10.6	8.1

Inland Eastern Maharashtra

Item	Food Share	Price (1000 Cal)
Rice	5.3	1.1
Wheat	9.2	0.8
Jowar	18.9	0.4
Bajra	0.2	0.8
Milk	8.7	4.0
Fish	0.4	14.8

consumption of processed food and less food consumption that occurs outside the house.

4.1 Tastes relate positively to relative endowments, negatively to prices

Habit formation in food consumption predicts that regional endowments above the Indian average should lead to the development of above average tastes for the food intensive in that endowment (hypothesis 1), as the food was relatively inexpensive in previous generations. Therefore, in the absence of the full history of past price data, I regress my taste estimates on a measure of the relative agro-climatic endowment for each good in each region, which would have determined the historic autarky prices.

The specific factors model suggests using the area planted of a crop, $area_{gr}$, as the initial endowment. In reality, land can be planted with a variety of crops, and the current cropping patterns may be affected by factors unrelated to the initial resource endowments that shaped tastes over many generations.[48] Therefore, the observed proportion of a region planted with a crop is a noisy measure of the historic agro-climatic endowment, and so I use a two-stage procedure that estimates the relative suitability of different regions for growing each crop.

In the first stage, I regress observed relative endowments on agro-climatic variables. My observed relative endowment measure is the proportion of a region's farmland planted with a specific crop, averaged over the 1970's, the earliest period available in my dataset and prior to the period in which I estimate tastes.[49] I find very similar results, shown in table D.1, when I use the proportion of a region's total production in either weight or value terms. I regress these observed relative endowments on agro-climatic variables, $AgClim_r$, allowing for crop-specific coefficients by interacting these variables with crop-specific dummies d_g:

$$\frac{area_{gr}}{\sum_{g'}^{G} area_{g'r}} = a_0 + \sum_{g'}^{G} a_{1g'} AgClim_r d_{g'} + u_{gr}. \tag{5}$$

Following Dev and Evenson (2003), I choose a selection of $AgClim_r$ variables that impact crop growth in the Indian sub-continent (the altitude, the mean temperature in January, April, July and October and mean rainfall for June, July and August).[50] I then use the predicted values from this regression as my estimate of the historic relative resource endowment that determines long-run autarky tastes, $endowment_{gr} = \frac{\widehat{area_{gr}}}{\sum_{g'}^{G} area_{g'r}}$.

In the second stage, I regress my taste estimates on predicted agro-climatic endowments:

$$tastes_{gr} = b_0 + b_1 endowment_{gr} + \varepsilon_{gr}. \tag{6}$$

Both stages are run simultaneously using Limited Information Maximum Likelihood.[51] As the magnitudes of tastes and observed endowments vary greatly over the 52 foods, I normalize both variables so that each food is mean zero, standard deviation one, across regions.[52]

[48]For example, recent technological advances, government subsidies or agricultural policies.

[49]Crop data can only be matched to 45 of the 52 goods, with animal products unmatched.

[50]Equation 5 resembles a test of the Rybczynski theorem. A more structural estimation in the context of the Heckscher-Ohlin model is shown in table D.1. I include an additional factor, population density, and assume both an equal number of goods as factors and factor price equalization in productivity equivalent units (necessitating a control for the real agricultural wage).

[51]This is an IV regression with over 300 instruments. I use the LIML Fuller-k estimator (c=1), as it is robust to a large number of weak instruments. The Kleibergen-Paap rk Wald F statistic is 3.2, higher than the Stock and Yogo (2002) 5 percent critical value of around 1.6 for rejecting the null of weak instruments.

[52]I also run a three-stage estimator by including the crop shares instrumented by agro-climatic endow-

Table 1: Tastes and Relative Resource Endowments

	$tastes_{gr}$
$endowment_{gr}$	1.739***
	(0.52)
Observations	3375

Note: Dependent variable, tastes, estimated using unexplained regional variation in food budget shares. Tastes normalized by crop. Endowment comprises the predicted values from regressing normalized crop shares on 7 monthly crop-specific rainfall and temperature variables and altitude. Two-stage estimation using LIML. Robust standard errors. Constant not reported. * significant at 10 percent, ** 5, *** 1.

If tastes develop to favor crops well-suited to local resource endowments, the coefficient on relative endowments, b_1, should be positive. In an alternative, supply driven, induced agricultural innovation story (Hayami and Ruttan, 1971; Acemoglu, 2002), consumption rises due to price declines in the locally abundant good. Therefore, I should find no correlation between tastes and endowments, as my taste measure picks up the demand variation that wan not attributable to price variation in my demand system.[53]

Table 1 presents the results of the two-stage regression. My habit formation hypothesis is supported:[54] Regions which have climates relatively suited to growing a certain food have larger tastes for that food, compared to other regions. In particular, a one standard deviation increase in the relative agro-climatic endowment is associated with tastes that are 1.7 standard deviations higher. Tables D.1 and D.2 show similar results using the alternative taste

ments directly in my demand system, equation 4. By restricting the coefficient on $endowment_{gr}$ to be identical across the 52 regressions, I find a hugely significant positive relationship between (unnormalized) endowments and the unexplained component of demand (coefficient of 0.17, t-statistic of 14298). The coefficients are not directly comparable to the two-stage results as I do not normalize variables in this specification.

[53]There are several other explanations for a positive b_1: First, my taste estimates may reflect regional quality differences, and high quality technologies for a food developed in regions with more suitable endowments via induced technical change. However, the quality differences between basic foodstuffs are generally small and in the next section I find that tastes are inversely correlated with prices. Such a correlation would only occur if the technical change led to absolutely lower prices alongside higher quality, and could only be maintained if there were enormous trade barriers across India for both food and technologies. Secondly, agro-climatic conditions may have initially attracted migrants with well-suited tastes. However, in section 5 I find that migrants in India do not move to regions ideally suited to their tastes. Finally, exogenously tastier crops may have been adapted to grow better in the local agro-climatic conditions. However, in this pure adaptation story there would be no correlation between tastes and historic agro-climatic conditions.

[54]The error term may be spatially correlated. I use the latitude and longitude of each region to apply the correction suggested by Conley (1999). The correction is only available for the GMM estimator, and both here, and in the later OLS regional regressions, the correction actually increases precision.

Table 2: Correlations Between Tastes and Prices

	Pearson's product-moment correlation Prices$_{gr}$	[95 Percent Conf. Interval]	
$tastes_{gr}$	-0.093***	-0.125	-0.061

Note: 3670 observations. Tastes estimated using unexplained regional variation in food budget shares. Prices are regional median unit values. Both variables normalized by good. * significant at 10 percent, ** 5, *** 1. Confidence intervals based on Fisher's transformation.

and price measures discussed in section 3.2.[55] Within India, tastes have become correlated with relative endowments when I estimate tastes using the unexplained regional variation in budget shares controlling for local price effects that are common across India.

I assumed in my theoretical model that habit formation did not bid up the prices of the region's relatively abundant foods so much that these foods actually became relatively expensive compared to other regions. Under this assumption, trade liberalization raises the relative price of preferred foods. I can verify whether this assumption holds for India by calculating the correlation between tastes and prices. There is no causation implied, with historic agro-climatic endowments the root determinant of both prices and tastes.

Table 2 reports this correlation, again normalizing tastes and prices separately by food. Tastes are inversely correlated with prices, so that a region in which a food is more preferred has a lower price for that food compared to other regions. Together with the endowment results above, both hypothesis 1 and assumption 1 are supported, namely that tastes are correlated with relative endowments and inversely correlated with prices. If India were to liberalize internal trade so that prices equalize across India, locally more preferred foods will systematically rise in price and less preferred foods will fall.

4.2 Preference changes correlate inversely with past price changes

The results of the previous section are consistent with my theory of habit formation that links tastes to resource endowments. However, I can test this mechanism more directly by

[55]These robustness checks address the concern that prices are poorly measured and my tastes are just picking up unmeasured price variation. Reassuringly, I obtain similar results when I use the estimated tastes of inter-state migrant households in place of their origin-state tastes. However, as migration data are only recorded at the state rather than regional level, the sample size falls and the standard errors increase substantially, and so the coefficient on agro-climatic endowment is similar in magnitude yet insignificant.

Table 3: Taste Changes and Price Changes Between 1993-94 and 2004-05

	(1) $\Delta_{05,94}$*relative tastes*$_{gr}$	(2)
$\Delta_{94,83} \ln p_{gr}$	-0.0149***	-0.0202***
	(0.0044)	(0.0042)
$\Delta_{05,94} \ln p_{gr}$		-0.00671
		(0.016)
Constant	0.0110	0.0232*
	(0.0079)	(0.012)
Observations	3329	3230
R^2	0.00	0.00

Note: Dependent variable $\Delta_{05,94}relative\ tastes_{gr}$ is the change in the taste coefficient, normalized mean 0, s.d. 1, for each region in each period. $tastes_{gr}$ estimated using unexplained regional variation in food budget shares. $\Delta_{t_2,t_1} \ln p_{gr}$ are log changes in regional median unit values between t_1 and t_2. Robust standard errors. * significant at 10 percent, ** 5, *** 1.

verifying that taste changes over time relate to past changes in relative prices.

My habit formation assumption implies that relative tastes among foods respond to the relative price changes in the previous generation. As I do not have comparable historic price data, I draw on other thick rounds of the NSSO surveys going back to 1983. In order to make full use of the time period spanned by the data, I estimate two new sets of taste parameters using the 2004-05 and 1993-94 cross-sections and normalize tastes over the 52 foods separately for each region and time period. I regress the change in these relative tastes for a food within a region over this decade, $\Delta_{05,94}relative\ tastes_{gr}$, on the log change in prices for that food in that region over the previous decade, $\Delta_{94,83} \ln p_{gr}$:

$$\Delta_{05,94}relative\ tastes_{gr} = b_0 + b_1 \Delta_{94,83} \ln p_{gr} + \varepsilon_{gr}.$$

Table 3 reports the results of this regression. The coefficient on the change in log price is negative, with foods becoming less preferred when in the previous decade their price rose, compared to other foods in the region. In the second column, I include the contemporaneous price change from 1993-94 to 2004-05, $\Delta_{05,94} \ln p_{gr}$. These price changes would be too recent to affect preferences in my habit formation model. Indeed, I find an insignificant coefficient on contemporaneous price changes.[56] This evidence suggests that the positive correlation

[56]This coefficient may be biased upwards, leading to a spurious insignificant coefficient, as demand

between tastes and endowments found in the previous section resulted in part through habit formation, in which preferences responded to relative prices over many generations.[57]

4.3 Caloric intake declines with the correlation of tastes and price changes

I now show that if prices rise in more preferred foods, as the results of section 4.1 imply, there will be negative caloric consequences. In particular, I test my fourth hypothesis: The greater the correlation between tastes and price changes, the more caloric intake decreases, conditional upon total food expenditure and the relative price per calorie. I examine the price and caloric changes that occurred between my main sample period, 1987-88, and the most recent sample, 2004-05, over the 76 regions sampled in both periods.[58]

I require a specification for how tastes and price changes impact caloric intake, and my regression specification comes directly from equation 2, the log linearization of caloric change. I separate the taste effects from the relative price per calorie effects by taking first order Taylor expansions around the average budget share and the average inverse relative price per calorie, $J_{gr} = (food_r/calories_r)/p_{gr}$.[59] This expansion suggests regressing caloric change on the regional sum of tastes interacted with price changes:

$$\Delta \ln calories_r = b_0 + b_1 \sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr} + b_2 \sum_{g=1}^{52} h_g(P_r, food_r) \ln p_{gr} + b_3 \Delta \ln food_r + b_4 \sum_{g=1}^{52} (J_{gr} - \overline{J_r}) \Delta \ln p_{gr} + \varepsilon_r. \quad (7)$$

I measure $\Delta \ln calories_r$ by the log change in the mean total calories consumed per person per day in region r, and $\Delta \ln food_r$ by the log change in the mean monthly total food expenditure per person in region r.

shocks may increase both $\Delta_{05,94} \ln p_{gr}$ and $\Delta_{05,94} relative\ tastes_{gr}$.

[57]An alternative explanation is that consumers purchase complementary stocks of durables which leads to behavior indistinguishable from habit formation. This is not plausible for food choices over decades, since food preparation equipment costs relatively little and can be used with a variety of foods.

[58]The two rounds are sufficiently far apart to observe significant changes in relative prices within regions, comparable to a trade liberalization. I obtain similar results for the regressions using other rounds of data. The price changes over this period were not related to comparative advantage, as there was little internal liberalization over the period, as discussed in appendix A. Regions with larger endowments and tastes for a good actually saw smaller price rises, perhaps due to vote maximizing local politicians attempting to keep their region's preferred local foods affordable for poor voters.

[59]The expansion is $\Delta \ln calories_r \simeq \Delta \ln food_r - \overline{J_r} \sum_{g=1}^{52} [tastes_{gr} + h_g(P_r, food_r)] \Delta \ln p_{gr} + \overline{bshare_r} \sum_{g=1}^{52} (J_{gr} - \overline{J_r}) \Delta \ln p_{gr}$.

Table 4: Caloric Change and the Correlation of Tastes with Temporal Price Changes

	(1)	(2)	(3)	(4)
	$\Delta \ln calories_r$ 1987-88 to 2004-05			
	(Unweighted)	(Weighted)	(Weighted, Excluding Farmers)	
	Full Sample		Poorest Quarter	<2000 Calories
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$	-0.354***	-0.600***	-0.556***	-0.364***
	(0.11)	(0.083)	(0.10)	(0.082)
$\sum_{g=1}^{52} h_g(P_r, food_r) \Delta \ln p_{gr}$	-0.296**	-0.566***	-0.648***	-0.382***
	(0.14)	(0.10)	(0.11)	(0.099)
$\Delta \ln food_r$	0.595***	0.732***	0.603***	0.371***
	(0.047)	(0.049)	(0.066)	(0.062)
$\sum_{g=1}^{52} (J_{gr} - \overline{J_r}) \Delta \ln p_{gr}$	2.828*	1.942*	3.571	1.542*
	(1.60)	(1.16)	(2.54)	(0.78)
Constant	-0.383***	-0.258***	-0.136	-0.0417
	(0.11)	(0.078)	(0.11)	(0.058)
Observations	76	76	75	75
R^2	0.58	0.71	0.61	0.50

Note: Dependent variable is log change in caloric intake per person between 1987-88 and 2004-05. Independent variables come from log linearizing caloric intake. Tastes estimated using unexplained regional variation in food budget shares. J_{gr} is inverse relative price per calorie. Where indicated, regressions weighted by a region's total survey weight in 1987-88. Robust standard errors. * significant at 10 percent, ** 5, *** 1. Subsamples are the poorest quartile by region, and those consuming fewer than 2000 calories per person per day, excluding any household who reports self-employment in agriculture as primary activity.

If the approximation is reasonable, there are sign predictions on the population-averaged slope coefficients. The change in regional caloric intake should decline with a measure of the correlation between tastes and price changes, $\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$, ($b_1 < 0$). There should be a similar effect for the explained component of the budget share, ($b_2 < 0$). Additionally, the change in regional caloric intake should increase with larger increases in food expenditure, ($b_3 > 0$), and when expensive calorie sources rise in price the most, ($b_4 > 0$).

Table 4 reports the results of regression 7.[60] Column 1 contains unweighted estimates

[60] The tastes are themselves estimated, which generates additional error that should be added to the standard errors reported in table 4. Therefore, I bootstrap the household sample 1000 times, estimating tastes and running regression 7 each time. However, the additional error generated through the taste estimation is minuscule, with the standard error on b_1 increasing by less than 0.005. There is also no evidence that $\hat{b}_1$ is attenuated, with the mean value of $\hat{b}_1$ from the bootstrap almost exactly equal to the observed coefficient.

of the average regional coefficients. In column 2, the coefficients are representative of the Indian population as a whole. Over the period, the average Indian caloric intake from the 52 foods declined from 2,200 to 2,000 calories per day. This decline was larger in regions where price changes over the period correlated more strongly with tastes (b_1 significantly less than zero).[61] Table 4 also provides strong support for the other three sign predictions.

Appendix D discusses and shows the results of three important robustness checks: I show that households reduce non-food expenditure in response to rising prices for more favored foods, however, the reallocation is not sufficient to maintain caloric intake. I relax the envelope theorem assumption that budget shares are fixed in the short run. Finally, I instrument for $\Delta \ln food_r$ with $\Delta \ln non\ food_r$ to bound any bias from food expenditure being correlated with shocks to caloric demand. In all cases, my results carry through.

Poor and undernourished households may put a higher weight on calories, relative to taste considerations, compared to wealthier households. Additionally, the theoretical model unambiguously predicts an increase in caloric intake at the time of trade liberalization for net producers of rice when the price of rice rises with trade liberalization. Therefore, by restricting my sample further to only the households who are either not working in the agricultural sector at all or are landless agricultural laborers,[62] I can determine the caloric elasticities for the key groups whose health is most likely to suffer permanent damage from a decline in caloric intake at the time of liberalization.

Columns 3 and 4 of table 4 shows that the data support the conjecture that the caloric intake of the poor and undernourished responds less elastically to relative food price changes. The b_1 coefficients for the two subgroups of non-farmers who either fall in the poorest quartile of each region or consume fewer than 2000 calories per person per day (already substantially below the recommended intake of 2400 calories) are 0.56 and 0.36 respectively, compared to 0.60 for the full sample.[63] In section 6, I calculate the potential impact of internal trade

[61] This result is not simply coming from larger aggregate price rises in some regions. Results, shown in tables D.5 and D.6, are similar when the price changes are demeaned by region or when a correlation is used.

[62] I exclude all households who report self-employment in agriculture as their primary activity.

[63] I estimate new tastes coefficients for each subgroup. Membership of these groups is clearly endogenous. However, since I only have a repeated cross-section, such problems are unavoidable, and the results still

liberalization in India on the consumption side, which provides a sense of the magnitude of these estimates. The predicted caloric impacts for these subgroups and the average rural Indian turn out to be very similar, as poor households have lower elasticities but much stronger tastes for the local foods that are expected to rise in price.

Putting the three central empirical results together, I have shown that tastes positively relate to relative endowments through habit formation, and that the correlation between tastes and price changes reduces caloric intake for a given level of food expenditure. Consequently, the caloric gains from trade liberalization will be muted and there may well be absolute caloric losses as landless laborers continue to purchase their favored local foods that have systematically risen in price.

5 Empirical Approach 2: Migrants as Small Open Economies

As further evidence for my hypotheses, I exploit spatial rather than temporal price differences. Migrants move with their labor endowment and their origin-state tastes, and face the prices of their destination state. This change mimics the experience of a household suddenly exposed to world prices upon trade liberalization. I verify that moving affects migrants similarly to how my model predicts trade liberalization affects consumers in a small economy.

I use the complete 1987-88 cross section and I define inter-state migrants as households in which either the household head or their spouse emigrated from another state in India.[64] To identify the causal caloric impacts of the price changes faced by a migrant, I must assume that migrants do not differ from non-migrants in unobservable ways, after controlling for total food expenditure and other variables in the dataset.[65]

I avoid the most severe selection problems, those that arise when household heads are choosing to migrate for better employment opportunities, by estimating all the regressions

provide useful insights into the heterogeneous responses of important subpopulations.

[64]This is the only round that contains both detailed migration data and a full consumption module. Many migrants move between rural and urban sectors, and so I now use both the rural and urban sample. I carry out a state rather than region level analysis as the NSS data includes only the migrant's state of origin. If both head and spouse emigrated, I use the household head's migration information.

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

using two additional specifications. In the *"wife move"* sample, I focus only on households in which the wife of the household head moved specifically for the purpose of marriage (either within or between states),[66] and I compare households in which the wife moved inter-state versus intra-state. Since Indian women typically live in their husband's village upon marriage, this covers about two-thirds of the wives in the dataset. The *"wife move 2"* sub-sample restricts the set of observations even further, only selecting households in which the husband still lives in the village of his birth and so the wife is likely to be moving into the extended household of her husband's family.[67] Table D.8 in the appendix contains descriptive statistics for these samples. Both of these strategies assume that the wife carries some of her preferences into household food purchasing decisions, and that inter-state and intra-state migrant wife households are identical after conditioning on the many observables.[68]

5.1 Migrants bring their tastes with them

The first test verifies that migrants maintain some of their origin-state's tastes when they move. This exercise provides evidence that, if India were to liberalize its internal agricultural trade, local tastes would remain concentrated in local foods even after relative prices change.

In order to test this hypothesis, I first calculate the mean food budget share spent on each of the 52 foods in each of the 31 states. I then compute the correlation between the average bundle of every state s and the bundle of every household. This produces 31 data points for each household i, who originally lived in origin state o and now resides in destination state d; $\rho_{iods} = corr_g(bshare_{ig}, bshare_{\bar{s}g})$.[69] I regress all of these ρ_{iods} correlations on five dummy variables: the household lives in that state $\mathbf{I}_{d=s}$, the household lives in that state but is an inter-state migrant $\mathbf{I}_{d=s,o\neq s}$, the household does not live in that state $\mathbf{I}_{d\neq s}$, the household does not live in that state but originally migrated from there $\mathbf{I}_{d\neq s,o=s}$ and the household

[66] I exclude the women who moved state jointly with their husbands at the time of marriage.

[67] Inter-state migrant households comprise 8.0 percent of the full weighted sample, 7.3 percent of the *wife move* sample and 5.4 percent of *wife move 2* sample.

[68] Although Rosenzweig and Stark (1989) suggest that long distance marriages are a risk mitigating mechanism, and so households who engage in inter-state marriages may have less variable expenditures.

[69] Since I create 31 data points from each household observation, I cluster standard errors at the household.

Table 5: Comparing Bundles of Migrants and Non-Migrants

	(1)	(2)	(3)
		$\rho_{iods} = corr_g(bshare_{ig}, bshare_{\bar{s}g})$	
	Full Sample	Wife Move Sample	Wife Move 2 Sample
$\mathbf{I}_{destination=s}$	0.811***	0.805***	0.804***
	(0.0012)	(0.0015)	(0.0016)
$\mathbf{I}_{destination=s,origin\neq s}$	-0.0210***	-0.0288***	-0.0251***
	(0.0037)	(0.0060)	(0.0072)
$\mathbf{I}_{destination\neq s}$	0.483***	0.450***	0.447***
	(0.0011)	(0.0015)	(0.0016)
$\mathbf{I}_{destination\neq s,origin=s}$	0.112***	0.139***	0.135***
	(0.0042)	(0.0063)	(0.0079)
$\mathbf{I}_{destination\neq s,nearby=s}$	0.181***	0.199***	0.199***
	(0.0014)	(0.0020)	(0.0021)
Observations	3,920,725	1,653,633	1,474,205
R^2	0.77	0.75	0.74

Note: Dependent variable is correlation between household food budget shares and mean shares for state s (31 observations per household). Independent variables are indicators for household's origin o and current (destination) d state. In the full sample there are 126,475 households of which 11,336 are migrants. Robust standard errors. All regressions survey weighted and clustered further by household. * significant at 10 percent, ** 5, *** 1.

does not live in that state but lives in a nearby state $\mathbf{I}_{d\neq s,nearby=s}$:[70]

$$\rho_{iods} = b_1\mathbf{I}_{d=s} + b_2\mathbf{I}_{d=s,o\neq s} + b_3\mathbf{I}_{d\neq s} + b_4\mathbf{I}_{d\neq s,o=s} + b_5\mathbf{I}_{d\neq s,nearby=s} + \varepsilon_{is}.$$

If migrants bring their origin-state tastes with them, then being a migrant reduces the correlation between their consumption bundle and the average consumption bundle of their destination state ($b_2 < 0$), and increases the correlation with the average consumption bundle of their origin state ($b_4 > 0$). As shown in table 5, the data support these sign predictions for all my samples.[71] Column 1 shows that the bundle of an average household has a correlation of 0.811 with its own state's average bundle, and this falls to 0.790 for migrant households.

[70] As migrants usually come from nearby states and nearby states have similar diets, omitting $\mathbf{I}_{d\neq s,nearby=s}$ would result in a spurious positive coefficient on $\mathbf{I}_{d\neq s,o=s}$. I group India into 5 zones, and define a nearby state as one in the same zone. I find similar results when controlling for the distance between states.

[71] Reassuringly, I also find that non-migrants' bundles are more similar to their current state than to other states ($b_1 > b_3$), and nearby states have more similar bundles than distant ones ($b_5 > 0$). The positive coefficient on $\mathbf{I}_{d=s}$, is partly mechanical, as individual budget shares are a component of state mean budget shares. However, this bias towards 1 is small as the average state sample contains 4,000 households.

The correlation is only 0.483 with state bundles from other areas of India, but rises by 0.112 if they originally migrated from that state. Table D.9 shows very similar results when I include a large number of additional controls. In conclusion, migrant households maintain some of the preferences of their origin state after migration.

5.2 Migrants consume fewer calories for a given level of food expenditure

I now investigate whether migrants consume fewer calories than those that stayed behind, controlling for total food expenditure. Since migrants' tastes are no longer inversely correlated with local prices, the foods they find particularly tasty now cost relatively more.[72] The expenditure controls absorb any income gain from moving (similar to the production gains from trade). Therefore, this test provides evidence that the local tastes identified in the previous section generate negative caloric impacts when prices change upon migration.

I regress caloric intake per person per day, $calories_i$, on a migrant-household dummy, $migrant_i$, flexible controls for total food expenditure, $food_i$, an extensive set of demographic and seasonal controls Z_i and a full set of origin-state dummies, d_o:[73]

$$calories_i = b_1 migrant_i + a_1 \ln food_i + a_2 \ln food_i^2 + \sum_o \gamma_o d_o + \Pi Z_i + \varepsilon_i. \quad (8)$$

As I include origin-state dummies, b_1 picks up the difference in caloric intake between those that stayed in the origin state and those that left. If the assumptions outlined at the start of section 5 hold, this comparison is equivalent to comparing the caloric intake of households in my model before and after trade liberalization, holding food expenditure constant.

Table 6 shows the results of this regression for the three samples. Inter-state migrants consume fewer calories for a given level of food expenditure ($b_1 < 0$),[74] even when just comparing

[72]The preferred origin-state foods are relatively expensive in the destination state because prices and tastes are inversely correlated (table 2). The same relationship holds across states, where the correlation between normalized state-level tastes and median state prices is -0.08 (95 percent interval of -0.13 to -0.02).

[73]I use log food expenditure as food expenditure is distributed log normally (unlike caloric intake). Additional controls are age variables, the education level of the household head, scheduled caste or tribe, religion, main household activity and household composition variables. Migrants may be moving to locations with higher prices for all goods. Accordingly, I also include dummies for urban-urban, rural-urban and urban-rural migration. As a further check, table D.12 shows the regressions including a full set of destination-sector dummies. Migrants still consume fewer calories when compared to non-migrants living in their destination state and sector (rural or urban) whose tastes are adapted to the local prices (although the magnitude of the caloric loss decreases by about one third).

[74]As with the regional regressions, I replace total food expenditure with total expenditure to check

Table 6: Caloric Intake of Migrants Compared to Non-Migrants

	(1)	(2)	(3)
	Daily Calories Per Person $calories_i$		
	Full Sample	Wife Move Sample	Wife Move 2 Sample
$migrant_i$	-107.2***	-115.0***	-43.87**
	(18.2)	(17.0)	(18.2)
$\ln food_i$	-2777***	-3787***	-3959***
	(963)	(588)	(630)
$\ln food_i^2$	478.1***	585.8***	609.6***
	(115)	(67.1)	(72.1)
Observations	124,578	52,836	47,501
R^2	0.50	0.66	0.67

Note: Daily calories per person regressed on inter-state migrant dummy, log food expenditure terms, origin-state dummies and controls Z_i, with the full results shown in tables D.10 and D.11. Robust standard errors. All regressions survey weighted. * significant at 10 percent, ** 5, *** 1.

inter-state to intra-state wife migrations in columns 2 and 3.[75] Being an inter-state migrant corresponds to the consumption of over 100 fewer calories per person per day controlling for food expenditure, or about 5 percent of total caloric intake. This caloric loss is commensurate with a 7 percent decline in food expenditure for the average Indian household.[76]

The caloric loss is smaller in the third column of table 6, where wives move to their husband's village of birth. These households are likely to be more traditional and contain other members of the husband's family. An out-of-state wife might be expected to have a smaller caloric impact due to a lesser say in household food purchasing decisions. The fact that a significant negative caloric effect remains implies that the isomorphic food technologies story cannot account for these results alone as the extended family will be familiar with the recipes and preparation techniques that take best advantage of the local foods.

for reallocation between food and non-food expenditure. The conditional caloric loss is reduced but only slightly, as shown in tables D.10 and D.11. Results are unchanged, and available on request, when I instrument food expenditure with non-food expenditure to control for potentially correlated measurement error (calories and food expenditure are calculated using the same raw data).

[75] I would also find $b_1 < 0$ if wives from more distant villages are more prized and fed higher quality foods. Alternatively, wives may be fed (or choose) cheaper calorie sources than other household members, as well as wives from further away being fed less food (or control of a smaller share of the household budget).

[76] Tables D.10 and D.11 show that migrants do not actually consume fewer calories, as their food expenditure is higher than those who stayed behind. Part of this derives from the production gains from trade that made the large migration costs worthwhile (migrants have significantly higher total food expenditure yet consume approximately the same number of calories).

Combining this evidence with the previous section suggests that migrant households consume substantially fewer calories for a given level of food expenditure as they continue to consume the favored products from their origin state that are now relatively expensive in their destination state. I will perform a stronger test of hypothesis 4 shortly, and demonstrate that the size of the caloric reduction for a given level of food expenditure depends on the particular tastes of the migrant's origin state and the price changes faced upon that specific migration.

5.3 Caloric loss from migration shrinks over time

Migrant households' tastes will gradually adapt to favor the relatively cheap goods in their destination state through the process of habit formation, bringing caloric gains (hypothesis 2). I test this hypothesis by supplementing the previous regression specification with the length of time since a wife moved, interacted with whether the wife is an inter-state migrant:[77]

$$calories_i = b_1 migrant_i + b_2 yrsaway_i + b_3(migrant_i \times yrsaway_i) \\ + a_1 \ln food_i + a_2 \ln food_i^2 + \sum_o \gamma_o d_o + \Pi Z_i + \varepsilon_i.$$

Table 7 shows the results of this regression, which once more support the habit formation hypothesis. For each additional year that an inter-state migrant wife lives in the destination state, daily caloric intake per person for a given level of food expenditure rises by 2.7 calories more than the rise experienced by an intra-state migrant wife ($b_3 > 0$). The coefficient is of a similar magnitude for the *wife move 2* specification, but is no longer significant. A crude estimate of the number of years required for the tastes of an inter-state wife's household to fully adapt is b_1/b_3, or between 80 and 96 years.[78] Over many years and several generations, habit formation alters tastes to favor locally cheap foods, and there are corresponding caloric gains.

I also include additional interaction terms, with the full results shown in table D.11. I find that the presence of a mother-in-law[79] in the household increases caloric intake for a given

[77]By focusing only on households in which the wives moved for marriage, I can control for caloric changes that depend on the length of time since the wife moved for marriage, $yrsaway_i$. This is necessary as the time since the wife moved will be correlated with demographic unobservables.

[78]These are out of sample estimates and should be treated with caution, as the average time since migration is only 20 years. This empirical strategy has no way of disentangling migrant-specific cohort effects effects from time-since-migration effects, as only the 1987-88 cross-section contains suitable migration data (national cohort effects will be absorbed in the $yrsaway_i$ term).

[79]I code $motherlaw_i = 1$ if a household head lives with their mother or mother-in-law. Given the Indian

Table 7: Caloric Intake of Intra-State and Inter-State Wife Households Over Time

	(1)	(2)
	Daily Calories Per Person $calories_i$	
	Wife Move Sample	Wife Move 2 Sample
$migrant_i$	-257.3***	-140.8***
	(33.05)	(37.82)
$yrsaway_i$	5.529***	4.478***
	(0.644)	(0.672)
$migrant_i \times yrsaway_i$	2.668**	1.768
	(1.143)	(1.208)
Observations	52,800	47,465
R^2	0.66	0.67

Note: Daily calories per person regressed on an inter-state wife dummy, $migrant_i$, the years since a wife moved village and years interacted with being an inter-state wife. Log food expenditure terms, controls Z_i and origin-state dummies omitted, with full results shown in table D.11. Robust standard errors. All regressions survey weighted. * significant at 10 percent, ** 5, *** 1.

level of food expenditure, potentially through her experience at preparing and purchasing affordable local foods. The effect is even stronger in inter-state wife households, either because the wife is less familiar with local foods and recipes, or because the wife has less influence over purchasing decisions when a mother-in-law is present.[80]

5.4 Caloric loss larger when tastes are correlated with price changes

In the final test, I use the information on the actual migration route traveled and the particular tastes of the migrant's origin state. I show that the caloric loss is larger when the relative price rises faced by a migrant moving from origin state o to destination state d are more concentrated in the preferred origin-state foods of that migrant, controlling for food expenditure. The specification mirrors that used in section 4.3 in order to test hypothesis 4, except here I use spatial price changes rather than temporal price changes within regions.

I rerun my migrant calorie regression, equation 8, except now rather than a single migrant household dummy, I allow the migrant effect to depend on the migrant's origin and

norm of patrilocality (Srinivas, 1980), the majority of these women will be the husband's mother.

[80]The second interaction term is the number of children. I find that children reduce calories per person since they require fewer calories than adults. However, for inter-state wife households, there is a positive net impact of children. If parents in my model have some foresight, they will partially adjust meals to include inexpensive local foods so that their children develop tastes for these foods and benefit as adults. Then, the degree of adjustment will increase in the number of children, as I find.

Table 8: Caloric Change and the Correlation of Tastes with Spatial Price Changes

	(1)	(2)	(3)
	Caloric Change with Migration (Controlling for Food Expenditure)		
	Full Sample	Wife Move Sample	Wife Move 2 Sample
$\sum_{g=1}^{52} tastes_{go}\Delta \ln p_{god}$	-0.873***	-0.705***	-0.560***
$\times migrant_i$	(0.0598)	(0.0672)	(0.0841)
$\sum_{g=1}^{52} h_{go}(P_o, food_o)\Delta \ln p_{god}$	-0.864***	-0.759***	-0.733***
$\times migrant_i$	(0.0713)	(0.0946)	(0.127)
$\sum_{g=1}^{52} (J_{go} - \overline{J_o})\Delta \ln p_{god}$	-2.924*	-6.597***	-9.597***
$\times migrant_i$	(1.765)	(2.075)	(2.598)
$migrant_i$	-0.0209***	-0.00127	0.0200**
	(0.00703)	(0.00697)	(0.00820)
Observations	124,578	52,812	47,482
Origin-Destination Pairs	484	341	256
R^2	0.722	0.695	0.699

Note: Dependent variable is the log of caloric intake per person per day. Independent variables shown come from the log linearization of caloric intake. Tastes estimated using unexplained state variation in food budget shares. J_{gr} is the inverse relative price per calorie. Log food expenditure terms, controls Z_i and origin-state dummies omitted, with full results shown in table D.13. Robust standard errors. All regressions survey weighted. * significant at 10 percent, ** 5, *** 1.

destination state. I interact the migrant dummy with a measure of the correlation between the migrant's origin-state tastes and the relative price changes between the migrant's origin and destination states ($\sum_{g=1}^{52} tastes_{go}\Delta \ln p_{god}$).[81] The model predicts that migrants whose origin-state tastes are especially well suited to their destination state's prices should experience smaller caloric declines, controlling for food expenditure. Following the regional specification 7, I use log calories and also include two additional terms that control for the relative price per calorie and the explained component of demand:

$$\ln calories_i = migrant_i[b_0 + b_1 \sum_{g=1}^{52} tastes_{go}\Delta \ln p_{god} + b_2 \sum_{g=1}^{52} h_{go}(P_o, food_o)\Delta \ln p_{god} + b_4 \sum_{g=1}^{52}(J_{go} - \overline{J_o})\Delta \ln p_{god}] + a_1 \ln food_i + a_2 \ln food_i^2 + \sum_o \gamma_o d_o + \Pi Z_i + \varepsilon_i.$$

[81]Since I do not have historical price data, I use the price differences at the time of the survey, and must assume there were similar relative price differences between states at whichever time the migrant moved.

The sample size is much larger than the regional regression, with 484 observed origin-destination migration routes. As before, I use price data and estimate taste parameters for 1987-88, but now by state rather than by region.

Table 8 shows the results of this regression. The greater the correlation between origin tastes and the price changes faced on individual migration routes, the larger the decline in the number of calories obtained from a given level of food expenditure ($b_1 < 0$). These coefficients are slightly larger than those estimated from temporal price variation across regions. The fact that the results are of the same sign and order of magnitude provides strong evidence that caloric intake is negatively impacted when tastes correlate with price changes.

6 The Caloric Impact of Internal Trade Liberalization in India

In this section, I use my previous empirical results to predict the nutritional impacts of the taste heterogeneity that I document within India if internal agricultural trade were liberalized, such that prices equalize across regions.[82] Appendix A details the many current restraints to agricultural trade within India.

I calculate the predicted caloric loss on the consumption side coming from regional tastes correlating with these equalizing price changes, and the rise in total food expenditure that would be required to avoid absolute caloric losses. These estimates can be contrasted with the predicted caloric loss from the same price changes if regional tastes were equal to the average Indian tastes, and were independent of regional endowments.[83] Comparing these two numbers provides an estimate of the bias on the consumption side from ignoring taste differences when predicting the nutritional impact of agricultural trade liberalization.

To proceed, I use the more conservative estimate of the elasticity of caloric change with the correlation between tastes and price changes, $\widehat{b_1}$, from the weighted regional regres-

[82] If high transport costs are a major contributor to the lack of agricultural market integration, then these estimates are relevant to a reform process that includes substantial infrastructure investments.

[83] This counterfactual only addresses the effects of habit formation on the consumption side as autarky prices, and hence price changes upon liberalization, would differ in a society without habit formation. Unfortunately, this cannot be verified because of the impossibility of observing autarky prices in the absence of habit formation. Habit formation may also change the distribution of these reduced production gains. However, investigating these production side effects requires data that span a period of agricultural trade liberalization.

Table 9: The Negative Caloric Impact on Rural Households Coming From Tastes Being Correlated with the Price Changes at the Time of Indian Internal Trade Liberalization

All-India Predicted Means	(1) Actual Tastes	(2) Identical Tastes	(3) Difference
$\Delta \ln \widehat{calories_r}$	-0.0265*	-0.00364	-0.0229***
	(0.016)	(0.017)	(0.0037)
$\Delta \ln food$ expend. to avoid caloric loss	0.0362	0.0050	0.0312
$\Delta \ln expenditure$ to avoid caloric loss	0.0333	0.0046	0.0287
$\Delta \ln \widehat{calories_r}$	-0.0445***	-0.0163	-0.0282***
(Poorest Quartile, Excl. Farmers)	(0.015)	(0.016)	(0.0034)
$\Delta \ln food$ expend. to avoid caloric loss	-0.0739	-0.0270	-0.0468
$\Delta \ln \widehat{calories_r}$	-0.0209	0.0007	-0.0216***
(<2000 Calories, Excl. Farmers)	(0.018)	(0.019)	(0.0033)
$\Delta \ln food$ expend. to avoid caloric loss	-0.0564	0.00189	-0.0583

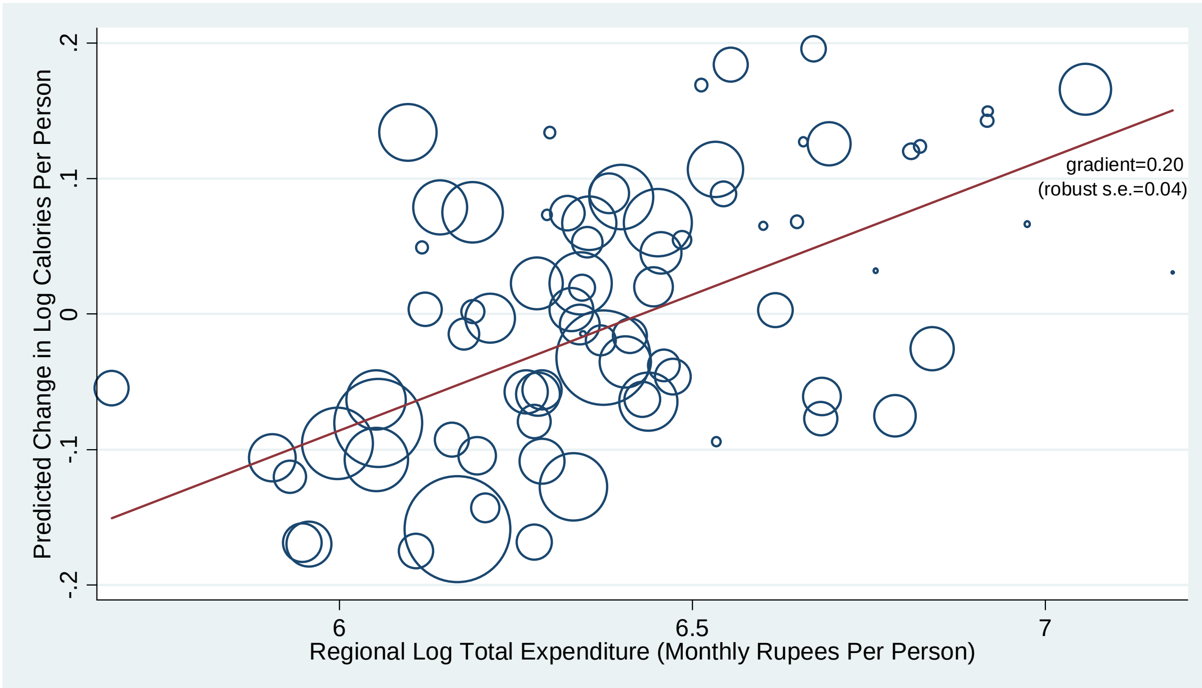
Note: 77 observations weighted by a region's total survey weight. $\Delta \ln \widehat{calorie_r}$ is the predicted log change in calories on the consumption side coming from regional tastes correlating with equalizing price changes in India using data from 2004-2005. Robust standard errors for means. * significant at 10 percent, ** 5, *** 1.

sion. Under the assumption that all regions have the same elasticity, the predicted caloric change attributable to tastes correlating with resource endowments is $\Delta \ln \widehat{calories}_r^{HF} = \widehat{b}_1 \sum_{g=1}^{G} tastes_{gr} \Delta \ln p_{gr}^{lib}$. I use regional taste estimates for rural households from the most recent thick NSSO survey (2004-05). For the predicted equalizing price changes, $\Delta \ln p_{gr}^{lib}$, I use the log difference between the Indian median price and the median price in region r, again from 2004-05.[84] For the counterfactual society, in which tastes are independent of endowments, I assume that tastes are identical across regions and equal to the average Indian tastes for each good, $\overline{tastes_g}$. In this case, the predicted reduction coming from tastes correlating with price changes is $\Delta \ln \widehat{calories}_r^{NHF} = \widehat{b}_1 \sum_{g=1}^{G} \overline{tastes_g} \Delta \ln p_{gr}^{lib}$.

The results shown in table 9 suggest that, holding total food expenditure constant, there will be an average caloric loss of 2.7 percent coming from the correlation between tastes and price changes (about 54 calories per person per day). With identical tastes across India, the caloric loss from this channel would be essentially zero. The estimated coefficient on total

[84]These are the median prices paid by the consumers who actually purchase the good.

Figure 5: The Predicted Caloric Losses from Liberalization and Regional Expenditure (Total Food Expenditure Held Constant, Markers Proportional to Population)



food expenditure in regression 7 was 0.73. Therefore, trade will have to generate income gains that increase average total food expenditure by 3.6 percent in order to compensate for the caloric loss that comes from the correlation of tastes and price changes.[85]

Table 9 also shows the predicted effects of trade-induced equalizing price changes on the hungriest and poorest subgroups discussed in section 4.3. The negative caloric impacts are of a similar magnitude to those found for the full sample, despite lower caloric elasticities for these subgroups. Poor and malnourished households have less diversified diets and accordingly developing stronger tastes for the relatively cheap local foods that rise in price at the time of liberalization. These stronger tastes for local foods negate the subgroups' smaller caloric elasticities for a given set of tastes and price changes. While the caloric impacts are similar, the compensating food expenditure increases are larger than for the full sample, as these subpopulations' caloric intakes were less responsive to food expenditure changes.

In geographic terms, the negative caloric impacts that come from tastes correlating with

[85]Using the coefficients from table D.7, 3.3 percent total expenditure gains are required.

price changes will not be spread uniformly across India. Figure 5 plots the predicted caloric loss from internal trade liberalization, holding food expenditure constant, against the mean per-capita expenditure of the region in 2004-05. There is a highly significant positive slope, with poorer regions more likely to suffer caloric losses on the consumption side, with predicted caloric losses of 20 percent in some of the poorest regions.[86] A similar relationship is shown in appendix figure D.1 for the non-farmers who consume fewer than 2000 calories per day . The poorer regions spend larger portions of their incomes on local staple foods, and so will be harder hit on the consumption side when comparative advantage foods rise in price.[87] Therefore, caloric inequality across India will increase unless agricultural trade liberalization brings the largest income gains to the hungriest rural households in the poorest regions.

The magnitudes of these required income gains on the consumption side are large enough to raise the concern that the static production gains may be insufficient to avoid caloric losses for some key groups at the time of internal trade liberalization. For example, general equilibrium trade models have been used to estimate the production gains from agricultural trade liberalization that come through greater specialization. Anderson and Valenzuela (2007) predict that full liberalization of world agricultural trade would result in the value added by farmers increasing by only 0.3 percent for lower-income developing countries.[88] In table 9, I find that for the poor and hungry landless laborers, food expenditure gains of 5.6 to 7.4 percent will be required just to maintain caloric intake, with even larger numbers in the poorest regions of the country. These households are the most at risk from malnutrition, and so caloric declines are the most damaging for these groups, and at the same time their incomes would not be expected to rise as much as those of landed farmers growing the local compara-

[86]The figure implicitly assumes that the elasticity of caloric change with respect to the correlation between price changes and tastes is not smaller in poorer regions. Table D.14 runs regression 7 separately for richer and poorer regions. The elasticity is actually larger in poorer regions, suggesting an even more skewed regional distribution than that shown in figure 5. I find a similarly significant relationship when I deflate expenditure by a Stone price index with national budget shares. However, it is not clear whether any price index makes sense when tastes differ across regions.

[87]The regional mean per-capita expenditure has a strongly negative correlation of -0.52 with the budget share spent on foods that are expected to rise in price with liberalization (95 percent confidence interval between -0.67 and -0.34). This result is partly driven poorer regions having lower agricultural prices on average.

[88]India in particular would suffer a 2.3 percent decline coming from reduced domestic protection, although Indian farmers would see an increase in their value added of 3.2 percent if India had no existing tariffs.

tive advantage crops. Combining these require food expenditure gains with the low estimates of the static income gains from trade reported above suggests that absolute caloric losses for rural landless laborers are quite possible if India liberalized internal agricultural trade.

7 Conclusions and Policy Implications

International trade theory generally assumes that tastes are identical across regions and independent of endowments. In this paper, I show that habit formation in food consumption leads to regional food tastes that favor crops relatively well-suited to local agro-climatic endowments. This connection erodes the short-run caloric gains from trade liberalization on both the production side (by bringing autarky prices together) and the consumption side (by limiting the substitution out of foods that rise in price at the time of liberalization).

I verify the empirical relevance of the consumption side of the model by exploring India's non-integrated domestic agricultural markets. Regional tastes, measured by the unexplained regional variation in household demand for agricultural products, correlate positively with agro-climatic endowments and negatively with local prices. These tastes respond over time to past relative prices as habit formation would suggest. Finally, caloric intake declines more in regions where price rises are more concentrated in locally favored foods, controlling for changes in food expenditure. I confirm these results by looking at the consumption patterns of inter-state migrants within India, who obtain fewer calories for a given level of food expenditure as their favored foods cost more outside their origin state. This effect dissipates over time, and the caloric intake from a given level of food expenditure declines more where price rises are more concentrated in migrants' preferred origin-state foods.

My findings imply that if India were to liberalize its internal agricultural trade, the prices of preferred foods will rise in each region. Consumers are reluctant to substitute away from these foods, and trade must generate larger income gains in order to avoid caloric losses, as compared to a society without habit formation. Poor and hungry landless households, who consume larger shares of local staple foods, will be especially hard hit through this mechanism.

These results have important policy implications. If agricultural trade liberalization reduces the caloric intake of the poor, from levels already bordering on malnutrition, there will be serious consequences as it is impossible to compensate later in life for nutritional shortfalls while young. The most harmful nutritional impacts can be avoided by accompanying agricultural liberalization with temporary food subsidies for favored local staples, specifically targeted at households on the edge of malnourishment. Many developing countries already have food subsidy systems in place, such as the Public Distribution System in India, making this measure easy to implement. As was the case with the introduction of the potato in Europe, governments can also take direct measures to encourage the adoption of foods made relatively cheap by trade, ensuring that the full caloric gains from trade arrive more quickly. More recently, governments in South Asia have promoted sweet potato as an alternative to rice in the production of noodles. Similar strategies were used extensively during the Irish potato famine to increase the consumption of imported maize that many Irish initially refused to eat, as they neither liked nor knew how to prepare it (Woodham-Smith, 1964).

Taking local taste differences seriously also has ramifications for estimating Computational General Equilibrium models. These have become a common tool for policymakers to understand the impacts of various trade liberalization scenarios. The use of home-biased Armington preferences, where consumers prefer domestic to foreign varieties of any good, has become commonplace in order to match the observed trade flows. Welfare effects hinge critically on the elasticity of substitution between foreign and domestic varieties, yet such preferences are ad hoc and improbable for homogenous agricultural commodities. Developing my model of habit formation that links local endowments with cross-price elasticities of substitution can produce Armington-like results but with a firm theoretical grounding for agricultural goods.

Tastes matter for trade. Neglecting their role overstates the caloric gains from agricultural trade liberalization, and masks potential nutritional losses for the poorest members of society.

References

Acemoglu, Daron, "Directed Technical Change," *Review of Economic Studies*, 2002, *69* (4), 781–809.

Anderson, Kym and Ernesto Valenzuela, "Do Global Trade Distortions Still Harm Developing Country Farmers?," *Review of World Economics*, 2007, *143* (1), 108–139.

Armington, Paul S., *A Theory of Demand for Products Distinguished by Place of Production.*, Washington DC: International Monetary Fund, Research Dept, 1968.

Barker, David J., *The Fetal and Infant Origins of Adult Disease*, London: BMJ Publishing Group, 1992.

Becker, Gary S. and Kevin M. Murphy, "A Theory of Rational Addiction," *The Journal of Political Economy*, 1988, *96* (4), 675–700.

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

Blanciforti, Laura and Richard Green, "An Almost Ideal Demand System Incorporating Habits: an Analysis of Expenditures on Food and Aggregate Commodity Groups," *The Review of Economics and Statistics*, *65.*

Conley, Tim G., "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Deaton, Angus, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Baltimore: Johns Hopkins University Press, 1997.

and Jean Dreze, "Nutrition in India: Facts and Interpretations," RPDS Working Papers, Princeton University 2008.

and John Muellbauer, "An Almost Ideal Demand System," *American Economic Review*, 1980, *70* (3), 312–326.

and , *Economics and Consumer Behavior*, Cambridge: Cambridge University Press, 1980.

Dev, S. Mahendra and Robert E. Evenson, "Rural Development in India: Rural, Non-farm, and Migration," SCID Working Paper 187 2003.

Du, Shufa, Bing Lu, Fengying Zhai, and Barry M. Popkin, "A New Stage of the Nutrition Transition in China," *Public Health Nutrition*, 2006, *5* (1a), 169–174.

Fieler, Ana Cecilia, "Non-Homotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," 2009. http://www.sas.upenn.edu/~afieler/.

Fogel, Robert W., "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy," National Bureau of Economic Research Working Paper 4638 1994.

Hayami, Yujiro and Vernan W. Ruttan, *Agricultural Development: An International Perspective*, John Hopkins University Press, 1971.

Hunter, Linda and James Markusen, "Per-Capita Income as a Determinant of Trade," in Robert Feenstra, ed., *Empirical Methods for International Trade*, MIT Press, 1987, chapter 4, pp. 89–109.

Jensen, Robert T. and Nolan H. Miller, "Giffen Behavior and Subsistence Consumption," *American Economic Review*, 2008, *98* (4), 1553–1577.

Jha, Raghbendra, K. V. Bhanu Murthy, and Anurag Sharma, "Market Integration in Wholesale Rice Markets in India," ASARC Working Papers 2005-03 2005.

Kelder, Steven H, "Longitudinal Tracking of Adolescent Smoking, Physical Activity, and Food Choice Behaviors," *American Journal of Public Health*, 1994, *84* (7), 1121–1126.

Mennella, Julie, Coren Jagnow, and Gary Beauchamp, "Prenatal and Postnatal Flavor Learning by Human Infants," *Pediatrics*, 2001, *107* (6).

Munshi, Kaivan and Mark Rosenzweig, "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth," National Bureau of Economic Research Working Paper 14850 2009.

Nunn, Nathan and Nancy Qian, "The Potato's Contribution to Population and Urbanization: Evidence from an Historical Experiment," National Bureau of Economic Research Working Paper 15157 2009.

Rosenzweig, Mark and Oded Stark, "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India," *The Journal of Political Economy*, 1989, *97* (4), 905.

Ruffin, Roy and Ronald W. Jones, "Protection and Real Wages: The Neoclassical Ambiguity," *Journal of Economic Theory*, 1977, *14* (2), 337–48.

Sen, Amartya K., *Development as Freedom*, Oxford: Oxford University Press, 1999.

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

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

Stock, James H. and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," National Bureau of Economic Research Working Paper T0284 2002.

Stone, Richard, *Quantity and Price Indexes in National Accounts*, Paris: Organisation for European Economic Co-operation, 1956.

Woodham-Smith, Cecil, *The Great Hunger: Ireland 1845-1849*, New York: Signet, 1964.

Appendices (Not For Publication)

A Background on Agricultural Trade in India

I briefly review the current state of Indian agricultural trade before assessing the potential impact of domestic liberalization. Despite wide ranging economic reforms over the last two decades, India's agricultural sector remains highly restricted. While there has been new legislation at the national (Union) level to liberalize domestic markets, these measures have been applied erratically at best because agricultural policy is under the exclusive constitutional remit of state governments.[1]

Interventionist food policies were initially enacted in response to the perceived failures of private trade in the Bengal famine of 1943. The Essential Commodities Act (1955) entitles both governments and states to impose restrictions on "trade and commerce in, and the production, supply and distribution of foodstuffs."[2] Other agricultural acts control to whom farmers and traders are allowed to sell and at what price. All traders require licenses, have restricted access to credit and must follow over 400 rules that govern food trade (Planning Commisson of India, 2001).

Internal trade is further restrained through state tariffs and district-level entry taxes, Octroi, collected at often corrupt checkpoints (Das-Gupta, 2006). This is in addition to the extremely poor transport infrastructure across India, which is perhaps the biggest hindrance to trading bulky agricultural goods within the country. State governments are also directly involved in the purchase and sale of food. The Commission on Agricultural Costs and Prices sets minimum support prices for farmers that are only available in certain regions, while state levies require private mills to supply grain at a fixed price, which is then sold to the poor through the Public Distribution System at prices chosen by each state. Jha et al. (2005) discuss these numerous restrictions in more detail, and show that as a result wholesale rice markets across India are not integrated. The lack of integration is evident in the NSS data, in which the dispersion of regional prices actually increased between 1987-88 and 2004-05.[3]

[1]For example, the Agricultural Produce Marketing Acts was amended in 2003 to allow farmers to sell their produce directly to buyers for the first time. Only about half of the states have so far incorporated the amendment and in most cases with substantial changes.

[2]FAO (2005) details some of the numerous state-level and even district-level restrictions that remain.

[3]The average over 52 foods of the cross-regional coefficients of variation of rural median food prices rose from 0.51 in 1987-88 to 0.53 in 2004-05. Similarly, the average pairwise correlation between the median prices of the 52 foods in any two regions declined from 0.85 to 0.83 between the two surveys.

Although there has been little progress reforming the domestic market, if India had fully liberalized all external trade, the domestic agricultural market would have become integrated. However, external agricultural trade has only seen limited reform in the years following India's 1991 liberalization. The initial tariff reductions did not cover agricultural goods at all. The impetus for agricultural liberalization came from the Agreement on Agriculture, which India committed to as a founding member of the WTO. This agreement required the conversion of all non-tariff barriers and quantitative restrictions into tariffs by 2002, but left domestic support untouched. However, tariff levels were set sufficiently high to choke imports in all but pulses and oilseeds.[4] As a result, the FAO (2003) reports that there was little impact from the liberalization of agricultural trade under the Agreement on Agriculture between 1997 and 2002.[5]

India still maintains high tariffs, agricultural import monopolies, state trading enterprises and export restrictions that maintain a "highly interventionist agricultural development policy regime" (Athukorala, Prema-chandra, 2005). Accordingly, alongside the domestic restraints detailed above, agricultural trade within India remains highly restricted, and internal markets are far from integrated.[6]

B Data Sources

The NSSO data used in both empirical sections of the paper are described in section 3.1. To measure agricultural endowments, I use district-level agricultural data from Indian Harvest produced by the Centre for Monitoring Indian Economy, aggregated up to NSSO regions. Further regional data come from the Indian District Database (Vanneman and Barnes, 2000) and the India Agriculture and Climate Data Set (Sanghi et al., 1998), while weather data come from Willmott and Matsuura (2001).

C A Simple Model of Habit Formation and Trade

C.1 Proof of Hypothesis 1:

In the specific factors model, the proof that preferences become correlated with relative resource endowments proceeds as follows. The production functions for rice and wheat have constant returns to scale and diminishing returns. One unit of each good provides one calorie.

I will first model the production side of the economy. The population is divided into laborers and factor (land) owners. Each labor owner possesses one unit of labor and there is L total labor in the economy. The distribution of rice land, T_r, and wheat land, T_w, among the

[4]In these two categories India is not self-sufficient and the government itself controls a substantial portion of imports via government agencies. According to (Gulati, 1998), the Indian Government followed the following rule: "Allow imports if there was a net deficit and allow exports if there was a comfortable surplus."

[5]Agricultural exports did, however, respond positively to the 20 percent devaluation of the rupee in 1991.

[6]Therefore, my theoretical mechanism cannot explain the decline in caloric intake that has occurred across India in the last 20 years. In fact relative prices across regions have moved in the opposite direction to that suggested by relative endowments. For example rice was already relatively cheap in large rice growing areas, and has become more so over the reform period.

population will be left unmodelled. Q_r and Q_w are the outputs of the two goods produced from labor and their specific factor:

$$Q_r = f_r(T_r, L_r),$$
$$Q_w = f_w(T_w, L_w),$$
$$\frac{\partial f_i}{\partial T_i} > 0, \frac{\partial f_i}{\partial L_i} > 0, \frac{\partial^2 f_i}{\partial T_i^2} < 0, \frac{\partial^2 f_i}{\partial L_i^2} < 0, \frac{\partial^2 f_i}{\partial T_i \partial L_i} > 0.$$

I normalize the price of wheat to 1:

$$p_w = 1, p_r = p.$$

Labor is divided between rice and wheat production and factor clearing implies that all the labor in the economy is exhausted. Labor is free to move between the two sectors and so marginal returns in the two sectors are equalized:

$$L = L_r + L_w,$$
$$p\frac{\partial f_r}{\partial L_r} = \frac{\partial f_w}{\partial L_w},$$
$$\frac{\partial L_r}{\partial p} = \left[\frac{-\frac{\partial^2 f_w}{\partial L_w^2}}{\frac{\partial f_r}{\partial L_r}} - \frac{\frac{\partial f_w}{\partial L_w}}{(\frac{\partial f_r}{\partial L_w})^2}\frac{\partial^2 f_r}{\partial L_r^2}\right]^{-1} > 0.$$

The specific factors T_w and T_r are fixed over generations. However, I calculate the derivatives with respect to the specific factors to see how initial endowments affect the initial equilibrium:

$$p\frac{\partial^2 f_r}{\partial L_r \partial T_r}\partial T_r + p\frac{\partial^2 f_r}{\partial L_r^2}\partial L_r = -\frac{\partial^2 f_w}{\partial L_w^2}\partial L_r + \frac{\partial^2 f_w}{\partial L_w \partial T_w}\partial T_w,$$
$$\frac{\partial L_r}{\partial T_r} = -\frac{p\frac{\partial^2 f_r}{\partial L_r \partial T_r}}{\frac{\partial^2 f_w}{\partial L_w^2} + p\frac{\partial^2 f_r}{\partial L_r^2}} > 0, \frac{\partial L_r}{\partial T_w} = \frac{\frac{\partial^2 f_w}{\partial L_w \partial T_w}}{\frac{\partial^2 f_w}{\partial L_w^2} + p\frac{\partial^2 f_r}{\partial L_r^2}} < 0.$$

I define the relative production of rice, r, as z. Individual producers take prices as given, and so I can calculate the relative supply response to price changes, as well as to differing quantities of the initial endowments:

$$z = \frac{Q_r}{Q_w} = \frac{f_r(T_r, L_r)}{f_w(T_w, L_w)},$$
$$\frac{dz}{dp} = \frac{1}{f_w}\frac{df_r}{dp} - \frac{f_r}{f_w^2}\frac{df_w}{dp} = \frac{1}{f_w}\frac{\partial f_r}{\partial L_r}\frac{\partial L_r}{\partial p} + \frac{f_r}{f_w^2}\frac{\partial f_w}{\partial L_w}\frac{\partial L_r}{\partial p} > 0,$$
$$\frac{dz}{dT_r} = \frac{1}{f_w}\frac{df_r}{dT_r} - \frac{f_r}{f_w^2}\frac{df_w}{dT_r} = \frac{1}{f_w}\frac{\partial f_r}{\partial T_r} + \frac{f_r}{f_w}\frac{\partial f_r}{\partial L_r}\frac{\partial L_r}{\partial T_r} + \frac{f_r}{f_w^2}\frac{\partial f_w}{\partial L_w}\frac{\partial L_r}{\partial T_r} > 0,$$
$$\frac{dz}{dT_w} = \frac{1}{f_w}\frac{df_r}{dT_w} - \frac{f_r}{f_w^2}\frac{df_w}{dT_w} = \frac{1}{f_w}\frac{\partial f_r}{\partial L_r}\frac{\partial L_r}{\partial T_w} + \frac{f_r}{f_w^2}\frac{\partial f_w}{\partial L_w}\frac{\partial L_r}{\partial T_w} - \frac{f_r}{f_w^2}\frac{\partial f_w}{\partial T_w} < 0.$$

I now turn to modelling the demand side of the economy. I assume demand is homothetic so that the distribution of factors across the economy does not impact relative consumption decisions. Relative tastes for rice, $tastes_r$, increase the budget share spent on rice,

$bshare_r = \theta_r$, conditional on the relative rice price:

$$\theta_r = tastes_r + h_r(p) \text{ where } h_r'(p) \leq 0.$$

I define the relative consumption of r as γ. Individual consumers take prices as given, and so I can calculate the price response of relative demand, as well as the impact of changes in tastes:

$$\gamma = \frac{\theta_r}{p(1-\theta_r)} = \frac{tastes_r + h_r(p)}{p(1 - tastes_r - h_r(p))},$$

$$\frac{d\gamma}{dp} = \frac{p(1-\theta_r)h_r'(p) - \theta_r((1-\theta_r) - ph_r'(p))}{p^2(1-\theta_r)^2} < 0,$$

$$\frac{d\gamma}{dtastes_r} = \frac{1}{p(1-\theta_r)^2} > 0.$$

Market clearing equilibrates relative supply and demand under autarky, with the equilibrium values superscripted with a:

$$z^a(p, T_r, T_w) = \gamma^a(tastes_r, p).$$

I will use the derivatives calculated for relative supply and demand to assess the impact of relative endowments on prices and tastes.[7]

Two different regions at the beginning of time have the same preferences that favor each good equally and technologies are identical in the two sectors. Region 1 is endowed with $T + x$ units of rice growing land T_r and T units of wheat growing land T_w. Region 2 has the reverse relative endowments, nT units of rice growing land T_r and $n(T + x)$ units of wheat growing land T_w. The population of region 1 is L, while region 2 has a population of nL. Region 2 is much bigger than region 1 as n is large.

In a hypothetical economy with the balanced endowments, $T = T_w = T_r$, raising the endowment of T_r to region 1's initial endowment will raise the relative production of rice since $\frac{dz}{dT_r} > 0$, bringing the economy out of equilibrium. The price will have to fall to reduce the relative production and raise the relative consumption of rice until equilibrium is reestablished because $\frac{dz}{dp} > 0$ and $\frac{d\gamma}{dp} < 0$. The opposite effect will occur when the endowment of T_w is increased to obtain region 2's initial endowment.[8] Therefore in the first generation, the region where rice is relatively abundant consumes relatively more rice and has a lower relative price for rice.

Tastes for rice, $tastes_{r,t}$, change over generations t through the process of habit formation. They are positively related to the past generation's relative consumption through adult tastes that favor foods consumed as a child:

$$tastes_{r,t} = g(\gamma_{t-1}^a) \text{ with } g'(\gamma_{t-1}^a) > 0.$$

[7]These were partial equilibrium derivatives as I did not impose equality between relative supply and relative demand until now.

[8]The homogenous of degree one production functions mean that the two regions are comparable since the scaling factor n only affects absolute quantities, not the relative measures referred to here.

In the second generation, habit formation increases the tastes for rice in region 1 (the relatively rice abundant region) compared to region 2, as γ^a_{t-1} is larger in region 1.[9]

This proves hypothesis 1: *A region develops tastes inversely related to the relative prices it faces. Therefore, tastes will become positively correlated with a region's relative resource endowments.*

C.2 Existence of Steady State with $P_s < 1$:

For the existence of a steady state I require that a fixed point exists where relative consumption does not change with an increase in tastes, so that

$$tastes_{r,t} = g(\gamma^a_t(tastes_{r,t}, p_t)).$$

For stability and a steady state which converges without oscillating, an increase in tastes today must lead to a less than proportional increase in tastes tomorrow, as tastes approach the steady state from below:

$$tastes_{r,t} = g(\gamma^a_{t-1}(tastes_{r,t-1}, p_{t-1})),$$
$$0 < \frac{dg(\gamma^a_{t-1}(tastes_{r,t-1}, p_{t-1}))}{dtastes_{r,t-1}} < 1.$$

I assume such a fixed point exists and is stable, and that the steady state has a price less than 1 and so is an interior steady state where both rice and wheat continue to be consumed. These steady state values are reached in generation s:

$$tastes_{rs} = g(\gamma^a_s(tastes_{rs}, p_s)),$$
$$p_s < 1.$$

These conditions are characterized for the Cobb-Douglas case in appendix C.5.

C.3 Proof of Hypothesis 2:

Assuming the rise in the rice price necessary to equilibrate the economy is not so large that p rises above the hypothetical balanced-endowment economy's autarky price of $p = 1$, tastes in period 2 will be inversely correlated with autarky prices. The aggregate caloric maximum is at $p = 1$, found by setting $\frac{dcalories}{dr} = 0$ along the edge of the production possibilities frontier. Caloric intake increases with a rise in p up to that point as the production possibilities frontier is concave.

$$calories = r + w(r)$$
$$\frac{dcalories}{dr} = 1 + \frac{dw}{dr} = 1 - p$$

The rice price increases in region 1 with habit formation, as the increased tastes for rice raise demand for rice. The rice price was initially below 1 and therefore the aggregate caloric intake increases with habit formation as long as rice remains relatively cheap ($p < 1$).

[9] This increases the relative consumption of rice, bringing about a price response (since endowments are fixed). To bring production into equilibrium, the price of rice must rise as $\frac{dz}{dp} > 0$ and $\frac{d\gamma}{dp} < 0$.

C.4 Proof of Hypothesis 3:

Region 1 integrates with the much larger region 2, and takes region 2's prices as given at the time of trade liberalization. Post trade values are superscripted with an asterisk. The new world price is $p^* > 1$. Tastes are fixed and I look only at region 1 and the generation initially affected by trade liberalization. Upon trade liberalization, rice is now relatively more expensive ($p^* > 1$) and caloric intake increases in wheat consumption. With an exogenous world price, the consumption decision can be separated from the production decision and relative consumption is determined by

$$\gamma = \frac{tastes_{rt} + h_r(p^*)}{p(1 - tastes_{rt} - h_r(p^*))}, \text{ where } \frac{\partial \gamma}{\partial tastes_{rt}} > 0,$$

with the budget set defined by

$$p^* r^* + w^* \leq p^* Q_r^* + Q_w^*.$$

Caloric intake is now decreasing in rice consumption and the taste for rice, as wheat is relatively cheap:

$$calories^* = r^* + w^* = (1 - p^*)(\frac{tastes_{rt} + h_r(p^*)}{p^*})(p^* Q_r^* + Q_w^*) + p^* Q_r^* + Q_w^*,$$

$$\frac{dcalories^*}{dtastes_{rt}} = \frac{(1 - p^*)}{p^*}(p^* Q_r^* + Q_w^*) < 0.$$

With habit formation, steady state $tastes_{rs}$ favor rice, and so caloric intake is lower than if tastes were neutral as in the first generation (without habit formation, the neutral tastes of the first generation, $tastes_{r1} = \frac{1}{2}$, are also the tastes of every subsequent generation). The relatively higher consumption of rice reduces caloric intake compared to the no habit formation society facing the same world price p^* and hence the same budget set.

The autarky price in region 2 will also be reduced by habit formation (as tastes for wheat bid up the price of wheat). Therefore, the caloric intake post trade in a world without habits will be even larger than if the post trade price was p^*, as wheat will be even cheaper at the time of liberalization and the greater production gains from trade will allow consumption at a point beyond that obtainable in the habit formation world:

$$\frac{dcalories^*}{dp^*} = (1 - p^*)(\frac{-(tastes_{rt} + h_r(p^*))}{p^{*2}} Q_w^* + (\frac{h_r'(p^*)}{p^*})(p^* Q_r^* + Q_w^*)) + (Q_r^* - r^*) > 0,$$

as $p^* > 1$ and region 1 exports rice, $(Q_r^* - r^*) > 0$.

Habit formation increased the aggregate caloric intake of each generation up to the steady state generation s (hypothesis 2). Therefore, since aggregate caloric intake was lower pre trade and higher post trade without habit formation, this implies that habit formation reduces both the aggregate caloric gains from trade as well as the caloric elasticities with respect to trade liberalization.

This proves hypothesis 3 for the case where $p^* > 1$: *Habit formation reduces the short-run* *aggregate caloric elasticities with respect to trade liberalization.*

If $p^* = 1$ in both the habit and no habit worlds, the caloric intake post trade is identical, but was higher pre trade with habit formation so that the absolute gains and elasticities are also reduced. If $p_s < p^* < 1$, the impact is ambiguous, although the change in caloric intake and the caloric elasticity are still likely to be reduced as long as the relative endowment $\frac{T_r}{T_w}$ is not very close to 1, as the economy-wide production gains are much smaller with habit formation (as p_s is higher than the autarky price without habit formation, p_1).

C.5 Characterization of the Steady State with Cobb-Douglas Functional Forms

Here I solve the model outlined in the theory section with specific functional forms. For simplicity I choose Cobb-Douglas production functions exhibiting constant returns to scale, so that $h_r(p) = 0$. The basic model is about preferences changing and so to abstract from other differences between the two goods I make the two production technologies equally labor intensive. I focus on region 1 where $T_r > T_w$:

$$Q_{rt} = L_{rt}^{\beta} T_r^{1-\beta},$$

$$Q_{wt} = L_{wt}^{\beta} T_w^{1-\beta},$$

$$0 < \beta < 1.$$

Let $p_{rt} = p_t$ and the price of one unit of wheat be the numeraire $p_{wt} = 1$. Factors earn marginal products in competitive equilibrium, resulting in the following factor pricing equations, where ω_t are wages, π_{rt} the returns to rice land and π_{wt} the returns to wheat land:

$$\omega_t = \frac{dp_t Q_{rt}}{dL_{rt}} = p_t \beta (\frac{T_r}{L_{rt}})^{1-\beta}, \tag{9}$$

$$\omega_t = \frac{dQ_{wt}}{dL_{wt}} = \beta (\frac{T_w}{L_{wt}})^{1-\beta}, \tag{10}$$

$$\pi_{rt} = \frac{dp_t Q_{rt}}{dT_r} = p_t (1-\beta)(\frac{L_{rt}}{T_r})^{\beta}, \tag{11}$$

$$\pi_{wt} = \frac{dQ_{wt}}{dT_w} = (1-\beta)(\frac{L_{wt}}{T_w})^{\beta}. \tag{12}$$

Factor clearing implies that all the labor in the economy is exhausted. Wages will be equalized across both sectors as workers are mobile. By feeding in this factor clearing condition I obtain relative prices as a function of the labor in each sector:

$$L_{rt} + L_{wt} = L, \tag{13}$$

$$p_t = (\frac{T_w}{T_r} \frac{L_{rt}}{(L - L_{rt})})^{1-\beta}. \tag{14}$$

Utility is Cobb-Douglas and so budget shares are independent of prices. The budget share spent on rice is therefore simply $tastes_{rt}$ and in the first generation tastes are neutral and $tastes_{r1} = \frac{1}{2}$:

$$U_t(r_t, w_t) = r_t^{tastes_{rt}} w_t^{1-tastes_{rt}}.$$

I provide a functional form for habit formation, where tastes for rice depend on the previous generations relative consumption of rice, but with a dampening parameter ν. This determines how much tastes for rice respond to an increase in relative rice consumption:

$$tastes_{rt} = \frac{(r_{t-1})^{\nu}}{(r_{t-1})^{\nu} + (w_{t-1})^{\nu}}, \quad \nu > 0. \tag{15}$$

The Cobb-Douglas preferences imply the following demand functions, where m_t is total factor income:

$$r_t = tastes_{rt}\frac{m_t}{p_t},$$
$$w_t = (1 - tastes_{rt})m_t.$$

Because demand is homothetic, everyone in the economy spends the same proportion of their income on each good. Therefore product market clearing for good r implies that:

$$L_{rt}^{\beta}T_r^{1-\beta} = tastes_{rt}\frac{(\omega_t L + \pi_{rt}T_r + \pi_{wt}T_w)}{p_t}. \tag{16}$$

Now I solve for equilibrium prices and labor allocation in generation t by combining the product market clearing condition 16 with the factor pricing equations 9-12 and the factor clearing equation 13.

$$L_{rt} = tastes_{rt}L,$$
$$p_t = (\frac{T_w}{T_r}\frac{tastes_{rt}}{(1 - tastes_{rt})})^{1-\beta},$$
$$\omega_t = \beta(\frac{T_w}{L})^{1-\beta}(\frac{1}{1 - tastes_{rt}})^{1-\beta}.$$

These are the prices and labor allocation in generation t conditional on $tastes_{rt}$. However $tastes_{rt}$ is a function of the previous generation's demands with habit formation. By feeding the demands and prices in generation $t-1$ into $tastes_{rt} = \frac{(r_{t-1})^{\nu}}{(r_{t-1})^{\nu}+(w_{t-1})^{\nu}}$, I obtain the difference equation for $tastes_{rt}$:

$$tastes_{rt} = \frac{1}{1 + ((\frac{tastes_{r,t-1}}{(1-tastes_{r,t-1})})^{-\beta}(\frac{T_w}{T_r})^{1-\beta})^{\nu}}. \tag{17}$$

Solving for the steady state, I set $tastes_r = tastes_{rt} = tastes_{r,t-1}$ and rearrange. Interior steady-state values are identified by the subscript s:

$$tastes_{rs} = \frac{1}{1 + (\frac{T_w}{T_r})^{\frac{(1-\beta)\nu}{1-\beta\nu}}},$$
$$p_s = (\frac{T_w}{T_r})^{\frac{(1-\nu)(1-\beta)}{1-\beta\nu}}.$$

The steady-state $tastes_{rs}$ are greater than a half, with tastes favoring rice consumption, as long as $T_r > T_w$ and $\beta\nu < 1$ (tastes do not respond excessively to relative consumption). Therefore tastes develop for the relatively abundant (comparative advantage) good, r in this example. This is hypothesis 1; tastes positively correlate with endowments. The steady state price remains less than 1 if $\nu < 1$.

The local stability of the steady state without oscillation requires that $0 < \frac{df(tastes_{r,t-1})}{tastes_{r,t-1}} < 1$ at the steady state, where $f(tastes_{r,t-1}) = tastes_{rt} = (1+((\frac{tastes_{r,t-1}}{(1-tastes_{r,t-1})})^{-\beta}(\frac{T_w}{T_r})^{1-\beta})^{\nu})^{-1}$ from the difference equation 17:

$$\frac{df(tastes_{r,t-1})}{dtastes_{r,t-1}} = -(1+(\frac{1}{tastes_{r,t-1}}-1)^{\beta\nu}(\frac{T_w}{T_r})^{(1-\beta)\nu})^{-2}[(\frac{T_w}{T_r})^{(1-\beta)\nu}\beta\nu(\frac{1}{tastes_{r,t-1}}-1)^{\beta\nu-1}(-\frac{1}{tastes^2_{r,t-1}})].$$

Feeding in the steady state value of $tastes_{rs}$ and simplifying:

$$\frac{df(tastes_{r,t-1})}{dtastes_{r,t-1}} = \beta\nu.$$

Additionally, there are two more unstable equilibria at $tastes_{r,t-1} = 0$ or $tastes_{r,t-1} = 1$. The unstable nature of these two points can be seen from subtracting $tastes_{r,t-1}$ from equation 17:

$$tastes_{rt} - tastes_{r,t-1} = \frac{1-(\frac{tastes_{r,t-1}}{(1-tastes_{r,t-1})})^{1-\beta\nu}(\frac{T_w}{T_r})^{(1-\beta)\nu}}{\frac{1}{1-tastes_{r,t-1}} + \frac{1}{1-tastes_{r,t-1}}(\frac{tastes_{r,t-1}}{(1-tastes_{r,t-1})})^{-\beta\nu}(\frac{T_w}{T_r})^{(1-\beta)\nu}}.$$

Therefore, for $0 < tastes_{r,t-1} < 1$ (and the conditions $T_r > T_w$ and $\beta\nu < 1$),

$$tastes_{rt} - tastes_{r,t-1} > 0 \; iff \; tastes_{r,t-1} < \frac{1}{1+(\frac{T_w}{T_r})^{\frac{(1-\beta)\nu}{1-\beta\nu}}} = tastes_{rs},$$

$$tastes_{rt} - tastes_{r,t-1} < 0 \; iff \; tastes_{r,t-1} > \frac{1}{1+(\frac{T_w}{T_r})^{\frac{(1-\beta)\nu}{1-\beta\nu}}} = tastes_{rs}.$$

Accordingly, the two equilibria at $tastes_{r,t-1} = 0$ and $tastes_{r,t-1} = 1$ are unstable, and the interior equilibrium at $tastes_{rs}$ is globally stable.

A sufficient condition for tastes to be correlated with endowments, a stable steady state to exist and the steady state relative price of rice to be strictly less than 1 is $\nu < 1$. This rules out the possibility that preferences respond to past consumption to such a large degree that they overturn the resource comparative advantage. In this case the high demand for rice actually makes it relatively more expensive, but it continues to be consumed in ever larger amounts. This is related to how much a consumer values variety, as with $\nu \geq 1$ tastes increase to such an extent that they overwhelm the disutility from consuming a less varied diet. In the empirical section I show that this assumption holds for India, as prices for a food are relatively cheaper in regions where tastes are stronger for that food.

C.5.1 Caloric Impact of Trade Liberalization on Labor

I will now look at landless workers' (owners of one unit of labor only) calorie consumption both in the first generation (the steady state without habit formation) and at the autarky steady state with habit formation. One unit of each good provides one calorie. Therefore total calories consumed, *calories*, equals $r+w$. Feeding the wage into the demand functions, I obtain the total calories consumed at time t for a worker possessing only a single unit of labor:

$$calories_t = tastes_{rt}\frac{\omega_t}{p_t} + (1 - tastes_{rt})\omega_t,$$

$$calories_t = tastes_{rt}^{\beta}\beta(\frac{T_r}{L})^{1-\beta} + (1 - tastes_{rt})^{\beta}\beta(\frac{T_w}{L})^{1-\beta}. \quad (18)$$

I differentiate caloric intake with respect to $tastes_{rt}$. Caloric intake increases as tastes adjust to favor rice ($tastes_{rt}$ rises) as shown in hypothesis 2:

$$\frac{dcalories_t}{dtastes_{rt}} = (\frac{1}{L})^{1-\beta}\beta^2[T_r^{1-\beta}tastes_{rt}^{\beta-1} - T_w^{1-\beta}(1 - tastes_{rt})^{\beta-1}],$$

$$\frac{dcalories_t}{dtastes_{rt}} > 0 \text{ if } (\frac{T_w}{T_r}\frac{tastes_{rt}}{(1 - tastes_{rt})})^{1-\beta} = p_t < 1.$$

What happens to the caloric consumption of landless labor for the adult generation alive at the time a small region at its autarky steady state, s, opens up to trade? The world price favors wheat, $p^* > 1$, and the small region is a price taker. The equalization of wages across the two sectors pins down the relative labor allocation through equation 14. I denote the new post-trade equilibrium values with an asterisk superscript:

$$p^* = (\frac{T_w}{T_r}\frac{L_{rs}^*}{(L - L_{rs}^*)})^{1-\beta},$$

$$L_{rs}^* = \frac{(p^*)^{\frac{1}{1-\beta}}L}{\frac{T_w}{T_r} + (p^*)^{\frac{1}{1-\beta}}},$$

$$\omega_s^* = \beta(\frac{T_w + T_r(p^*)^{\frac{1}{1-\beta}}}{L})^{1-\beta}.$$

I calculate the total caloric intakes before ($calories_s$) and after ($calories_s^*$) trade liberalization, as a function of tastes at the steady state, $tastes_{rs}$:

$$calories_s = tastes_{rs}^{\beta}\beta(\frac{T_r}{L})^{1-\beta} + (1 - tastes_{rs})^{\beta}\beta(\frac{T_w}{L})^{1-\beta},$$

$$calories_s^* = tastes_{rs}\frac{1}{p^*}\beta(\frac{T_w + T_r(p^*)^{\frac{1}{1-\beta}}}{L})^{1-\beta} + (1 - tastes_{rs})\beta(\frac{T_w + T_r(p^*)^{\frac{1}{1-\beta}}}{L})^{1-\beta}.$$

Therefore, the caloric elasticity with respect to trade liberalization is simply:

$$\frac{calories_s^* - calories_s}{calories_s} = \frac{[\frac{tastes_{rs}}{p^*} + (1 - tastes_{rs})]\beta(\frac{T_w + T_r(p^*)^{\frac{1}{1-\beta}}}{L})^{1-\beta}}{tastes_{rs}^{\beta}\beta(\frac{T_r}{L})^{1-\beta} + (1 - tastes_{rs})^{\beta}\beta(\frac{T_w}{L})^{1-\beta}} - 1.$$

Differentiating this expression with respect to $tastes_{rs}$ shows how the caloric elasticity varies with preferences:

$$\frac{d\frac{calories_s^*}{calories_s}}{dtastes_{rs}}N = [\frac{1}{p^*} - 1][\frac{1}{p_s} + (\frac{1 - tastes_{rs}}{tastes_{rs}})] - \beta[\frac{1}{p^*} + (\frac{1 - tastes_{rs}}{tastes_{rs}})][\frac{1}{p_s} - 1],$$

$$N = \frac{[tastes_{rs}^{\beta}(\frac{T_r}{T_w})^{1-\beta} + (1 - tastes_{rs})^{\beta}][\frac{1}{p_*} + (\frac{1 - tastes_{rs}}{tastes_{rs}})]}{(1 + \frac{T_r}{T_w}(p^*)^{\frac{1}{1-\beta}})^{1-\beta}} > 0, p_s = (\frac{T_w}{T_r}\frac{tastes_{rs}}{1 - tastes_{rs}})^{1-\beta}.$$

I compare a society where habits favor the comparative advantage good with a society where there are fixed neutral preferences. To do this I feed in the neutral preferences,

$tastes_{r1} = tastes_{rs} = \frac{1}{2}$, and calculate the change in caloric elasticity with respect to trade liberalization as tastes for the comparative advantage good increase. This does not assume that preferences evolve precisely as described in the previous section, only that they are positively related to past relative consumption:

$$\frac{d\frac{calories_s^*}{calories_s}}{dtastes_{rs}}N = [\frac{1}{p^*} - 1][\frac{1}{p_s} + 1] - \beta[\frac{1}{p^*} + 1][\frac{1}{p_s} - 1]. \quad (19)$$

I sign this expression when $T_r > T_w$ and $p^* > p_s$, so the area has a comparative advantage in its relatively abundant good r. Here $tastes_{rs} > \frac{1}{2}$ with habit formation. The standard case is $p^* \geq 1 > p_s$ where the world is evenly endowed with the two factors or has a relatively more of the factor required to produce good r. Both terms of equation 19 are negative (or zero for the first term if $p^* = 1$). Therefore, the elasticity of caloric intake with respect to trade liberalization is reduced when preferences develop to favor the comparative advantage good.[10] This is hypothesis 3 in the paper.

C.6 Welfare Implications of Model with Quality Improvements

The utility function can be rewritten as follows:

$$U(r,w) = \widetilde{r}^{\frac{1}{2}}\widetilde{w}^{\frac{1}{2}},$$
$$\widetilde{r} = A_t r^{2\alpha_t},$$
$$\widetilde{w} = A_t w^{2(1-\alpha_t)}.$$

The actual quantities of rice and wheat consumed are r and w, while $\widetilde{r}$ and $\widetilde{w}$ are the quality-adjusted quantity consumed using the local technologies. This model is isomorphic to the model where tastes change over generations. Here, relative technologies for converting raw food ingredients into meals, α_t and $(1 - \alpha_t)$, respond to the previous generation's physical consumption in the same way tastes responded to past relative consumption. A_t is the absolute technological progress in generation t.[11] This model is discussed in section 2.5, and since preferences are fixed, welfare gains from trade can be evaluated.

Proceeding in the same way as above. I show that the welfare elasticity of trade liberalization for landless labor decreases if transformation technologies favor the relatively abundant

[10] In the case $p_s < p^* < 1$, so that the area has a comparative advantage in rice but the world endowment also favors the production of rice, the sign cannot be determined. The sign will generally be negative unless p^* is much smaller than 1 or the price change $p^* - p_s$ is very small. In these cases the rice-loving preferences that develop with habit formation are still more suited to world prices than the neutral preferences, and so being less willing to substitute into the expensive calorie source (wheat) actually makes consumers better off in caloric terms.

[11] Over generations there will be absolute utility gains as the transformation technologies improve. This enters the utility function multiplicatively, while α_t determines how those technology gains are shared between the two goods. Since I am analyzing the instantaneous gains from trade upon liberalization, the equivalent comparison to the case of habits and no habits becomes a situation where there is equal total technological progress A_t, but it is either primarily focused on the more consumed good or shared evenly between the two goods.

comparative advantage good: $\frac{d(U^W - U^A)/U^A}{d\alpha_s} > 0$ if $T_r > T_w$ and $p^* > 1 > p_s$:[12]

$$U^A = A_s(\alpha_s \frac{\omega_s}{p_s})^{\alpha_s}((1-\alpha_s)\omega_s)^{1-\alpha_s},$$

$$U^W = A_s(\alpha_s \frac{\omega_s^*}{p^*})^{\alpha_s}((1-\alpha_s)\omega_s^*)^{1-\alpha_s},$$

$$\frac{U^W - U^A}{U^A} = \frac{(\alpha_s \frac{1}{p^*})^{\alpha_s}((1-\alpha_s))^{1-\alpha_s}\beta(\frac{T_w + T_r(p^*)^{\frac{1}{1-\beta}}}{L})^{1-\beta}}{(\alpha_s^\beta \beta(\frac{T_r}{L})^{1-\beta})^{\alpha_s}((1-\alpha_s)^\beta \beta(\frac{T_w}{L})^{1-\beta})^{1-\alpha_s}} - 1.$$

The log change is a monotonic transform of $\frac{U^W - U^A}{U^A}$ and is more easily differentiated:

$$\log U^W - \log U^A = \log \frac{\alpha_s^{(1-\beta)\alpha_s}(1-\alpha_s)^{(1-\beta)(1-\alpha_s)}(\frac{1}{p^*})^{\alpha_s}(T_w + T_r(p^*)^{\frac{1}{1-\beta}})^{1-\beta}}{T_r^{(1-\beta)\alpha_s}T_w^{(1-\beta)(1-\alpha_s)}},$$

$$\frac{d(\log U^W - \log U^A)}{d\alpha_s} = (1-\beta)[\log \frac{T_w}{T_r}(\frac{1}{p^*})^{\frac{1}{1-\beta}}\frac{\alpha_s}{(1-\alpha_s)}].$$

Therefore, if $\alpha_s = \frac{1}{2}$, the counterfactual where local transformation technologies develop independent of relative consumption, this derivative is negative when $T_r > T_w$ and $p^* > 1 > p_s$. Increasing α_s, so that transformation technologies favor the good made with the relatively abundant factor, reduces the welfare elasticity of trade liberalization. This is the amended hypothesis 3* in the paper.

C.7 A Parametrized Example with Non-Homothetic Preferences

In this section I relax the assumption of homothetic Cobb-Douglas preferences, and show that for a simple two-period model with non-homothetic preferences, landless labor develops the strongest preferences for the local comparative advantage food, and that at the time of trade liberalization landless labor will lose in caloric terms, although the group would have gained at the time of liberalization in a world without habit formation.

The production side is identical to the model outlined in section C.5.1, and so the factor pricing equations 9-12 and the factor clearing equation 13 continue to hold as before. Similarly, tastes evolve as in equation 15. However, the product market clearing equation is altered by assuming a more general demand specification.

My demand specification comes from the Almost Ideal Demand System used in the empirical section of the paper (equation 3),

$$bshare_{rt}(m_t, .) = tastes_{rt} + \gamma \ln p_t + b \ln m - b\alpha_0 + btastes_{rt} \ln p_t + \frac{b}{2}\gamma \ln p_t \ln p_t,$$

where m is factor income. The Cobb-Douglas specification used above is a special case where $\gamma = b = 0$.

[12]With Cobb-Douglas production functions, Melvin and Waschik (2001) show that labor's welfare is minimized at autarky prices, and so any price changes are welfare improving. However, here i show that the welfare gain with trade is smaller in a world with habit formation. For other constant elasticity of substitution production functions, Melvin and Waschik (2001) show that welfare losses for labor upon trade liberalization are possible.

I assume that a proportion ϕ of the population is landless, and posses only 1 unit of labor each. The landowner population, $(1-\phi)L$, possess one unit of labor and a proportion $\frac{1}{(1-\phi)L}$ of each of the two specific land factors. Product market clearing for good r implies:

$$L_{rt}^{\beta}T_r^{1-\beta} = \frac{(bshare_{rt}(\omega_t,.)\omega_t\phi L + bshare_{rt}(\omega_t + \frac{\pi_{rt}T_r+\pi_{wt}T_w}{(1-\phi)L},.)(\omega_t(1-\phi)L + \pi_{rt}T_r + \pi_{wt}T_w)}{p_t}.$$

I can now solve for all the unknown variables in the first period using the product market clearing equation alongside equations 9-13, and feeding in $tastes_{r1} = \frac{1}{2}$. I calculate tastes in period 2 using equation 15, and then solve for the unknown variables in the second period in a similar manner.

For a reasonable set of parameters,[13] landless labor has stronger preferences for rice in the second period, the food that the region has a comparative advantage at producing. Intuitively, as the landless are poorer than landowners, they consumed a proportionally larger amount of the relatively cheap rice in the first period. If trade liberalization with a world that has a comparative advantage in wheat occurs during the second period, landless labor will see their caloric consumption decline at the time of liberalization. However, in the case of no habit formation (the equivalent to trade liberalization in the first period when $tastes_{r1} = \frac{1}{2}$), landless labor would gain in caloric terms at the time of liberalization.

D Robustness Results

D.1 Robustness of Taste Estimates

There are several econometric reasons why the LA/AIDS taste estimates may be inconsistent. The endogeneity of prices is a general issue in demand estimation, with the literature highlighting differentiated products as a particular concern since these often have promotions and quantity discounts (Dhar et al., 2003). As all the foods in the sample are raw agricultural commodities, this should not be a substantial problem for rural India. My paper details how tastes vary at the level of the agro-climatic region, and these regional taste differences are picked up by the regional dummy variables. However, if tastes also vary at the village level and village markets are not fully integrated within regions, village taste peculiarities will change local demand and therefore local prices. Since I cannot include a village-level taste dummy and village-level prices, this is a case of omitted variable bias.

To clarify the situation, I rewrite equation 4 by sweeping out the region dummies, and omitting the expenditure and demographic terms for neater exposition:

$$\widetilde{bshare}_{gri} = \sum_{g'} \gamma_{gg'}\widetilde{\ln p}_{g'riv} + \tilde{\varepsilon}_{gri}, \tag{20}$$

$$\hat{\boldsymbol{\gamma}}_g = \boldsymbol{\gamma}_g + (\frac{1}{n}\sum_i \widetilde{\ln \boldsymbol{p}}_{riv}\widetilde{\ln \boldsymbol{p}}'_{riv})^{-1}\frac{1}{n}\sum_i \widetilde{\ln \boldsymbol{p}}_{riv} \cdot \tilde{\varepsilon}_{gri}. \tag{21}$$

[13] These parameters are as follows: $L = 10, T_w = 5, T_r = 15, \nu = 0.8, \beta = 0.4, b = -0.1, \alpha_0 = 0.4, \gamma = -0.3, P = 1.5, \phi = 0.3$.

The price faced by all households i in village v is $p_{g'riv}$.[14] The $G \times 1$ vector of the 52 coefficients on prices, $\gamma_{gg'}$s, for the good g regression is $\boldsymbol{\gamma}_g$. I define a region r average for a generic variable $x_{g'}$ as $\overline{x}_{g'r}$, a region demeaned variable $\widetilde{x}_{g'ri} = x_{g'ri} - \overline{x}_{g'r}$ and a $G \times 1$ vector $\widetilde{\boldsymbol{x}}_{ri}$ of the G variables $\widetilde{x}_{g'ri}$ for each household i. The estimated parameters from this regression will be identical to those when the region effects are included via the Frisch-Waugh theorem. I can recover the fixed effects by using the OLS first order conditions from regression 4:[15]

$$\widehat{tastes}_{gr} = \overline{bshare}_{gr} - \sum_{g'} \widehat{\gamma}_{gg'} \overline{\ln p}_{g'r}. \tag{22}$$

The village-level taste deviations (α_{grv}) for village v are the omitted variable and are mean zero at the region level, $\widetilde{\varepsilon}_{gri} = \alpha_{grv} + \widetilde{\epsilon}_{gri}$. The price is determined by equalizing village-level aggregate supply y_v^s and aggregate demand y_v^d:

$$y_v^s = \sum_{j \in v} z_{grj}(p_{grv}) + Z_{gr}(p_{grv}),$$

$$y_v^d = \sum_{i \in v} \frac{food_i \times bshare_{gri}(\alpha_{grv})}{p_{grv}} + Y_{gr}(p_{grv}).$$

In the equations above, z_{grj} is one producer j's supply of good g, Z_{gr} is the out-of-village supply that increases with the village price p_{grv} and Y_{gr} is the out-of-village demand that decreases with p_{grv}. $E[\widetilde{\ln p}_{griv} \alpha_{grv}] > 0$ since α_{grv} raises $bshare_{gri}$, and so raises the equilibrium price.

Instrumenting the 52 prices requires 52 instruments that are correlated with $\ln p_{grv}$ but uncorrelated with α_{grv}. Hausman (1994) suggests using prices from other markets which have been partly determined by the same supply shifters $Z_{gr}(p_{grv})$ but are not correlated with village tastes α_{grv}.[16] Accordingly, I instrument each village price with the price in a nearby village in the same district, which should be affected by similar supply shocks.[17] The main results are robust to using these instrumented taste measures and are reported in tables D.1 through D.6.[18] For these instrumented results to be consistent, village tastes and hence deviations from regional average tastes cannot be spatially correlated. If they are,

[14] Only one of the village and household identifiers are necessary, and I will use only the village identifier when referring to village level supply and demand. Otherwise I use both identifiers on price terms.

[15] As shown by Kennan (1989), there may also be an additional bias in estimating $\gamma_{gg'}$ from individual demand shocks, such as income, that does not disappear with aggregation. However this bias will become small as long as there is a sufficient village level component to this individual shock.

[16] There has been a heated debate between Hausman and Bresnahan about the validity of these instruments. Most of the discussion centers around whether such promotions as national advertising campaigns shift tastes simultaneously across all markets. This is not an issue here as my food products are generally undifferentiated and I explicitly control for regional taste shifters.

[17] I instrument prices in the village with prices in the next village in the district according to the NSS village number. For the highest numbered village in the district, I use prices in the lowest numbered village.

[18] There are 52 first-stage regressions. These instruments may be somewhat weak. The average first stage F-stat is 13.6. The Kleibergen-Paap rk Wald F-statistic is 0.375 for the full first-stage. Stock and Yogo (2002) do not report critical values for more than 3 endogenous regressors.

more distant prices may be suitable instruments, although these prices will also be much more weakly correlated with village prices.

I can approach the endogeneity of prices in another way and avoid instrumentation altogether. The bias in $\widehat{tastes}_{gr}$ should only increase the dispersion of the taste estimates and not their rank ordering under certain conditions. I derive the bias by combining equation 21 and 22, where $\overline{\ln \boldsymbol{p}}$ is the $G \times R$ matrix of regional prices $\overline{\ln p_{g'r}}$ and $\boldsymbol{tastes}_g$ is the $R \times 1$ vector of regional tastes for good g:

$$\widehat{\boldsymbol{tastes}}_g - \boldsymbol{tastes}_g = -\overline{\ln \boldsymbol{p}}'((\frac{1}{n}\sum_i \widetilde{\ln \boldsymbol{p}}_{riv} \widetilde{\ln \boldsymbol{p}}'_{riv})^{-1} \frac{1}{n}\sum_i \widetilde{\ln \boldsymbol{p}}_{riv} \cdot \alpha_{griv}).$$

If this bias increases monotonically with $tastes_{gr}$, $\frac{d(\widehat{tastes}_{gr} - tastes_{gr})}{dtastes_{gr}} > 0$, then the rank ordering will remain unchanged. To proceed I simplify the problem further and assume that all goods are substitutes, so that village-specific tastes for good g lower the price for good g':

$$E[\widetilde{\ln p}_{griv} \alpha_{grv}] = c_1 > 0,$$
$$E[\widetilde{\ln p}_{g'riv} \alpha_{grv}] = c_2 < 0.$$

I assume the variance-covariance matrix of region-demeaned prices is approximately diagonal, meaning that the deviations from regional average prices within a village for each good are approximately independent:

$$E[\widetilde{\ln \boldsymbol{p}}_{riv} \widetilde{\ln \boldsymbol{p}}'_{riv}]^{-1} \approx \begin{bmatrix} \omega_{g=1} & ... & \omega_1 \\ \omega_1 & ... & \omega_1 \\ \omega_1 & ... & \omega_{g=G} \end{bmatrix},$$
$$\omega_1 \approx 0, \omega_{g'} \approx [\frac{1}{n}\sum_i \widetilde{\ln p}^2_{g'riv}]^{-1} > 0.$$

Finally, I replace $\overline{\ln p_{gr}}$ with its best linear predictor conditional upon regional tastes, $tastes_{gr}$. The theory of habit formation and regional endowments outlined in this paper, and verified for India, provides signs for the taste terms. Strong regional tastes for food g are associated with lower regional prices for that food (both are determined by endowments). Similarly, strong regional tastes for good g are associated with higher prices for good g', as relative endowments are lower:

$$\overline{\ln p_{gr}} = \psi_1 tastes_{gr} + \sum_{g' \neq g} \psi_2 tastes_{g'r} + u_{gr},$$
$$\psi_1 < 0, \psi_2 > 0.$$

Under the null hypothesis, tastes do not evolve through habit formation and are independent of regional prices, $\psi_1 = \psi_2 = 0$. The estimates of regional tastes should then be zero and unbiased.

With this simplifying structure in place I can calculate how the bias in $\widehat{tastes}_{gr}$ changes with the size of $tastes_{gr}$:

$$\frac{d(\widehat{tastes}_{gr} - tastes_{gr})}{dtastes_{gr}} = -\psi_1 \omega_g c_1 - \sum_{g' \neq g} \psi_2 \omega_{g'} c_2 > 0.$$

The dispersion of the taste estimates increases if tastes vary at the village-level and markets within regions are not integrated, but the rank ordering of tastes remains unchanged. Therefore, normalized taste measures across regions should still pick up the relative tastes for food g. Normalized tastes were already used to test the prediction that tastes are positively correlated with endowments and negatively with prices. Results for regression 7 using normalized tastes (mean 0 standard deviation 1) are shown in tables D.5 and D.6. Normalization removes the relative importance of each good in caloric consumption (for example changes in the rice price will have a larger impact on calories than changes in the price of black pepper). Accordingly, instead of the summation measure, I use a correlation between normalized tastes and price changes, weighted by the national budget shares[19] for each good, $\rho_r^T = corr_g(tastes_{gr}, \Delta \ln p_{gr})$. As before, the coefficient on $corr_g(tastes_{gr}, \Delta \ln p_{gr})$ is negative.

Total food expenditure may also be correlated with the demand for individual foodstuffs. Fortunately, I know how much the household spent on other expenditures and this allows me to bound the bias. If food expenditure increases with higher demand for a food, other expenditures will necessarily decline with a fixed income, biasing the coefficients in the other direction. Therefore $food_i$ can be instrumented with other expenditures, and the true coefficients should lie somewhere between the uninstrumented and instrumented results. The estimated coefficients, shown in tables D.1 through D.6, are very similar, suggesting that the endogeneity of food expenditure in the demand estimation is not a major concern. If the measurement error in food and non-food expenditure is independent, this instrumentation strategy also avoids biased parameters that would result from an imperfect measure of food expenditure appearing both in the denominator of the food budget share and on the right hand side of equation 4.

The $\gamma_{gg'}$ terms on prices may also vary by region and be correlated with taste differences. This will lead to biased estimates of $tastes_{gr}$ as the region dummy absorbs the region specific price terms. Including the region specific price effect in the error term makes this bias clear:

$$\varepsilon_{gri} = \sum_{g'} (\gamma_{gg'r} - \gamma_{gg'}) \ln p_{g'r} + \nu_{gri},$$

$$\widehat{tastes}_{gr} = tastes_{gr} + \sum_{g'} (\gamma_{gg'r} - \gamma_{gg'}) \overline{\ln p_{g'r}}.$$

The population-averaged $\gamma_{gg'}$s are consistently estimated assuming the rank condition holds, the mean zero error term ν_{gri} is strictly exogenous and $E(\gamma_{ggr} - \gamma_{gg} \mid \widetilde{\ln p_{gri}}) = 0$ (Wooldridge,

[19]Results are unchanged if I use the national caloric share instead.

2005).[20] However the $tastes_{gr}$ are still biased as they include the regional $\gamma_{gg'}$ deviation.

Allowing elasticities to vary by region substantially reduces the degrees of freedom in estimating equation 4 and asks too much of the limited village price variation within regions. One possibility is to use the 38th and 50th thick rounds (1983 and 1993/4) to add extra price variation. I assume tastes are constant over the short-run (10 years in this case) and then estimate equation 4 on a region-by-region basis. Tables D.1 through D.6 show the main regressions rerun using this measure of tastes as a further robustness check. While allowing $\gamma_{gg'}$ to vary by region attenuates the coefficients, the signs remain the same and the coefficients are significantly different from zero.

It is also possible to make a similar argument to the one above, with regional variation in $\gamma_{gg'}$ only increasing the dispersion of tastes but not changing their rank ordering. In this case, I require $\frac{d\gamma_{gg'r}}{dtastes_{gr}} > 0$ and $\frac{d\gamma_{ggr}}{dtastes_{gr}} < 0$ for $\frac{d(\widehat{tastes}_{gr} - tastes_{gr})}{dtastes_{gr}} > 0$, so that the budget share spent on good g increases when there are price rises in other goods and decreases when the price of good g rises. If this is satisfied, by the same logic as I outlined above in the case of village-specific tastes, the dispersion of tastes will increase but not the ranking despite $\gamma_{gg'}$ varying by region, and therefore the normalized taste results shown in tables D.5 and D.6 should be valid.

D.2 Robustness of Regional Caloric Change Regression

There are several concerns regarding the parameters estimated by running regression 7. If households reduce non-food expenditure in response to rising prices for more favored foods, the caloric decline will be tempered. Table D.7 shows the results of rerunning regression 7, but replacing $\Delta \ln food_r$ with the change in total expenditure on all goods, $\Delta \ln expenditure_r$. The magnitude of the caloric reduction coming from tastes correlating with price changes declines by about half as expenditure is partially reallocated towards food. However, conditional upon total expenditure, caloric intake still declines with the correlation between tastes and price changes.

The approximation used to obtain regression 7 assumed that budget shares were fixed in the short run. This ignores any income effects that lower demand for inferior goods with low prices per calorie. The omission is likely to reduce the coefficient on $\Delta \ln food_r$. Since tastes are positively related to relative endowments, $\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$ will also be correlated with the size of the income gains from price changes and may bias b_1 downwards. The full log linearization includes an additional term, $\sum_{g=1}^{52} calshare_{gr} \Delta \ln bshare_{gr}$, which represents the decline in calories from a shift in budget shares to more expensive calorie sources:

[20] The last assumption is that price deviations within regions are uncorrelated with differences in $\gamma_{gg'}$ across regions. This is plausibly satisfied.

$$\Delta \ln calories \simeq \Delta \ln food - \sum_{g=1}^{52} [bshare_g] \left[\frac{food}{calories} \Big/ p_g \right] \Delta \ln p_g + \sum_{g=1}^{52} calshare_g \Delta \ln bshare_g,$$

$$\Delta \ln calories_r = b_0 + b_1 \sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr} + b_2 \sum_{g=1}^{52} h(P_r, food_r) \Delta \ln p_{gr} + b_3 \Delta \ln food_r$$
$$+ b_4 \overline{bshare_r} \sum_{g=1}^{52} (J_{gr} - \overline{J_r}) \Delta \ln p_{gr} + b_5 \sum_{g=1}^{52} calshare_{gr} \Delta \ln bshare_{gr} + \varepsilon_r,$$

where $calshare_g$ is the share of good g in total caloric intake. This regression is run and the results presented in columns 2 and 5 of table D.7. The coefficient on $\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$ actually becomes more negative, implying that the omission of the additional $\sum_{g=1}^{52} calshare_{gr} \Delta \ln bshare_{gr}$ term was not responsible for the negative coefficient on b_1.

As a further robustness check, I instrument for $\Delta \ln food_r$ with the log change in non-food expenditure, $\Delta \ln non\ food_r$. A shock that increases the demand for calories, such as changing work patterns, will also affect food expenditure and result in a positive correlation between $\Delta \ln food_r$ and the error term, biasing b_3 upwards. However, there will be a negative or no correlation with $\Delta \ln non\ food_r$, and the true value of b_3 will be bounded between the instrumented and uninstrumented estimates. These results are also shown in table D.7, and b_1 is essentially unchanged in the two specifications, implying that the endogeneity of food expenditure is not a major problem.

Figure D.1: The Predicted Caloric Losses from Liberalization and Regional Expenditure For Subsample Consuming Fewer than 2000 Calories Per Person Per Day, Excluding Farmers (Total Food Expenditure Held Constant, Markers Proportional to Population)

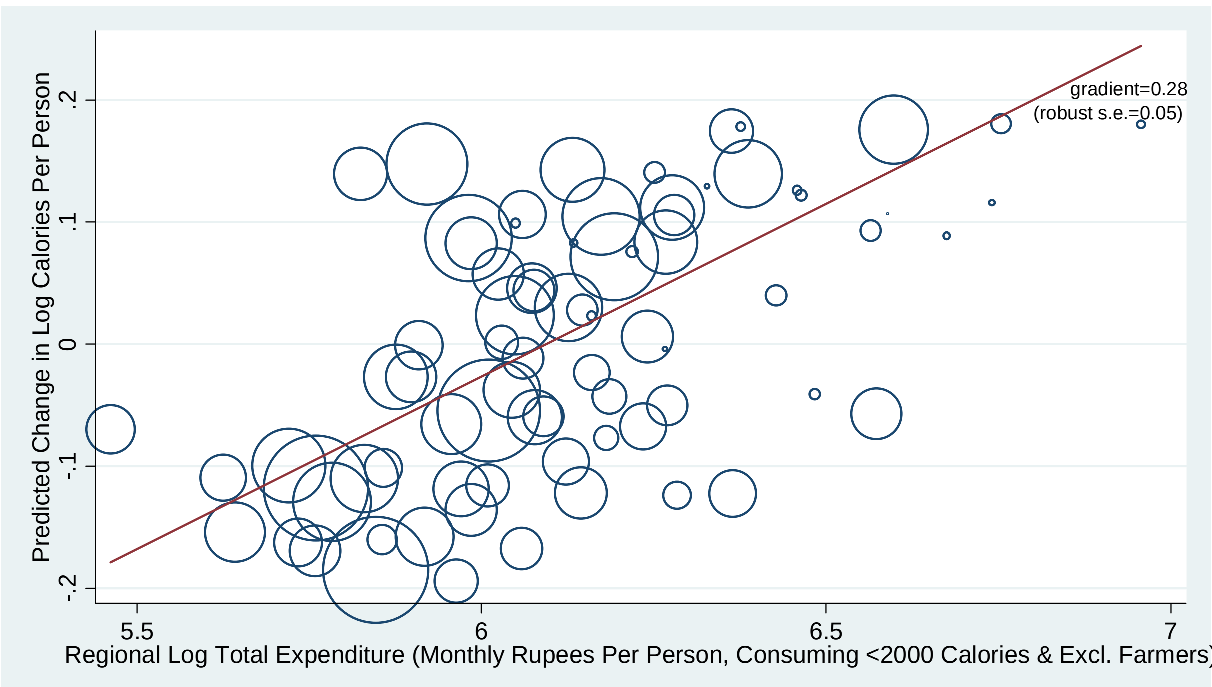


Table D.1: Robust Taste Estimates and Relative Resource Endowments I

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				$tastes_{gr}$			
	(Prod. Endow.)	(Value Endow.)	(Price Instr.)	($food_i$ Instr.)	(Regional $\gamma_{gg'r}$)	(Hh. Price Interact.)	(Wage & Density)
$endowment_{gr}$	1.764**	1.824**	1.891	1.751***	0.117**	1.736***	1.700***
	(0.71)	(0.90)	(1.17)	(0.51)	(0.058)	(0.52)	(0.65)
Observations	3000	1463	3375	3375	3278	3375	2385

Note: Dependent variable, tastes, estimated using the unexplained regional variation in food budget shares, with common price and food expenditure controls. Independent variable, $endowment_{gr}$, predicted from regressing observed relative endowments on agro-climatic endowments by crop using Limited Information Maximum Likelihood as in table 1. Production endowment (column 1) uses total production in tonnes rather than are planted for relative endowment in first stage. Value endowment (column 2) uses total value instead of area planted. Price instruments (column 3) for taste estimation are prices for 52 goods in nearby village. $food_i$ instrument (column 4) for taste estimation is non-food expenditure. Regional $\gamma_{gg'r}$ $tastes_{gr}$ (column 5) are estimated by running LA/AIDS separately on each region. The household price interactions (column 6) include interactions of all of the control variables with the own price term in the expenditure system. Agricultural wage and population density (column 7) includes two additional endowments, the real agricultural wage and the population density for the subset of regions covered by the India Agriculture and Climate Data Set. Both tastes and observed relative endowments normalized mean 0, s.d. 1 by good. Robust standard errors. * significant at 10 percent, ** 5, *** 1.

Table D.2: Robust Taste Estimates and Relative Resource Endowments II

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				$tastes_{gr}$			
Village Price:	Mean	Min	Max	25^{th} pctile.	75^{th} pctile.	Trans 1	Trans 2
$endowment_{gr}$	1.727***	1.698***	1.587***	1.897***	1.624***	1.721***	1.701***
	(0.55)	(0.47)	(0.47)	(0.67)	(0.51)	(0.56)	(0.51)
Observations	3375	3375	3375	3375	3375	3375	3375

Note: Dependent variable, tastes, estimated using the unexplained regional variation in food budget shares, with common price and food expenditure controls. Various village prices are used instead of the median price that is used in the main specification (mean, minimum, maximum, 25th percentile and 75th percentile of the reported unit values, a unit price including a 5 percent ad-valorem transport cost when good is not available locally and so nearby price used instead, trans1, and an ad-valorem transport cost based on sugar prices differences, trans 2). $endowment_{gr}$ are predicted values from regressing observed relative endowments on agro-climatic endowments by crop using Limited Information Maximum Likelihood as in table 1. Both tastes and observed relative endowments normalized mean 0, s.d. 1 by good. Robust standard errors. * significant at 10 percent, ** 5, *** 1.

Table D.3: Correlations Between Robust Taste Estimates and Prices I

Pearson's product-moment correlation			
	Prices$_{gr}$	[95 Percent Conf. Interval]	
$tastes_{gr}$ (Prices Instrumented)	-0.036**	-0.069	-0.004
$tastes_{gr}$ ($food_i$ Instrumented)	-0.084***	-0.116	-0.051
$tastes_{gr}$ (Regional $\gamma_{gg'r}$)	-0.027*	-0.059	0.006
$tastes_{gr}$ (Hhold Price Interactions)	-0.096***	-0.128	-0.064

Note: 3670 observations. Tastes estimated using the unexplained regional variation in food budget shares, with common price and food expenditure controls. Price instruments for taste estimation are prices for 52 goods in nearby village. $food_i$ instrument for taste estimation is non-food expenditure. Regional $\gamma_{gg'r}$ $tastes_{gr}$ are estimated by running LA/AIDS separately on each region. The household price interactions include interactions of all of the control variables with the own price term in the expenditure system. Prices are regional median unit values. Both variables normalized mean 0, s.d. 1 by good. * significant at 10 percent, ** 5, *** 1. Confidence intervals based on Fisher's transformation.

Table D.4: Correlations Between Robust Taste Estimates and Prices II

Pearson's product-moment correlation			
	Prices$_{gr}$	[95 Percent Conf. Interval]	
$tastes_{gr}$ (Mean Village Price)	-0.074***	-0.106	-0.041
$tastes_{gr}$ (Minimum Village Price)	-0.139***	-0.171	-0.108
$tastes_{gr}$ (Maximum Village Price)	-0.019	-0.052	0.013
$tastes_{gr}$ (25^{th} Pctile. Village Price)	-0.122***	-0.154	-0.090
$tastes_{gr}$ (75^{th} Pctile. Village Price)	-0.035**	-0.670	-0.003
$tastes_{gr}$ (5 Percent Transport Cost)	-0.054***	-0.086	-0.022
$tastes_{gr}$ (Sugar Price Transport Cost)	-0.075***	-0.107	-0.043

Note: 3670 observations. Tastes estimated using the unexplained regional variation in food budget shares, with common price and food expenditure controls. Various village prices are used instead of the median price that is used in the main specification (mean, minimum, maximum, 25th percentile and 75th percentile of the reported unit values, a unit price including a 5 percent ad-valorem transport cost when a good is not available locally and so a nearby price used instead, and a unit price including an ad-valorem transport cost based on sugar prices differences). Regional prices are regional median unit values. Both variables normalized mean 0, s.d. 1 by good. * significant at 10 percent, ** 5, *** 1. Confidence intervals based on Fisher's transformation.

Table D.5: Caloric Change, Robust Taste Estimates and Temporal Price Changes (Unweighted)

	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta \ln calories_r$ 1987-88 to 2004-05 (Unweighted)					
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$	-0.355***					
(Prices Instrumented)	(0.12)					
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$		-0.351***				
($food_i$ Instrumented)		(0.12)				
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$			-0.345***			
(Regional $\gamma_{gg'r}$)			(0.11)			
$corr_g(tastes_{gr}, \Delta \ln p_{gr})$				-0.118***		
(Normed Tastes)				(0.039)		
$\sum_{g=1}^{52} tastes_{gr} \widetilde{\Delta \ln p_{gr}}$					-0.201***	
(Demeaned $\Delta \ln p_{gr}$)					(-2.97)	
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$						-0.354***
(Hhold Price Interactions)						(0.11)
$\sum_{g=1}^{52} h_g(P_r, food_r) \Delta \ln p_{gr}$	-0.369***	-0.338**	-0.335***		-0.102	-0.292**
	(0.12)	(0.13)	(0.11)		(-1.09)	(0.14)
$\Delta \ln food_r$	0.589***	0.591***	0.594***	0.506***	0.526***	0.595***
	(0.048)	(0.048)	(0.048)	(0.052)	(10.6)	(0.047)
$\sum_{g=1}^{52} (J_{gr} - \overline{J_r}) \Delta \ln p_{gr}$	2.878	2.882*	2.806*	1.941	2.499	2.787*
	(1.57)	(1.58)	(1.61)	(1.47)	(1.66)	(1.60)
Constant	-0.369***	-0.381***	-0.389***	-0.677***	-0.673***	-0.387***
	(0.11)	(0.11)	(0.11)	(0.061)	(-11.5)	(0.11)
Observations	75	76	76	76	76	76
R^2	0.58	0.58	0.60	0.55	0.54	0.58

Note: Dependent variable is the log change in caloric intake per person between 1987-88 and 2004-05. The independent variables come from the log linearization of caloric intake. J_{gr} is the inverse relative price per calorie. Tastes estimated using the unexplained regional variation in food budget shares, with common price and food expenditure controls. Price instruments for taste estimation are prices for 52 goods in nearby village. $food_i$ instrument for taste estimation is non-food expenditure. Regional $\gamma_{gg'r}$ $tastes_{gr}$ are estimated by running LA/AIDS separately on each region. Normed tastes uses correlation between $\Delta \ln p_{gr}$ and tastes normalized mean 0 s.d. 1 by good, with correlation weighted using national food budget shares for each good. Demeaned price changes remove the region average price change from $\Delta \ln p_{gr}$. The household price interactions include interactions of all of the control variables with the own price term in the expenditure system. Robust standard errors. * significant at 10 percent, ** 5, *** 1.

Table D.6: Caloric Change, Robust Taste Estimates and Temporal Price Changes (Weighted)

	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta \ln calories_r$ 1987-88 to 2004-05 (Weighted)					
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$	-0.597***					
(Prices Instrumented)	(0.082)					
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$		-0.566***				
($food_i$ Instrumented)		(0.090)				
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$			-0.604***			
(Regional $\gamma_{gg'r}$)			(0.086)			
$corr_g(tastes_{gr}, \Delta \ln p_{gr})$				-0.0967***		
(Normed Tastes)				(0.033)		
$\sum_{g=1}^{52} tastes_{gr} \widetilde{\Delta \ln p_{gr}}$					-0.273***	
(Demeaned $\Delta \ln p_{gr}$)					(-3.58)	
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$						-0.600***
(Hhold Price Interactions)						(0.082)
$\sum_{g=1}^{52} h_g(P_r, food_r) \Delta \ln p_{gr}$	-0.655***	-0.522***	-0.600***		-0.101	-0.567***
	(0.094)	(0.11)	(0.088)		(-1.17)	(0.10)
$\Delta \ln food_r$	0.739***	0.724***	0.735***	0.511***	0.541***	0.732***
	(0.049)	(0.049)	(0.050)	(0.057)	(9.74)	(0.049)
$\sum_{g=1}^{52} (J_{gr} - \overline{J_r}) \Delta \ln p_{gr}$	1.533	1.911*	1.858	1.173	1.261	1.943*
	(1.17)	(1.09)	(1.20)	(1.64)	(0.68)	(1.16)
Constant	-0.248***	-0.290***	-0.255***	-0.665***	-0.679***	-0.259***
	(0.074)	(0.085)	(0.080)	(0.066)	(-10.3)	(0.078)
Observations	75	76	76	76	76	76
R^2	0.72	0.72	0.71	0.56	0.59	0.71

Note: Dependent variable is the log change in caloric intake per person between 1987-88 and 2004-05. The independent variables come from the log linearization of caloric intake. J_{gr} is the inverse relative price per calorie. Tastes estimated using the unexplained regional variation in food budget shares, with common price and food expenditure controls. Price instruments for taste estimation are prices for 52 goods in nearby village. $food_i$ instrument for taste estimation is non-food expenditure. Regional $\gamma_{gg'r}$ $tastes_{gr}$ are estimated by running LA/AIDS separately on each region. Normed tastes uses correlation between $\Delta \ln p_{gr}$ and tastes normalized mean 0 s.d. 1 by good, with correlation weighted using national food budget shares for each good. Demeaned price changes remove the region average price change from $\Delta \ln p_{gr}$. The household price interactions include interactions of all of the control variables with the own price term in the expenditure system. Regressions are weighted by a region's total survey weight. Robust standard errors. * significant at 10 percent, ** 5, *** 1.

Table D.7: Caloric Change, Tastes and Temporal Price Changes: Additional Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta \ln calories_r$ 1987-88 to 2004-05					
	(Unweighted)			(Weighted)		
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$	-0.146	-0.428***	-0.324**	-0.335***	-0.730***	-0.581***
	(0.14)	(0.12)	(0.14)	(0.11)	(0.093)	(0.11)
$\sum_{g=1}^{52} h_g(P_r, food_r) \Delta \ln p_{gr}$	-0.0693	-0.372**	-0.270*	-0.309**	-0.725***	-0.545***
	(0.17)	(0.14)	(0.16)	(0.13)	(0.12)	(0.12)
$\Delta \ln expenditure_r$	0.364***			0.445***		
	(0.062)			(0.055)		
$\Delta \ln food_r$		0.645***			0.806***	
		(0.052)			(0.048)	
$\Delta \ln food_r$ (Instrumented with $\Delta \ln nonfood_r$)			0.547***			0.710***
			(0.13)			(0.12)
$\sum_{g=1}^{52} (J_{gr} - \overline{J_r}) \Delta \ln p_{gr}$	2.649*	2.304	2.768*	2.154	1.685*	1.972*
	(1.38)	(1.55)	(1.51)	(1.87)	(1.00)	(1.12)
$\sum_{g=1}^{52} calshare_{gr} \Delta \ln bshare_{gr}$		0.159**			0.152***	
		(0.061)			(0.050)	
Constant	-0.427***	-0.314***	-0.363***	-0.315**	-0.158*	-0.255***
	(0.14)	(0.10)	(0.12)	(0.13)	(0.084)	(0.080)
Observations	76	76	76	76	76	76
R^2	0.37	0.63	0.58	0.45	0.76	0.71

Note: Dependent variable is the log change in caloric intake per person between 1987-88 and 2004-05. The independent variables come from the log linearization of caloric intake. J_{gr} is the inverse relative price per calorie. Tastes estimated using the unexplained regional variation in food budget shares, with common price and food expenditure controls. $\Delta \ln food_r$ instrumented by two stage least squares using $\Delta \ln nonfood_r$ in columns 3 and 6, with a first stage F-stat of 14.35 and 16.74 respectively. Regressions are weighted by a region's total survey weight where indicated. Robust standard errors. * significant at 10 percent, ** 5, *** 1.

Table D.8: Table of Means for Migrants and Non-Migrants

Variable	Full Sample		Wife Move Sample		Wife Move 2 Sample	
	Non-Mig	Migrant	Non-Mig	Migrant	Non-Mig	Migrant
Calories	2160.7	2228.8	2194.3	2222.8	2194.8	2257.7
(Per Person Per Day)	(4.2)	(34.0)	(5.6)	(18.3)	(5.8)	(22.1)
Food Expenditure	99.7	128.7	96.2	119.0	94.5	109.7
(Monthly Rupees/Person)	(0.3)	(1.8)	(0.4)	(1.5)	(0.4)	(1.7)
ln(Food Expenditure)	4.494	4.706	4.465	4.655	4.450	4.580
(Monthly Rupees/Person)	(0.003)	(0.010)	(0.003)	(0.013)	(0.003)	(0.014)
Total Expenditure	171.6	258.1	162.2	229.0	156.6	196.0
(Monthly Rupees/Person)	(0.7)	(4.9)	(1.0)	(7.8)	(0.8)	(3.8)
Years Since Moved			21.1	20.1	21.3	21.1
for Marriage			(0.1)	(0.3)	(0.1)	(0.3)
Age of Household Head	44.7	44.0	44.1	43.7	44.2	44.1
	(0.1)	(0.2	(0.1)	(0.3)	(0.1)	(0.4)
Illiterate	0.49	0.32	0.49	0.36	0.51	0.43
	(0.00)	(0.01)	(0.00	(0.01)	(0.00)	(0.01)
Above Primary	0.20	0.37	0.19	0.34	0.17	0.28
Education	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Household Size	6.37	6.17	6.64	6.64	6.68	6.89
	(0.02)	(0.06)	(0.02	(0.09)	(0.02)	(0.11)
Rural Household	0.80	0.47	0.86	0.57	0.88	0.71
	(0.00)	(0.01)	(0.00	(0.01)	(0.00)	(0.01)
Observations	115,069	9,898	49,431	3,868	44,945	2,565

Note: Means of non-migrant and migrant households from 1987-1988 NSS survey. For wife move samples, migrants are wives who moved inter-state, as opposed to intra-state, at the time of marriage separately from their husbands. Wife Move 2 sample only includes wives who moved to their husband's village of birth. Standard errors in parentheses. All means survey weighted.

Table D.9: Comparing Bundles of Migrants and Non-Migrants (Reporting Controls)

	(1)	(2)	(3)
		$\rho_{iods} = corr_g(bshare_{ig}, bshare_{\bar{s}g})$	
	Full Sample	Wife Move	Wife Move 2
$\mathbf{I}_{destination=s}$	-0.458***	-1.026***	-1.009***
	(0.085)	(0.10)	(0.11)
$\mathbf{I}_{destination=s,origin\neq s}$	-0.0451***	-0.0425***	-0.0316***
	(0.0040)	(0.0061)	(0.0071)
$\mathbf{I}_{destination\neq s}$	-0.786***	-1.382***	-1.367***
	(0.085)	(0.10)	(0.11)
$\mathbf{I}_{destination\neq s,origin=s}$	0.108***	0.125***	0.127***
	(0.0040)	(0.0063)	(0.0077)
$\mathbf{I}_{destination\neq s,nearby=s}$	0.181***	0.203***	0.203***
	(0.0014)	(0.0020)	(0.0021)
$\ln food$	0.493***	0.731***	0.725***
	(0.038)	(0.045)	(0.047)
$\ln food^2$	-0.0479***	-0.0733***	-0.0725***
	(0.0042)	(0.0048)	(0.0051)
age household head	0.000813**	0.00656***	0.00624***
	(0.00033)	(0.00067)	(0.00070)
age household head^2	-0.00000608*	-0.00000601	-0.00000409
	(0.0000034)	(0.0000054)	(0.0000056)
age spouse	-0.00354***	-0.000757	-0.000519
	(0.00081)	(0.0013)	(0.0014)
adult males		-0.00659***	-0.00642***
		(0.00041)	(0.00043)
adult females	0.00507***	0.00464***	0.00461***
	(0.00086)	(0.0013)	(0.0014)
children	-0.0000212	-0.000431	-0.000485
	(0.00051)	(0.00072)	(0.00075)
head literate ($\leq$ primary)	0.0377***	0.0372***	0.0372***
	(0.0021)	(0.0028)	(0.0029)
head > primary educ.	0.0317***	0.0391***	0.0416***
	(0.0024)	(0.0033)	(0.0035)
urban-urban mig.	-0.00856***	-0.0194***	-0.0261***
	(0.0033)	(0.0051)	(0.0057)
rural-urban mig.	-0.0181***	0.00600	-0.000791
	(0.0039)	(0.0049)	(0.0054)
urban-rural mig.	-0.0458***	-0.0355***	-0.0416***
	(0.0050)	(0.0057)	(0.0065)
Observations	3,864,925	1,637,916	1,472,531
R^2	0.77	0.76	0.75

Note: Dependent variable is the correlation between household food budget shares and mean shares for state s (31 observations per hhold). Independent variables are indicators for origin o and current d state. Constant, religion, caste, household type and subround dummies not shown. Robust standard errors. All regressions survey weighted and clustered further at individual household. * significant at 10 percent, ** 5, *** 1.

Table D.10: Caloric Intake of Migrants Compared to Non-Migrants (Reporting Controls)

	(1)	(2)	(3)
	Daily Calories Per Person $calories_i$ (Full Sample)		
$migrant_i$	-107.2***	-92.23***	38.66
	(18.2)	(32.9)	(33.6)
$\ln food$	-2777***		
	(963)		
$\ln food^2$	478.1***		
	(115)		
$\ln total\ expenditure$		284.6	
		(236)	
$\ln total\ expenditure\ ^2$		67.34***	
		(23.8)	
age household head	9.707***	18.30***	22.51***
	(0.94)	(1.23)	(1.40)
age household head^2	-0.0957***	-0.156***	-0.169***
	(0.011)	(0.012)	(0.014)
adult males	10.02***	5.510**	19.44***
	(2.10)	(2.60)	(3.16)
adult females	1.072	-0.398	-19.06***
	(4.26)	(2.78)	(3.43)
children	-6.115	-35.55***	-106.1***
	(5.22)	(1.73)	(2.16)
head literate ($\leq$ primary)	-69.61***	-37.71***	110.2***
	(5.29)	(8.26)	(9.59)
head $>$ primary educ.	-261.3***	-153.8***	257.4***
	(35.8)	(10.2)	(9.21)
urban-urban mig.	-440.6***	-400.7***	-170.2***
	(26.2)	(12.2)	(12.5)
rural-urban mig.	-290.3***	-276.8***	-114.0***
	(10.4)	(16.5)	(17.7)
urban-rural mig.	-39.81***	-41.61***	8.542
	(11.9)	(13.9)	(16.3)
Observations	124,578	124,578	124,578
R^2	0.50	0.29	0.10

Note: Daily calories per person regressed on an inter-state migrant dummy that takes the value 1 if either the household head or his spouse migrated from another state. Constant, religion, caste, household type, subround and origin-state dummies not shown. Robust standard errors. All regressions survey weighted. * significant at 10 percent, ** 5, *** 1.

Table D.11: Caloric Intake of Intra-State and Inter-State Wife Households (Reporting Controls)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Daily Calories Per Person $calories_i$							
	Wife Move Sample				Wife Move 2 Sample			
$migrant_i$	-115.0***	-99.16***	21.91	-257.3***	-43.87**	-43.02**	35.72	-140.8***
	(17.0)	(17.1)	(18.7)	(33.05)	(18.2)	(18.7)	(22.3)	(37.82)
$\ln food$	-3787***			-3794***	-3959***			-3957***
	(588)			(588.0)	(630)			(630.0)
$\ln food^2$	585.8***			586.7***	609.6***			609.4***
	(67.1)			(67.13)	(72.1)			(72.19)
$\ln total$		-783.6				-1676***		
$expenditure$		(489)				(545)		
$\ln total$		179.2***				272.1***		
$expenditure^2$		(49.5)				(55.5)		
$yrsaway_i$				5.529***				4.478***
				(0.644)				(0.672)
$migrant_i$				2.668**				1.768
$\times yrsaway_i$				(1.143)				(1.208)
mother-in-law				66.25***				64.20***
				(8.420)				(8.636)
$migrant_i \times$				79.57**				56.36
mother-in-law				(33.77)				(36.43)
children	-12.10***	-33.88***	-102.0***	-13.59***	-12.06***	-33.03***	-100.0***	-12.70***
	(2.02)	(2.12)	(2.63)	(2.000)	(2.08)	(2.20)	(2.72)	(2.089)
$migrant_i \times$				27.55***				17.39**
children				(7.992)				(8.606)
age head	9.576***	14.06***	21.09***	9.952***	9.862***	14.01***	20.91***	10.34***
	(1.37)	(1.62)	(2.09)	(1.364)	(1.43)	(1.69)	(2.17)	(1.427)
age head^2	-0.114***	-0.148***	-0.179***	-0.124***	-0.115***	-0.145***	-0.175***	-0.124***
	(0.013)	(0.016)	(0.021)	(0.0130)	(0.014)	(0.016)	(0.021)	(0.0135)
age spouse	2.652***	3.845***	3.022***	-1.102	2.425***	3.217***	2.555**	-0.526
	(0.69)	(0.86)	(1.12)	(0.823)	(0.73)	(0.89)	(1.17)	(0.858)
adult males	7.021**	7.848**	24.31***	7.959***	4.504	6.224	25.32***	5.533*
	(3.08)	(3.86)	(4.89)	(3.078)	(3.18)	(4.02)	(5.08)	(3.183)
adult females	2.833	-2.653	-9.414*	-11.99***	0.157	-3.579	-8.098	-13.68***
	(3.17)	(3.99)	(5.08)	(3.615)	(3.28)	(4.10)	(5.32)	(3.751)
literate $\leq$	-66.96***	-50.05***	99.15***	-68.59***	-70.40***	-51.32***	98.78***	-72.03***
primary	(6.41)	(7.89)	(9.95)	(6.372)	(6.71)	(8.15)	(10.4)	(6.664)
$>$ primary	-215.1***	-145.7***	248.1***	-215.7***	-210.2***	-148.8***	237.3***	-211.6***
educ.	(12.2)	(12.5)	(12.5)	(12.18)	(12.6)	(13.2)	(13.3)	(12.70)
urban-urban	-486.8***	-455.1***	-177.0***	-476.2***	-456.5***	-430.0***	-198.9***	-450.1***
mig.	(16.6)	(18.5)	(17.7)	(16.45)	(17.1)	(18.8)	(19.3)	(17.07)
rural-urban	-287.9***	-264.0***	-112.1***	-281.8***	-241.3***	-219.3***	-109.2***	-238.2***
mig.	(11.5)	(13.5)	(15.2)	(11.48)	(12.1)	(14.5)	(16.4)	(12.10)
urban-rural	-61.85***	-51.77***	7.354	-58.63***	-79.22***	-55.41***	-1.979	-75.60***
mig.	(12.4)	(15.3)	(19.0)	(12.32)	(13.7)	(16.7)	(21.9)	(13.57)
Observations	52,836	52,836	52,836	52,800	47,501	47,501	47,501	47,465
R^2	0.66	0.45	0.14	0.66	0.67	0.47	0.14	0.67

Note: Daily calories per person regressed on an inter-state wife dummy, $migrant_i$, and the years since moving with an interaction for being an inter-state wife. Wife Move 2 sample only includes wives who moved to their husband's village of birth. Constant, religion, caste, household type, subround and origin-state dummies not shown. Robust standard errors. All regressions survey weighted. * significant at 10 percent, ** 5, *** 1.

Table D.12: Caloric Intake of Migrants Compared to Non-Migrants (Destination-State-Sector Controls)

	(1)	(2)	(3)			
	Daily Calories Per Person $calories_i$					
	Destination-State-Sector Dummies Only			Also Including Origin-State Dummies		
	Full Sample	Wife Move	Wife Move 2	Full Sample	Wife Move	Wife Move 2
$migrant_i$	-69.90***	-84.40***	-31.42*	-70.90***	-80.20***	-28.11
	(13.99)	(15.42)	(17.08)	(16.76)	(15.81)	(17.46)
ln $food$	-2780***	-3755***	-3966***	-2788***	-3769***	-3968***
	(960.2)	(585.0)	(630.3)	(962.0)	(585.3)	(630.5)
ln $food^2$	479.1***	582.3***	610.7***	480.2***	584.2***	611.1***
	(114.3)	(66.78)	(72.24)	(114.6)	(66.82)	(72.26)
age household head	9.737***	9.892***	10.06***	9.711***	9.880***	10.07***
	(0.932)	(1.366)	(1.432)	(0.930)	(1.363)	(1.430)
age household head^2	-0.0971***	-0.112***	-0.114***	-0.0969***	-0.113***	-0.115***
	(0.0110)	(0.0130)	(0.0135)	(0.0109)	(0.0129)	(0.0135)
age spouse		2.033***	2.111***		2.065***	2.096***
		(0.688)	(0.720)		(0.687)	(0.719)
adult males	10.95***	7.553**	4.387	11.03***	7.695**	4.443
	(2.061)	(3.068)	(3.179)	(2.060)	(3.062)	(3.182)
adult females	-0.941	2.381	0.947	-0.639	2.558	1.006
	(3.943)	(3.128)	(3.260)	(4.069)	(3.128)	(3.261)
children	-6.262	-11.72***	-11.85***	-6.337	-11.80***	-11.84***
	(5.464)	(2.028)	(2.085)	(5.422)	(2.022)	(2.085)
head literate ($\leq$ primary)	-68.99***	-65.24***	-70.07***	-68.70***	-65.26***	-69.94***
	(5.676)	(6.349)	(6.675)	(5.715)	(6.354)	(6.682)
head > primary educ.	-268.1***	-224.9***	-218.7***	-267.3***	-224.2***	-217.9***
	(36.80)	(12.29)	(12.78)	(36.52)	(12.27)	(12.78)
rural household	-113.5	753.8***	-899.1	-8.194	1336***	-974.1
	(92.23)	(22.59)	(720.7)	(139.6)	(155.6)	(736.3)
Observations	124967	53335	47547	124967	53335	47547
R^2	0.500	0.662	0.670	0.501	0.663	0.671

Note: Daily calories per person regressed on an inter-state migrant dummy. Constant, religion, caste, household type, subround, origin-state and destination-state-sector dummies not shown. Sector is either rural or urban. Robust standard errors. All regressions survey weighted. * significant at 10 percent, ** 5, *** 1.

Table D.13: Caloric Change and the Correlation of Tastes with Spatial Price Changes (Reporting Controls)

	(1)	(2)	(3)
		Daily Calories Per Person $calories_i$	
	Full Sample	Wife Move Sample	Wife Move 2 Sample
$\sum_{g=1}^{52} tastes_{go} \Delta \ln p_{god}$	-0.873***	-0.705***	-0.560***
$\times migrant_i$	(0.0598)	(0.0672)	(0.0841)
$\sum_{g=1}^{52} h_{go}(P_o, food_o) \Delta \ln p_{god}$	-0.864***	-0.759***	-0.733***
$\times migrant_i$	(0.0713)	(0.0946)	(0.127)
$\sum_{g=1}^{52} (J_{go} - \overline{J_o}) \Delta \ln p_{god}$	-2.924*	-6.597***	-9.597***
$\times migrant_i$	(1.765)	(2.075)	(2.598)
$migrant_i$	-0.0209***	-0.00127	0.0200**
	(0.00703)	(0.00697)	(0.00820)
$\ln food$	1.692***	1.242***	1.257***
	(0.0746)	(0.105)	(0.111)
$\ln food^2$	-0.111***	-0.0665***	-0.0677***
	(0.00814)	(0.0113)	(0.0120)
age household head	0.00668***	0.00514***	0.00527***
	(0.000409)	(0.000575)	(0.000593)
age household head^2	-6.32e-05***	-5.75e-05***	-5.70e-05***
	(4.13e-06)	(5.16e-06)	(5.33e-06)
age spouse		0.00114***	0.000875***
		(0.000300)	(0.000304)
adult males	0.00322***	0.00285**	0.00257*
	(0.000949)	(0.00132)	(0.00138)
adult females	0.000893	0.000526	-0.000688
	(0.00101)	(0.00135)	(0.00140)
children	-0.00376***	-0.00676***	-0.00670***
	(0.000596)	(0.000752)	(0.000770)
head literate (≤ primary)	-0.0386***	-0.0341***	-0.0341***
	(0.00216)	(0.00285)	(0.00293)
head > primary educ.	-0.0877***	-0.0735***	-0.0720***
	(0.00296)	(0.00352)	(0.00369)
urban-urban mig.	-0.183***	-0.188***	-0.183***
	(0.00415)	(0.00558)	(0.00621)
rural-urban mig.	-0.137***	-0.122***	-0.105***
	(0.00452)	(0.00487)	(0.00519)
urban-rural mig.	-0.0265***	-0.0258***	-0.0381***
	(0.00498)	(0.00550)	(0.00607)
Observations	124578	52812	47482
Origin-Destination Pairs	484	341	256
R^2	0.722	0.695	0.699

Note: Dependent variable is the log of caloric intake per person. Tastes estimated using unexplained state variation in food budget shares, with common price and expenditure controls (LA/AIDS). J_{gr} is the inverse relative price per calorie. Constant, religion, caste, household type, subround and origin-state dummies not shown. Robust standard errors. All regressions survey weighted. * significant at 10 percent, ** 5, *** 1.

Table D.14: Caloric Change and the Correlation Between Tastes and Temporal Price Changes for Richer and Poorer Regions

	(1)	(2)
	$\Delta \ln calories_r$ 1987-88 to 2004-05	
	Richer Half of Sample	Poorer Half of Sample
$\sum_{g=1}^{52} tastes_{gr} \Delta \ln p_{gr}$	-0.158	-0.645***
	(0.13)	(0.12)
$\sum_{g=1}^{52} h_g(P_r, food_r) \Delta \ln p_{gr}$	-0.109	-0.633***
	(0.13)	(0.17)
$\Delta \ln food_r$	0.720***	0.691***
	(0.089)	(0.067)
$\sum_{g=1}^{52} (J_{gr} - \overline{J_r}) \Delta \ln p_{gr}$	3.462**	1.042
	(1.32)	(2.18)
Constant	-0.760***	-0.152
	(0.13)	(0.095)
Observations	38	37
R^2	0.76	0.75

Note: Dependent variable is the log change in caloric intake per person between 1987-88 and 2004-05. The independent variables come from the log linearization of caloric intake. J_{gr} is the inverse relative price per calorie. Regressions are weighted by a region's total survey weight. The sample is split by a region's average monthly per-capita expenditure in 1987/88. Robust standard errors. * significant at 10 percent, ** 5, *** 1.

Appendix References

Athukorala, Prema-chandra, "Agricultural Trade Policy Reforms in India," *South Asia Economic Journal*, 2005, *6* (1), 23–26.

Das-Gupta, Arindam, "Internal Trade Barriers in India: Fiscal Check-posts," *South Asia Economic Journal*, 2006, *7* (2), 231.

Dhar, Tirtha, Jean-Paul Chavas, and Brian W. Gould, "An Empirical Assessment of Endogeneity Issues in Demand Analysis for Differentiated Products," *American Journal of Agricultural Economics*, 2003, *85* (3), 605–617.

FAO, "WTO Agreement on Agriculture: The Implementation Experience - Developing Country Case Studies," Report, Food and Agriculture Organization 2003.

FAO, "Towards An Indian Common Market: Removal of Restrictions on Internal Trade in Agriculture Commodities," Report, Food and Agriculture Organization 2005.

Gulati, Ashok, "Indian Agriculture in an Open Economy: Will it Prosper?," *India's Economic Reforms and Development: Essays for Manmohan Singh*, 1998, pp. 122–46.

Hausman, Jerry A., "Valuation of New Goods under Perfect and Imperfect Competition," National Bureau of Economic Research Working Paper 4970 1994.

Jha, Raghbendra, K. V. Bhanu Murthy, and Anurag Sharma, "Market Integration in Wholesale Rice Markets in India," ASARC Working Papers 2005-03 2005.

Kennan, John, "Simultaneous Equations Bias in Disaggregated Econometric Models," *The Review of Economic Studies*, 1989, *56* (1), 151–156.

Melvin, James and Robert Waschik, "The Neoclassical Ambiguity in the Specific Factor Model," *Journal of International Trade & Economic Development*, 2001, *10* (3), 321–337.

Planning Commisson of India, "Internal Trade," in "Tenth Five Year plan (2002–2007)," New Delhi, 2001, chapter 7.8.

Sanghi, Apurva, K. S. Kavi Kumar, and James W. McKinsey, "India Agriculture and Climate Data Set," 1998. http://chd.ucla.edu/dev_data.

Stock, James H. and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," National Bureau of Economic Research Working Paper T0284 2002.

Vanneman, Reeve and Douglas Barnes, "Indian District Data, 1961-1991: Machine-readable Data File and Codebook," 2000. http://www.inform.umd.edu/districts/index.html.

Willmott, Cort J. and Kenji Matsuura, "Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series 1950–1999," 2001. http://climate.geog.udel.edu/~climate.

Wooldridge, Jeffrey M., "Fixed-Effects and Related Estimators for Correlated Random-Coefficient and Treatment-Effect Panel Data Models," *Review of Economics and Statistics*, 2005, *87* (2), 385–390.