Norwegian Institute
of International
Affairs

Food Price Differences Across Indian States: Patterns and Determinants

Arne Melchior

NUPI Working Paper 869

Publisher: Norwegian Institute of International Affairs

Visiting address: C.J. Hambros plass 2d
Address: P.O. Box 8159 Dep.
NO-0033 Oslo, Norway
Internet: www.nupi.no
E-mail: post@nupi.no
Fax: [+ 47] 22 99 40 50
Tel: [+ 47] 22 99 40 00

Food Price Differences Across Indian States: Patterns and Determinants

Arne Melchior
Norwegian Institute of International Affairs (NUPI), Oslo, Norway
Oslo, 16 November 2016

Abstract
The paper examines food price differences across Indian states during 2004-2014 using food consumer prices from household surveys and wholesale/retail prices for selected goods. At the individual product level there are large price differences across states, with prices doubling or trebling across India for a typical case, but with considerable variation across products. Price dispersion is still high but considerably lower for food on average; measured at this level price dispersion between Indian states is considerably lower than between countries within the same income range, and Indian states are slightly more integrated than countries in Western Europe. At the product level, the most important determinants of price differences across states are limited access to supply from other states, and the extent of own production in the state. Richer states have higher consumer prices, but this income-price link is weaker for wholesale prices. Food price dispersion within India has decreased during the period studied. For policy, the results suggest that India should eliminate obstacles to inter-state trade in order to promote food security and the real income of its citizens. The magnitude and importance of price level differences also suggest that better price level data should be provided in the future, to facilitate further study of India's regional development.

JEL Codes: F11, F15, O13, Q11, R12.

Published by Norwegian Institute of International Affairs

Contents

1. Introduction[1]

Food prices and food price inflation play a key role for food safety, welfare and nutrition in all countries but particularly in developing countries, due to the larger share of food in overall consumption. In India, food constitutes about 40 per cent of urban consumer expenditure and about 50% of rural consumption, with even higher shares for the lowest income classes (NSSO 2013).[2] Food price inflation in India has recently (from 2008-09) accelerated, giving rise to concern that this development could hurt the poor and a debate about causes and policy measures to contain price inflation (see e.g. Gulati and Saini 2013).

Food price inflation in India is often debated as a national issue and a key issue often neglected is the variation in food prices across states. In this article, we demonstrate that there is large inter-state variation in food prices in India; for a typical food item the highest price across Indian states may be 2-4 times higher than the lowest one. With so large price gaps, national averages provide an inaccurate picture of food issues at the state level and this is particularly important for a huge country with considerable heterogeneity across regions. We present new evidence and analyse the determinants of this price variation, and discuss the implications.

Inter-state price variation is important not only for food markets and consumption, but also for the macroeconomic analysis of India's development. Without explicitly measuring price level gaps, regional developments cannot be fully understood, especially for emerging countries with fast change and large regional gaps. National statistical agencies in some countries (e.g. India and China) actually do collect price data and construct price indexes at the regional level, but standard practice is to

---

1 The paper is written as part of the research project FOODSEC (Food Security in India: The Interactions of Climate Change, Economics, Politics and Trade), Project No. 226057 funded by the Research Council of Norway (RCN) during 2013-2016. Funding from RCN is gratefully acknowledged. I thank Prof. P. G. Chengappa for facilitating and Dr. M. Umanath for assistance in retrieving wholesale price data. I thank A. Ganesh Kumar, Parmod Kumar, Pratap Singh Birtal and Fenella Carpena for highly useful comments to an earlier draft. I also thank the participants at the FOODSEC seminar at ISEC (Institute for Social and Economic Change), Bangalore on 19 January 2016 for comments when preliminary results were presented, and participants at the FOODSEC workshop in Delhi 4 Ocober 2016 for comments when an earlier version of the paper was presented. As usual, the responsibility for remaining errors stays with the author.

2 The exact numbers depend on the method of measurement, and we therefore refer to the approximate levels without going into detail. For more information, see NSSO (2013).

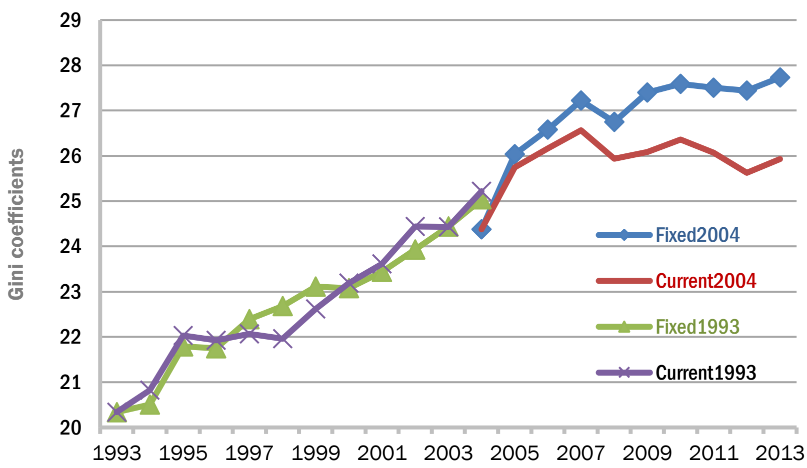
set price levels in all regions equal to 100 in the statistical base year. Thereafter we can follow the evolution of price level differences for a while; but when a new base year is introduced and we start all over again at 100.

As a result of such procedures, the measurement of regional disparities depends on the base year, and results based on fixed prices differ from those based on current prices. As an illustration, Diagram 1 shows regional income inequality in India from 1993-94 to 2013-2014, using GDSP (gross domestic state product) per capita in current and fixed prices and with 1993-94 as the base year for the first period, and 2004-05 as the base year for the latter part. We use population-weighted Gini coefficients (0=no inequality; 1=maximum inequality).

Diagram 1: Regional inequality in India

Population-weighted GINI coefficients across states.

Data source: Government of India, MOSPI and NITI Aayog.



Regional inequality increased until about 2007 and thereafter remained more stable. For the period from 1993-94 until 2004-05, the trend was similar with current and fixed prices. From 2004-05 to 2013-14, however, the two curves diverge, with slightly rising inequality based on fixed prices but falling inequality measured by current prices. Here we should believe more in the fixed price measurement, and the figure suggests that there was faster price growth in relatively poor states that "undermined" their nominal income gains. Observe also how the curves diverge in the overlapping year 2004-05, due to the change of basis for GDP price measurement. Hence with the new base year, inter-state inequality drops significantly due to the new basis (and product classifications).

Figure 1 suggests that India could have a similar debate as the one that raged about international income disparities some time ago: During the 1990's, population-weighted Gini coefficients for international income inequality across countries showed a considerable decline measured in purchasing power parities (PPPs, adjusted for price level differences across countries), but an increased based on current exchange rates (Milanovic 2005, Melchior and Telle 2001). Hence measuring price level differences is a key for understanding international as well as regional (intra-national) income gaps in large countries.

This study of inter-state price variation in India is therefore motivated partly by interest in issues about food consumption and food safety, but also aims to shed light on inter-state price variation more generally. The idea is that such price variation may be of paramount importance for large countries such as the USA, India, China, Russia and Brazil – with considerable heterogeneity and large geographical spaces. There has been considerable research on intra-national price gaps in the USA, or for the USA and Canada focusing in the difference between intra-national effects and the border effect (see e.g. O'Connell and Wei 1997, Gorodnichenko and Tesar 2009, Hajzler and MacGee 2014). There is now also a considerable literature on regional price differences within China; see e.g. Brandt and Holz (2006), Gong and Meng (2008) and Li and Gibson (2014). According to Li and Gibson (2014, 100), there has been price level convergence so that China is by now a "relatively well integrated market economy" – contrary to some earlier literature suggesting that China was a fragmented market with large price dispersion. Also for Russia, Gluschenko and Khimich (2007) show intra-national convergence for food prices; although some regions (notably Russia Far East) appear to be segmented from the rest of the country. Intra-national price gaps could be more important for emerging economies with considerable regional heterogeneity; while e.g. the USA has large inter-personal income inequality, its regional inequality is limited (Melchior 2008). Hence inter-state price variation should be a BRIC issue, due to country size combined with regional heterogeneity.[3]

For India, the literature on regional price differences is limited. Deaton and Dupriez (2011) use household survey data to examine rural-urban and spatial price differences in India and Brazil. For India, they use household survey data for 2004/2005 and calculate Törnquist price indexes for 21 larger states. Their results indicate that the food price gap across states from top to bottom is about 20%, and that approximately 1/3 of the price gaps may be due to quality differences. For India, regional price gaps are very important for measuring poverty (see e.g. Deaton 2003), and this is a motivating factor for Deaton and Dupriez (2011). The authors also show that regional price gaps in Brazil are much smaller than for India.

3 Observe that while the USA has large income inequality across persons, it has low income inequality across states. Hence the USA is likely more integrated than the BRICs, potentially also with respect to price level differences (Melchior 2008).

In this paper, we add to the analysis of regional price differences in India in different ways. Similar to Deaton and Dupriez, we use unit values for household surveys in parts of the analysis, but covering three years (also 2009/10 and 2011/12). Based on this, we present new descriptive evidence on the extent of food price variation across states and over time. Second, we extend the analysis for a subset of food products where we also collect wholesale and retail data. This provides a double check on the unit values from household survey data, and new evidence on price gaps at different levels of the value chain. Third, we examine the determinants of inter-state price variation, adding data on state-level characteristics as well as the spatial pattern of production for the different products. A main contribution of the paper is to add trade and the supply side into the analysis of food prices in India. An innovation is to analyse the impact of trade with no data on inter-state trade; using a gravity-based index of proximity to supply that turns out to have strong explanatory power. Parts two and three of the analysis are selective, focusing on some products and not all, and the aim is not to derive the "true food price index" for India, but to find out more about the drivers of regional price variation.

Why do price levels differ across regions or countries? Here we draw on the growing body of research on international price differences. To some extent, this has been stimulated by the huge data collection project undertaken by the World Bank in the International Comparison Project (ICP), where price data from an increasing number of countries has been collected at regular time intervals since the 1960s (see icp.worldbank.org). The latest round of data collection was in 2011 and it is interesting to observe that for food products, India had the lowest price level among 177 countries in the world, with a value of 50, with the world average at 100 and the highest prices at 232 (Norway and Japan). The themes addressed in research on international price differences are generally relevant also for the study of intra-national price differences in India.

A theoretical workhorse in the literature on international price differences has been the so-called Balassa-Samuelson hypothesis; saying that prices for internationally traded goods are equalized but international price differences are caused by price level differences for non-traded goods and services (Balassa 1964, Samuelson 1964). According to trade-based explanations, price levels for traded goods and services should generally be equalized if trade costs are eliminated. A paradox in this perspective is that price level differences are persistent also for traded goods and services, even in cases where formal trade barriers have been eliminated. A key finding in the literature is that even for traded goods, there is internationally a strong correspondence between income and price levels (see e.g. Hallak 2006, Bekker et al. 2012), and this is not easily explained in the trade-theoretical approach or with the Balassa-Samuelson hypothesis. International trade data provide a rich source of information that has been exploited in recent research to shed light on the source of price differences. This literature suggests that apart from trade costs, price differences can be caused by quality differences or particular demand patterns (see e.g. Hallak and Schott 2011, Feenstra and Romalis 2014), or by pricing-to-market behaviour (see e.g. Alessandri and Kaboski 2009, Simonovska 2010). Similar lines of research have also been followed and shown to be of importance for food prices in India (Deaton and Dupriez 2011, Atkin 2013). An alternative explanation for the price-income relationship is that traded goods and services are made in value chains that include a services component also in the consuming nation or state, or along the way through intermediaries within or between countries. According to this, a product also includes a bundle of related services, e.g. for food products transport, storage, cooling and intermediation. The recent literature on global value chains (see e.g. Timmer et al. 2014) sheds light on this at the international level, and the theme may be relevant also in the analysis of India. In the analysis of the determinants of inter-state price gaps, we revert to some of these issues.

While international price differences can be examined using detailed trade data, such data is not available for inter-state trade in India and we therefore rely on price data combined with data on production and state-level characteristics. The article's focus is motivated by the excellent supply of Indian price data, especially for food items. We use several different data sources:

- In section 2 we use data from the NSSO (National Sample Survey Organization) household surveys to shed light on inter-state price gaps for all food products. We use data from the surveys in 2004-05, 2009-10 and 2011-12 for per capita consumption of individual products at the state level, and analyse dispersion in the unit values for each product across states. Using population-weighted Gini coefficients like the ones shown in Diagram 1 (with unit values instead of income levels), we show that for food products there is large and persistent inter-state variation in India, but with some decline over time.

- In section 3 we analyse price dispersion between all the major mandis (wholesale markets) for 15 important food items, based on data from the National Horticulture Board (NHB). We show that retail prices are on average about 50-60 per cent higher than wholesale prices, and the NHB retail prices closely correspond to the consumer prices derived from the NSSO household surveys. For the products included, the extent of inter-state price dispersion is as large as for consumer prices.

- In section 4 we analyse the determinants of inter-state price variation, using the price data referred to above combined with other data for states or markets (production, income levels, etc.). We show that the standard trade cost explanation is the strongest driver of price differences, since states with larger production or shorter distance to main suppliers, or with better roads, have lower price levels. Along with the evidence for international price dispersion, we also find that richer states have higher price levels. This income-price link is important for consumer prices, but not for wholesale prices. We do not have data to show whether value chains, quality differences or pricing-to-market is explaining the price-income correlation.

Section 5 finally discusses the implications of the analysis. The analysis suggests that better infrastructure and access to supply is important in order to lower food prices, so India should promote better infrastructure and reduce all obstacles to trade across states in order to better food security and the real income of its citizens.

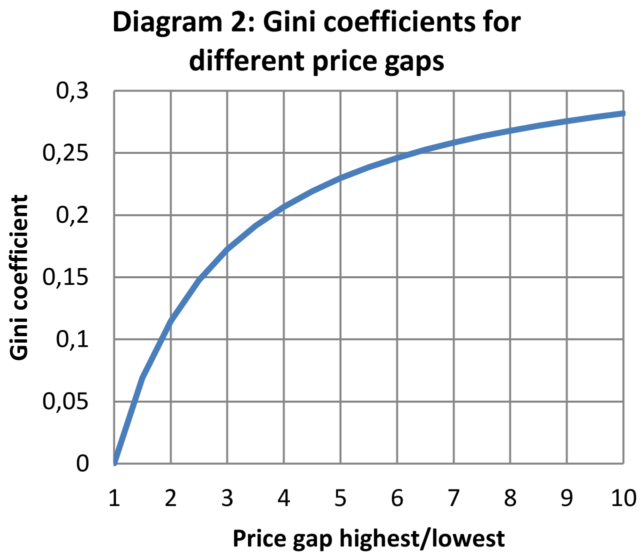
The correspondence between prices and income levels suggests that with development and growth, India should expect rising food prices as part of its development, so in the future the country cannot expect to maintain the lowest food price level in the world. We argue that the magnitude of price gaps warrants better statistics in the field, particularly for emerging large nations such as the BRICs. For this reason, statistical agencies should provide regular data on price levels and not only price changes over time.

2. Food price dispersion across Indian states: evidence from household survey data

India's household surveys provide a rich data source related to food consumption and food prices. Not only are the surveys extensive with a large number of respondents across all states; data is collected on an extensive list of issues and published online by the NSSO (National Sample Survey Organization) in a large number of reports including a lot of detailed material even if one does not have access to the raw data. In the following, we present evidence from three surveys where detailed data on food consumption was also collected, namely the 61st, 66th and 68th household surveys, undertaken by NSSO in 2004-05, 2009-10 and 2011-12, respectively. The results are documented in a number of reports, of which we mainly use here NSSO 2007, 2012, 2013, 2014a, 2014b. In the surveys, data is collected for about 125 different food products (in addition to non-food products), reporting the quantity and value of consumption per capita in each of 35 states and union territories. There are some classification changes over time but there are 113 items where the description is the same in all three years and we use only these when we compare results across the three years.

In the survey, consumption of non-food items is mostly based on collected data on *expenditures*. For food and fuel, however, the consumed *quantity* is the main variable and the value of food consumption is not directly observed but computed based on different relevant prices: "Consumption out of purchase is evaluated at the purchase price. Consumption out of home produce is evaluated at ex farm or ex factory rate. Value of consumption out of gifts, loans, free collections, and goods received in exchange of goods and services is imputed at the rate of average local retail prices prevailing during the reference period." (NSSO 2014a, 8). Hence the valuation of food prices is a "composite" based on different prices. Nevertheless, the unit value of products, obtained by dividing the value of expenditure by its quantity, provides a useful indicator for analysing price variation across states. Later, we shall compare these prices to other price indicators.

NSSO (2007, 2012, 2014b) reports results for 35 states/territories, split into rural and urban areas. We are interested in inter-state variation rather than urban-rural price gaps so we merge urban and rural areas and use average unit values by state. For the majority of products, there is consumption in most of the 35 states. Since states vary greatly in size and importance, there is a need for weighting. This is a main motivation why we use *population-weighted Gini coefficients* as our measure of price dispersion across states. An advantage is also that the Ginis provide a comparable and easy-to-recall measure of price spread across products and settings.[4] The Gini coefficient varies between 0 (prices equal in all states) and has an upper limit of 1 (can be thought of as zero prices in all states except one). Given our use of Gini coefficients, it may be useful to illustrate what a Gini at a certain level tells about price gaps. For this purpose, think of 30 states with equal populations. We assume that one of these has a price of one, and this is the lowest price observed. We let the highest price observed vary and assume that the price increases uniformly across states, between the lowest and highest value. Calculating the Ginis for price gaps between 1:1 (no price differences at all) and 10:1 (the highest price is 10 times the lowest), we obtain the Gini coefficients shown in Diagram 2.



A top/bottom price gap at 2:1 produces a Gini at 0.11, and a gap of 3:1 gives a Gini of 0.17. This is in fact the most typical range observed in the data for India. In the real data, states differ in size and some smaller states sometimes have more extreme prices but we use population-weighted Ginis so that such outliers have less influence. This may be compared to the Törnquist indexes of Deaton and Dupriez (2011), indicating lower aggregate variation. If states have high prices for some

4 An alternative would be to use quantities or values consumed per state as weights. This would have the advantage of "correcting" for variation in the consumption shares (which are endogenous and depend on prices, income levels and taste patterns). Using consumed quantities would however inflate the Ginis (shifting Lorenz curves to the right), while using the values of consumption would have an ambiguous impact depending on the demand elasticities. As a purely descriptive measure we therefore prefer population-weighting. This also gives gives rich and poor the same weight in measurement. In the later analysis of determinants, we will take the variation in consumption shares into account.

products and low prices for others, there may be modest aggregate price variation even if the product-level dispersion is higher. Since our focus is analysis of supply-side determinants and not to derive the overall price index or cost of living, we focus on the product-level variation where price gaps are generally larger.

Table 1 shows some key indicators summarizing the results based on NSSO data, using the 113 food items where the product description is the same in all three years.

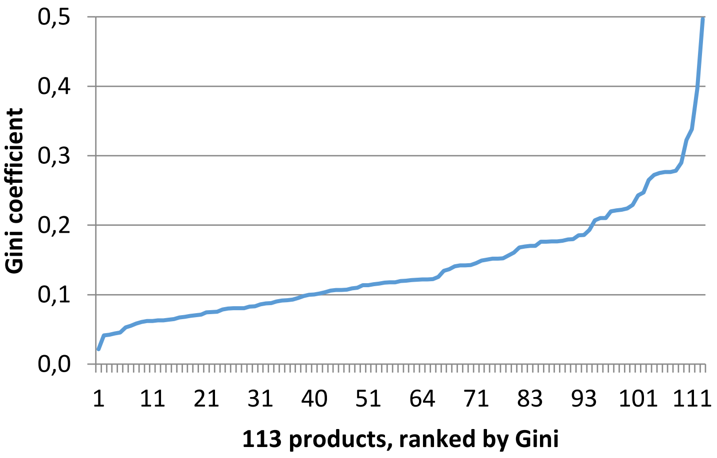
Table 1: Inter-state price variation for food products in India, based on NSSO household surveys: Summary of Gini coefficients for 113 food products.

	2004-2005	2009-2010	2011-2012
Median	0.171	0.122	0.120
Mean	0.200	0.141	0.140
Minimum	0.020	0.017	0.021
Maximum	0.551	0.457	0.497
Source: Own calculations based on NSSO 2007, 2012 and 2014b.			

The mean and median values confirm the statement above – there is large price variation across states and the highest prices are typically 2-3 times the lowest ones. As shown by the minimum and maximum values, however, there is huge dispersion across products. Appendix Table A1 lists the results for all the 113 products. Using data from the 68th survey in 2011-12, Diagram 3 shows the calculated Gini's for these products in 2011-12.

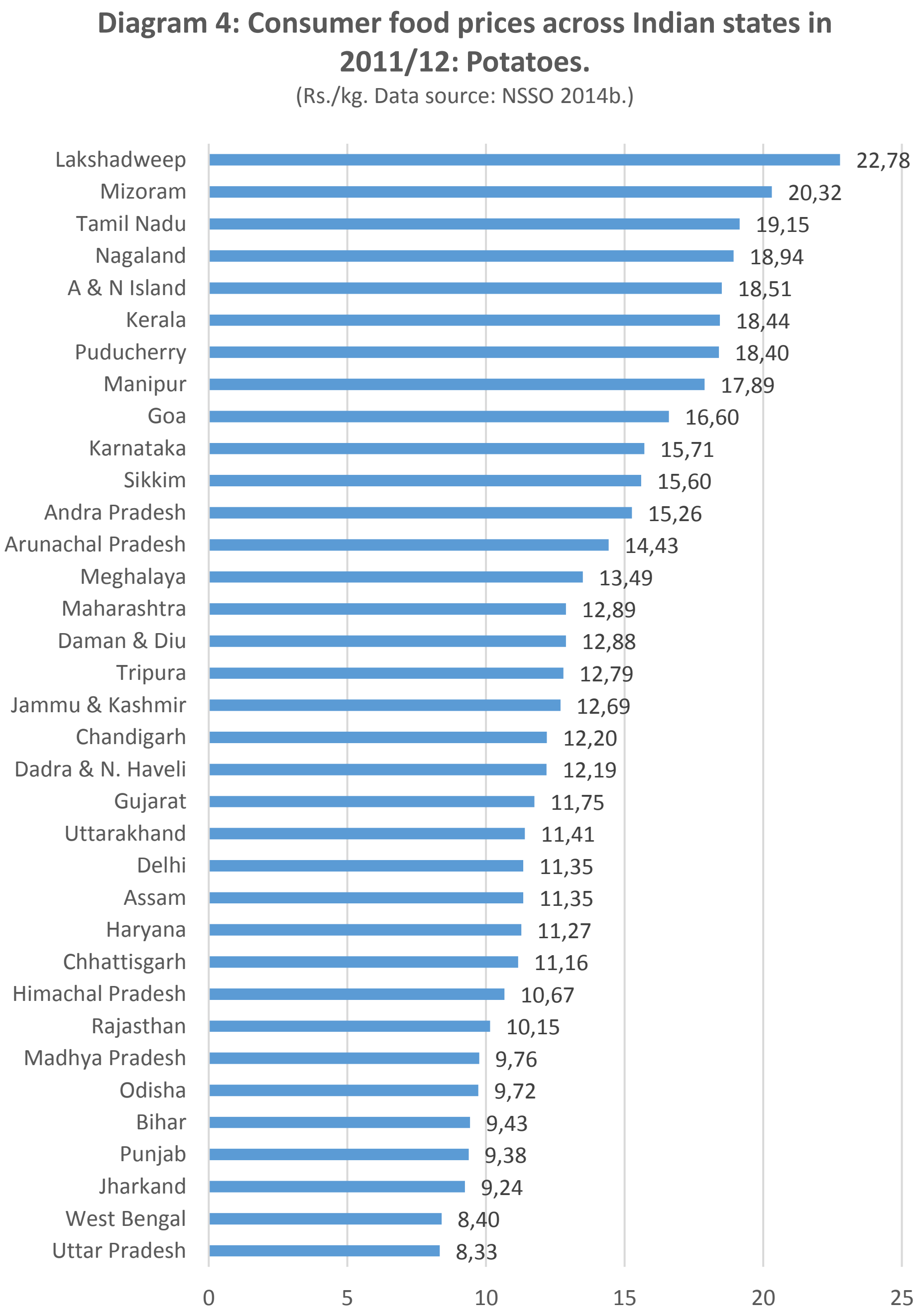
Diagram 3: Gini coefficients of inter-state price differences for 113 products in 2011/12

Source: Own calculations based on NSSO (2014b).



The Gini's vary from 0.02 to 0.50, hence patterns vary strongly across products. Most values are below 0.3, but a few products have more extreme values; here it should be noted that for some products, consumption in some states may be low so that data inaccuracy or differing consumption patterns may create outliers. Here we have used all observations with positive quantity and value of consumption, with no lower threshold.

As an illustration of what a Gini of a "normal" magnitude means, Diagram 4 shows the case of potatoes, with a Gini coefficient at 0.15 in 2011-12. This is an important product, accounting for 2.14 % of India's food consumption. Prices vary across states from 8 to 23 Rupees per kilogram, with a resulting Gini at 0.15. Hence prices vary not only by a few percentage points, but they double and treble across states. Some of the small peripheral states have very high prices, but these are given less weight in the Gini due to their small populations. The Gini of 0.15 therefore indicates considerable price variation also across the major states.



For the later analysis of determinants, a relevant observation is that the largest producers of potatoes in India are Uttar Pradesh and West Bengal. They also have the lowest prices, and this is a preliminary indication that proximity to production is important for price levels, as we would expect from the "trade cost" explanation of price gaps.

Diagram 5, with the corresponding data in Table 2, shows averages for broader product groups. Here the point of departure is the Gini coefficients for each product in 2011-12, and averages have been made using all-India value of consumption as weights.

Diagram 5: Food price variation across indian states in 2011/2012: Gini's (weighted average across individual products) for main product groups. Data source: NSSO 2014b.

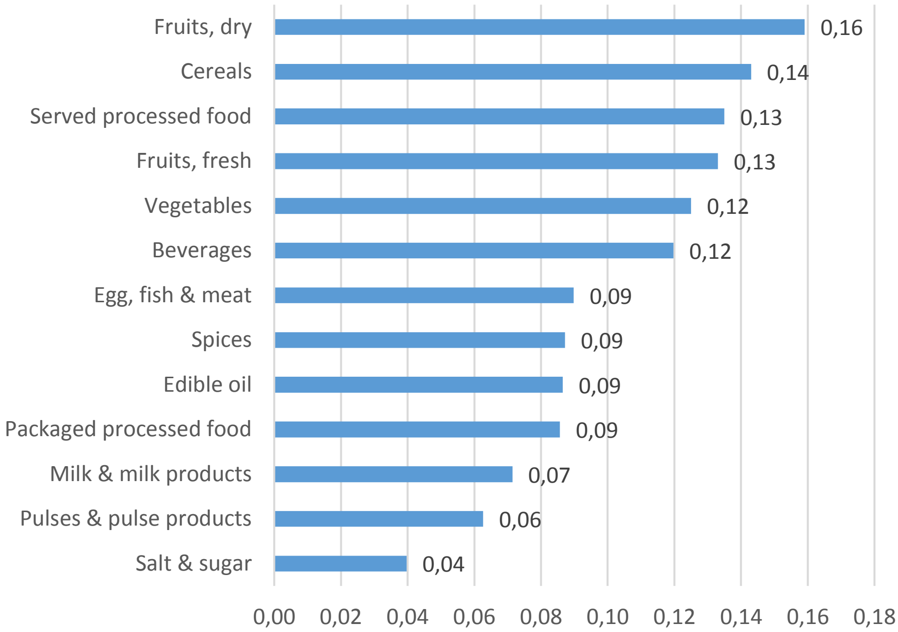


Table 2: Inter-state price dispersion in India for subgroups of food. Weighted averages of Gini coefficients for individual products, 2011/12.

NSSO number	Sub-category of food	% of food consumption	Gini (average)
179	Salt & sugar	3.17	0.040
159	Pulses & pulse products	5.23	0.063
169	Milk & milk products	15.50	0.071
299	Packaged processed food	1.27	0.086
189	Edible oil	6.76	0.087
269	Spices	6.27	0.087
199	Egg, fish & meat	8.87	0.090
279	Beverages	4.53	0.120
219	Vegetables	11.89	0.125
239	Fruits, fresh	4.72	0.133
289	Served processed food	5.33	0.135
129	Cereals	18.46	0.143
249	Fruits, dry	1.38	0.159

Source: Own calculations based on NSSO (2014b).

According to the trade-related explanation, price dispersion could be supported by trade costs – policy-driven or "natural" in the form of infrastructure, or by limited competition due to regulations of trade and commerce. For food trade, logistics requirements and perishability are main factors that create natural variation in trade costs across products. In the table, we find Salt & sugar with the lowest price dispersion; plausible since this item is easily transportable and not very perishable. In the diagram and table, we interestingly find the main staple categories Pulses and Cereals at opposite ends of the scale, with low price dispersion for pulses and large price gaps for cereals. Cereals are more regulated and regulation is one candidate explanation for the high extent of market segmentation for grains. Vegetables and fruit, which are the focus of the further analysis here, are in the upper range with considerable price dispersion. According to the value chain explanation, price dispersion could be larger for products with a larger services component. This could explain why e.g. fresh fruit and served processed food have high price dispersion. As shown by Deaton and Dupriez (2011), quality differences play some role for inter-state price variation. For some products, there could also be sub-varieties with different prices that could affect measurement and comparison across states.

From Appendix Table 1 an observation is also that for products in the public distribution system (PDS), price dispersion is systematically higher in PDS than outside. This is shown in Table 3.

Table 3: Price dispersion and the public distribution system

Product description	Gini coefficients		
	2004-05	2009-10	2011-12
Rice - P.D.S.	0.171	0.337	0.323
Rice - other sources	0.096	0.108	0.100
Wheat/atta - P.D.S.	0.227	0.185	0.277
Wheat/atta - other sources	0.149	0.145	0.180
Sugar - P.D.S.	0.057	0.057	0.136
Sugar - other sources	0.020	0.017	0.021

Source: Own calculations based on NSSO (2007, 2012, 2014b).

In Section 4 we revert to the analysis of determinants of price dispersion. As a preliminary step, we will examine here whether some states have systematically high or low prices, or if this varies across products. If price dispersion is driven by the "trade cost" explanation and states specialize in different products, price gaps should be related to trade costs from the supplying states to the destinations. Then states may have low prices for goods produced at home (as seen for the major potato growing states) and higher prices for those imported. If that is the case, price rankings for different products might not be so strongly correlated – states may have low prices for some goods and high prices for others. If, on the other hand, states with high income levels systematically have higher prices, it might be in line with the Balassa-Samuelson hypothesis or other explanations of the income-price relationship, as discussed in the introduction. If the income-price link is the main driver of price gaps, price differences should to a larger extent be correlated across products.

As a first measure of whether some states have particularly high or low price levels, we calculate a simple average of relative prices. By relative prices we mean the unit value for the state in question divided by the average price across all states for each product. We use 2011-12 data for all the individual products where we have price observations for at least 33 of the 35 states. In this way, we obtain a data set with 82-88 individual products for each state. Diagram 6 plots the average relative price against Gross State Domestic Product (GSDP) per capita for 30 states with data on both.[5]

Diagram 6: Relative food prices versus income levels across Indian states, 2011-2012.

Data sources: NSSO 68th Round and CSO.

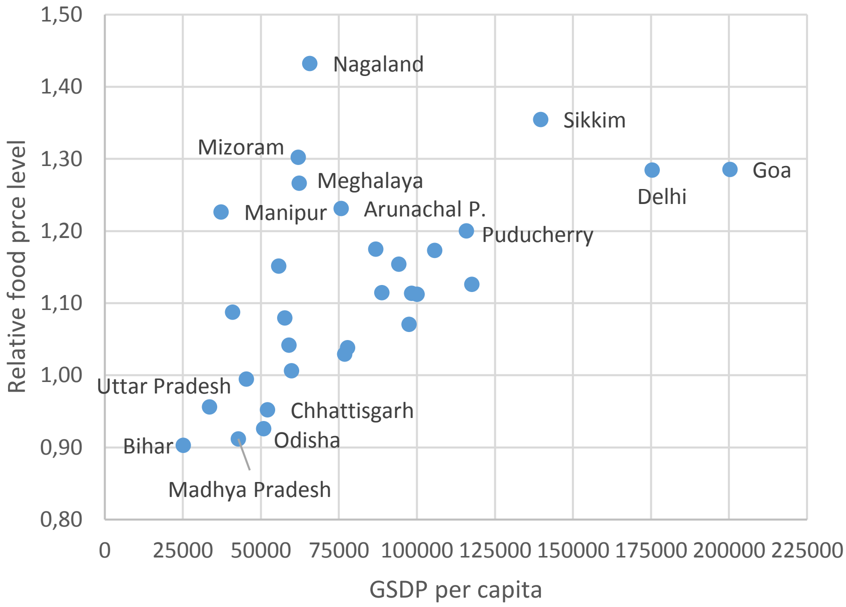


Diagram 6 shows that the average price dispersion is much lower than what we find for individual products, and closer to the range found by Deaton and Dupriez (2011). Using the average food price levels in Diagram 6, we obtain a Gini at 0.052; i.e. much lower than the median at the individual product level.

5 GSDP data are in Rupees and in current prices for 2011/12 and downloaded from data.gov.in, based on data from CSO (Central Statistical Office) that were processed and generated by the financial resources division, Planning commission.

There is a significant correlation between price and income levels (correlation coefficient: 0.54). As noted in the introduction, this phenomenon is well known from cross-country comparisons and it is interesting to observe that the same applies for inter-state comparisons within India. For China, a similar price-income correlation was found by Brandt and Holz (2006). For India, Deaton and Dupriez (2011) generally interpret this correlation as evidence of quality differentiation. While this surely matters, the value chain explanation is also a candidate. With respect to the Balassa-Samuelson hypothesis, we may observe that the prices we analyse are not for hair-dressers and non-traded goods, but for food products that should indeed be tradable within India.

Based on this tentative evidence, some states stand out with particularly high or low price levels for food. [6]

– A & N Island, Nagaland, Delhi, Goa, Maharashtra, Tamil Nadu and Gujarat can be said to have particularly high price levels for food. This is a mixture of different types; with some rich states or union territories, and some small and peripheral.

– West Bengal, Tripura, Jharkand, Chhattisgarh, Madhya Pradesh, Uttar Pradesh, Odisha and Bihar have particularly low price levels for food. In general, these are states with relatively low income levels.

This analysis therefore suggests that price an income levels are closely correlated but in addition, the high prices in some small and peripheral states indicate that economic geography and trade costs also play a role.[7]

During the last 25 years, India has undergone significant development and reforms, and an issue is also whether price gaps have remained the same over time or whether there has been a change. Has increased trade integration led to less market integration and therefore

---

6 As an additional check, we run a simple panel regression with fixed effects for the states, regressing the relative prices for each state/product on the value of consumption for the same state/product. Here the right hand side variable is mainly auxiliary; the main output is the fixed effects, measuring whether a state's price level deviates systematically from the average. Table A2 in the Appendix shows the average relative price and the corresponding rank for each state, and the rank based on the fixed effects. The correlation between the two rankings is 0.52. The fixed effects in the regressions are significantly different from zero in 18 out of 35 cases. In general, the fixed effects were often not significant for the intermediate cases, whereas states with low or high price levels are often confirmed by significant fixed effects in the regressions.

7 Observe also that although there is a correlation between price and income levels, the population-weighted Gini coefficient based on relative price averages is 0.052; i.e. clearly lower than the median for product-level Ginis. This suggests that some of the price differences are product-specific so ranking vary across products, but some are related to the properties of states and common across products.

lower price dispersion? For example, the EU Commission has the clearly stated aim and belief that increased market integration should lead to price convergence (Goldberg and Verboven 2005). From our results and arguments so far, this is however not so clear: if price dispersion has one "income driver" and another "trade cost driver", it is not certain the trade integration – which affects the trade cost driver – is enough to cause price convergence.

Our NSSO sample with three years is limited but nevertheless suggests that there has been a decline in price dispersion from 2004-05 to the last two years. Table 1 shows a clear drop in median and average Gini coefficients from 2004-05 to later years. Regressing Gini coefficients for 2011/12 and 2009/10 on the preceding years, the slope coefficients are significant and suggest a systematic decline in price dispersion; especially from 2004/05 to the later years but even from 2009/10 to 2011/12.[8] With some caution due to the limited number of years, this suggests that food price dispersion in India declined during the last decade. One possible explanation is that intra-Indian trade integration has improved so that trade cost-driven price differences have been reduced. Another possibility, linked to the observation in Diagram 1, is that price inflation has been higher in poorer states and this has led to some price convergence.

Our result here is contrary to Atkin (2013), who used data from household surveys in 1987/88 and 2004/05 and concluded that there was some increase in agricultural price dispersion. The trend may therefore have changed after 2004/5. A possible reason for this difference could be the pattern observed in Diagram 1, which suggests that something new occurred with regional price gaps after 2007. Since regional income disparities grew only modestly after 2007 (or even fell in current values), a possibility is that increased trade integration in India could be the driver of convergence. In the later analysis, however, we are not able to obtain clear results on the drivers underlying this convergence so it remains a puzzle, and a kind of hypothesis only until we have been able to trace the causes.[9]

---

8 The regression slope coefficients/R^2 were 0.32/0.25 (2011/12 Ginis regressed on 2004/05 Ginis); 0.35/0.33 (2009/10 on 2004/05); and 0.75/0.57 (2011/12 on 2009/10), with small standard errors indicating that the values were all significantly below one – confirming a decline. Correlation coefficients are 0.58 (2004/05 and 2009/10); 0.50 (2004/05 and 2011/12); and 0.76 (2009/10 and2011/12), showing that Ginis are strongly cortrelated but not very stable over time and there are many outliers, perhaps caused by "marginal" products have a small share oif consumptioin in some states.

9 An issue could be related to data, e.g. India's population in 2004/2005 should be 981 million according to the household survey data while data from NITI Ayyog suggest 1089 million. The Ginis for 2004/05 were therefore recalculated with population data from the Planning Commission; however, the results were only marginally affected.

3. Price dispersion along the value chain: Evidence from mandi price data

From a value chain perspective, it should be recalled that there are different prices at each stage in the value chain, depending on how many times the goods change ownership. The number of transactions could vary across products and their use; e.g. Chand (2006) present four stages as a standard case but there could be more or less. If a product is sold directly from the farmer to the village neighbour, the producer may retain a large share of the consumer price; but if the product passes through several stages the producer may get only a small fraction of the final price. The NSSO unit values used in Section 2 are measured at the consumer end, but cover urban as well as rural inhabitants, and they are not directly observed prices but based on different options for the valuation of consumption.

The inter-state variation in prices may be different at different steps in the value chain, and it is not certain that the inter-state price dispersion using NSSO unit values pertain to other prices along the value chain. As a second source of price information we will therefore use mandi (wholesale market) prices. These prices are more upstream in the value chain; at some stage between farmers and consumers. The mandis buy locally produced food but also food from other regions so we do not know to what extent the mandi prices include transport costs. The mandi traders may sell products onward for local consumption (e.g. via retailers) but also to other cities and states (or even internationally). The magnitude of large mandis such as the famous Azadpur mandi in Delhi suggests that proximity to consumption is a major determinant so we should expect that mandi prices are significantly higher than producer prices, due to transport costs.[10]

While the NSSO data cover all products, it takes more effort to collect mandi prices. Such data is however available from various web sources. For this more in-depth analysis we focus on horticulture since data is more easily available via the National Horticulture Board (NHB, www.nhb.gov.in).[11] For a number of horticulture products, NHB has systematically collected prices for a number of products from all the major mandis of India, aggregated for various market areas. For example,

[10] The difference between wholesale and producer prices is not examined further, since we do not have producer prices at hand for this study.

[11] Another web source of price data is agmarknet.nic.in.

Delhi has several markets but in the NHB data these are aggregated under the heading Delhi.

A subset of products was selected using the following criteria:

- Products should be relatively homogenous so that prices are comparable across states.
- Products should match the products covered by NSSO data.
- Products should be important in the sense that they have a significant share of consumption. This assures that data is "thicker" and we obtain more reliable price observations across states.
- For the later analysis of determinants, production data should be available, in order to facilitate the analysis of "trade cost/economic geography" drivers of price disparities.
- We also deliberately drop the more regulated food markets such as grains since we are interested in finding evidence on how markets and trade work.

Using these criteria, we end up with a list of 11 products, covering 10 percent of food consumption in India. In the NHB price data, four products are split into two subcategories (potatoes, apples, tomatoes, brinjal) so we examine prices for 15 items. Table A3 in the Appendix shows the shares of food consumption for each product, as well as the Gini for interstate price disparities calculated earlier on the basis of NSSO data. These are products with varying degrees of price dispersion based on NSSO data, and the products also vary in terms of perishability, including e.g. fresh fruits such as grapes and banana with high perishability, and vegetables with varying degrees of perishability. [12] It should however be emphasized that our sample is not representative and our aim is not to find the true cost of living for all food items.

For the selected products, we collected monthly wholesale price data from January 2004 to June 2015. While monthly data allow analysis of the considerable short-term price fluctuations in agriculture, this is not the focus here and we aggregate data into years. With NSSO data covering July-June and the Indian fiscal year April-March, we aggregate as appropriate for the purpose at hand.

The data collected for the 11/15 products cover 31 market areas across 23 states. Hence some of the states have more than one market area, and some states or union territories have no market areas (in this

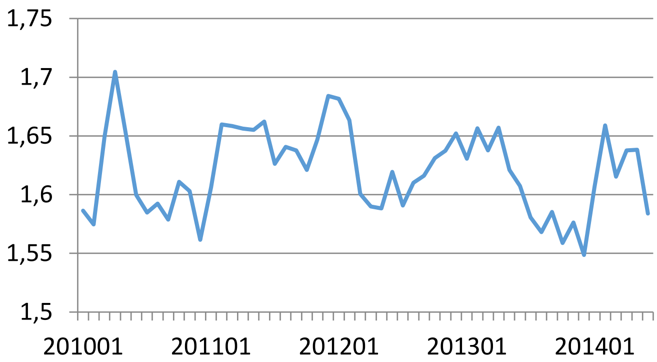
---

[12] Some other products were also included at first but were dropped due to lack of data for the later analysis.

data set). Each market area covers different mandis.[13] For each product, market area and month, the data includes arrivals (volume) and the wholesale price. The data set has 51601 observations; however, with zero arrivals (and therefore no wholesale prices) in 5195 cases (10% of all observations). The data set also includes retail prices, however with many missing observations for 2004-2009 so we use retail price data only from 2010, or from 2008 if we do not need complete data. The retail prices are collected from retail outlets in the respective mandi areas.[14] As we would expect, the retail prices are higher than wholesale process. Diagram 7 shows a simple average across all observations for each month since 2010, for the retail/wholesale price ratio.

Diagram 7: The average ratio between retail and wholesale prices for 15 food products in India

Source: Own calculations based on data from NHB.



Throughout the period, the ratio fluctuates around a level of 1.6, suggesting that retail prices are on average about 60% higher than the wholesale prices. The median (mean) across all 20194 observations is 1.52 (1.62). There is some variation across products and markets but not extremely so. Table 4 shows the range of values remaining if we delete the top and bottom 1% of the distribution, then 5% and so on. The table shows e.g. that 95% of the observations are in the range 1.181-2.401.

[13] See "Directory of Wholesale Agricultural Produce Assembling Markets in India", available at agmarknet.nic.in and published by the Indian Ministry of Agriculture in 2004. There is no metadata telling explicitly which mandis are covered by the data.

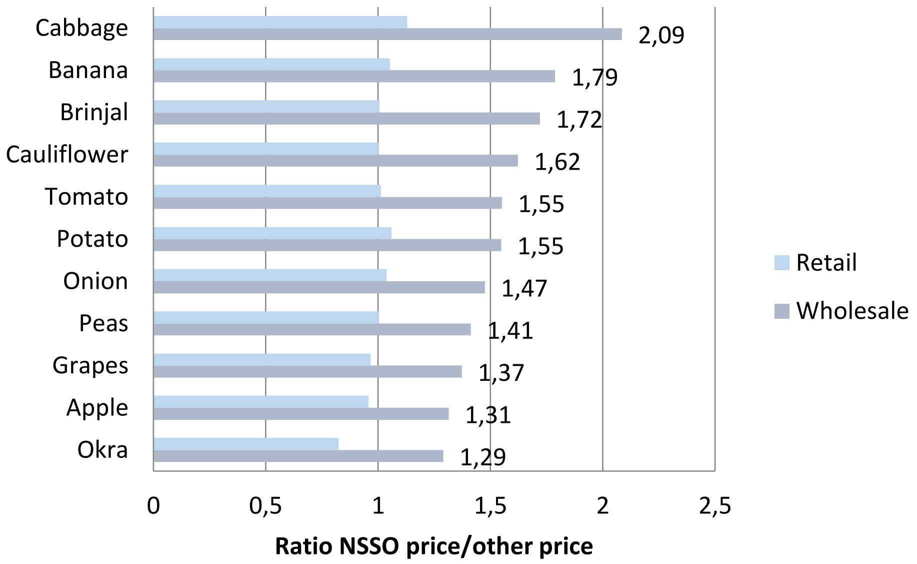
[14] The statement is based on interviews with NHB staff in Bangalore.

Table 4: Quintile distributions and value ranges for the retail/wholesale price ratio, for 15 food products in India		
Share cut at top and bottom	Remaining range of values	
	Lower	Upper
1%	1.096	3.380
5%	1.181	2.401
10%	1.231	2.105
25%	1.336	1.777
Note: Own calculations based on data from NHB.		

It is of interest to examine how the mandi price data compare to the NSSO unit values. We expect that the NSSO prices should be comparable to the retail price data from NHB, but clearly higher than the wholesale prices. This comparison can only be undertaken for the years covered by NSSO data, and for the 23 states that have mandis in the NHB data. We therefore compare the NHB prices with NSSO prices for the states where the respective mandis are located. We aggregate the NHB data into the 11 products covered by NSSO, and the same time periods (July to June). Calculating the ratios between NSSO unit values and the NHB prices for each product, year and state/mandi observation, Diagram 8 presents averages for the 11/15 product groups. For wholesale prices the average includes 2004/5, 2009/10 and 2011/12; for retail prices only the last of these years is included.

Diagram 8: Ratio between NSSO unit values and NHB wholesale/retail prices (averages)

Source: Own calculations based on data from NSSO and NHB.



There is a very close correspondence between NSSO unit values and the NHB retail prices, with values close to one in most cases. This confirms that the NSSO unit values are close to directly measured retail prices and provide a valid source for the analysis of inter-state price differences.

The average ratio NSSO unit value/NHB wholesale price is expected to be above one and ranges from 1.29 to 2.09 across the 11 products. The simple average is at 1.56; i.e. in the same range as shown by using the NHB data alone (as shown in Diagram 7). Hence we have a double indication that this is an appropriate estimate for the average range for the retail/wholesale markup in India for these food products during the time period covered. Diagram 8 shows that there is considerable variation across products.

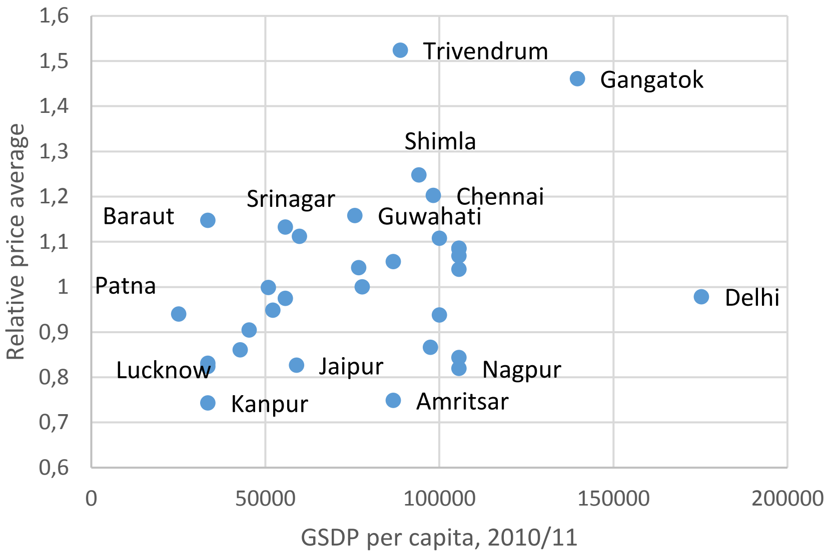
An issue is whether price dispersion is lower for wholesale prices, since the value chain driver of price dispersion is weaker for the wholesale prices than for consumer prices. We also calculate Ginis for wholesale price dispersion across mandis, using volume shares as weights. For the 11 products in Diagram 8, there is some variation but on the whole, the levels of price dispersion are quite similar. For 2011/12, the average Gini for the 11 products was 0.129 for wholesale prices, compared to 0.125 for NSSO consumer prices. The two results are not strictly comparable due to the different weights used (population versus the volume of arrivals), but suggests that price dispersion is comparable to that observed for consumer prices. Diagram 9 shows (similar to Diagram 6 for consumer prices) average price levels for the different market (average relative prices for the whole period covered by data) areas against the 2011/12 income levels of the states where they are located. There is also here a positive correspondence but perhaps weaker than suggested for consumer prices in Diagram 6.[15]

---

[15] The average price levels in different market areas are reported in Appendix Table A5, together with results from fixed effect regressions that provide a similar ranking of price levels. For brevity we do not discuss the details.

Diagram 9: Mandi price levels versus GSDP per capita, 2004-2014

Source: Own caculations based on NHB and Niti Ayyog data.



Hence Delhi is an outlier with relatively low prices in spite of its high income level.

4. Determinants of inter-state price differences

Motivated by research on international price differences, we are interested in finding out whether the "trade and economic geography" explanation or the income-price link is the key driving force behind price gaps, or if they both matter. We therefore combine the price data with annual data on state characteristics and the statewise production of the different crops. The data sources are listed in Appendix Table A6 and the variables are described in the following section. In general, more data is missing for the latest years so most of the analysis will be undertaken for the period 2004/05 to 2011/12.

It is important to observe that we do not intend to estimate demand or supply elasticities but undertake a meso-level analysis of what determines relative price differences. For example, the mandi price data is a rich source of information on short-term price fluctuations but these are not addressed here since the main focus is on inter-state gaps and not to estimate the demand and supply curves. It is still important to be aware of the classical econometric issues about identification and endogeneity in estimation. As an illustration, Diagram 10 shows a hypothetical case where we assume that two states have the same supply curves but that demand for the product is higher in state 2, so the demand curve D_2 is above D_1. The traded quantity and price in state 2 are therefore both higher.

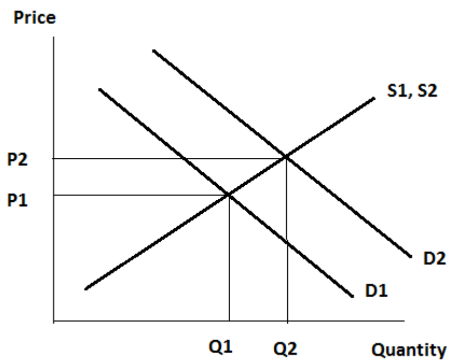


Diagram 10: Identification issues in estimation

Diagram 10 illustrates that shifts in demand identifies the supply curve (since the observed data points are along the supply curve). In a similar way, shifts in supply would identify the demand curve. In Diagram 10 the slope will be positive; with supply shifts the slope would be negative.

At equilibrium we must have supply=demand in both countries which gives the "observed" points (Q1, P1) and (Q2, P2) along the supply curve. Now assume that both countries have the same supply curves $S_i = \alpha P_i^{\beta s}$ and the demand curves $D_i = \gamma_i P_i^{\beta d}$ where the constant γ_i is assumed higher for country 2. Now at equilibrium we must have

$$\frac{Q_1}{Q_2} = \left(\frac{P_1}{P_2}\right)^{\beta s} = \frac{\gamma_1}{\gamma_2}\left(\frac{P_1}{P_2}\right)^{\beta d}$$

Or we can multiply by P1/P2 to get the "observed" values V1 and V2:

$$\frac{V_1}{V_2} = \left(\frac{P_1}{P_2}\right)^{1-\beta s} = \frac{\gamma_1}{\gamma_2}\left(\frac{P_1}{P_2}\right)^{1-\beta d}$$

From this we can express P_1/P_2 as a function of V_1/V_2 or Q_1/Q_2, in two different ways depending on whether we use the supply or demand parameters. Hence if we regress relative prices on volume or value shares of consumption (as part of an equation), the resulting estimates could differ depending on whether we include a constant term or not, and the parameter estimates would vary depending on the specification.

In the following estimation, there will be several forces at work, and demand and supply shifts at the same time. The simple relationships above demonstrate that endogeneity is not necessarily a problem as long as we are not trying to identify the true demand or supply elasticities. The results could also be sensitive to the specification, e.g. whether we include dummies or not. With simultaneous demand and supply shifts, it is also possible that the results are insignificant even if there is an underlying relationship as in Diagram 9.

4.1. Explanation of variables: Analysing trade without trade data!

For the trade and economic geography explanation of price differences, we should ideally have data on production, inter-state trade and inter-state trade barriers. While state-level production data is at hand, data on inter-state trade in food products, or on trade costs between states, are not available. We therefore have to find an indirect way of capturing the role of trade and trade frictions. In order to do this, we use the gravity equation as a point of departure (Head and Mayer 2014). The gravity equation is essentially a strongly supported empirical relationship in international trade, with the key relationship taking the form

$$T_{ij} = \alpha * S_i^{\beta_i} * S_j^{\beta_j} * d_{ij}^{\gamma}$$

where T_{ij} is trade from region i to j, S_i and S_j are measures of economic size, and d_{ij} is geographical distance between i and j. In analysis of international trade flows, the β's are often close to but below one, and γ is negative and often close to minus one. Given the strong empirical support for this relationship, we assume that is also applies to inter-state trade within India.

According to Head and Mayer (2014, 154), a distance elasticity of minus one is representative for the empirical literature. The size elasticity β_j is often close to one (ibid., 133). Letting T_{ij} denote the imports of region i from region j, we therefore simplify by assuming that the the β's are equal to one, and γ is equal to minus one; i.e. in the typical range of results in empirical studies. We then obtain the relationship

$$\frac{T_{ij}}{S_i} = \alpha * \frac{S_j}{d_{ij}}$$

Summing this for imports from all sources, we obtain an expression for total imports of a product, relative to the importing region's size. We assume that the larger are imports, the greater is the competitive impact and the lower are prices. In the analysis here we let S_j be represented by the exporting region's share of all-India production.

With this motivation, we construct the variable DISTPROD (distance to production), which is an index of proximity to other supplying states in India but not including a state's own production. The index takes the form

$$DISTPROD_{ikt} = \sum\nolimits_{j \neq i} \frac{s_{jkt}}{(a + d_{ij})}$$

Here s_{jkt} is the share of region j in India's production of product k in year t. d_{ij} is the geographical distance between states i and j, using great circle distances based on the geographical coordinates of state/union territory administrative centres. The parameter *a* is a constant added to distance in order to avoid that observations with zero distance unduly inflate the measure and affect the results.[16] It also matters for the scaling of the DISTPROD variable and we choose a relatively low value that renders a relatively even distribution in the variable range. Using a=100, we obtain 11 (products) * 35 (states and union territories) * 8 (years) = 3080 observations of DISTPROD, ranging from 0.19 to a maximum of 3.64. The smaller is DISTPROD, the more remote are the large suppliers of a

16 In the data set, geographical distance between regions varies from zero to 3003.There are some off-diagonal observations at zero since some states have shared administrative capitals. Without the constant a, these terms in the would go to infinity. The constant a eliminates the problem, and whether it is set at e.g. 50 or 100 does not have too much to say.

product. If trade frictions are high, we expect from the trade-based explanation that differences in DISTPROD will be a more important driver of price differences across states.[17]

From a trade perspective, a state's own production will also be a key determinant of the price level. If a state has a comparative advantage for a product, domestic supply will be larger and this tends to drive the price downward. From the Heckscher-Ohlin-Samuelson (HOS) theory, we know that if there is no trade, the autarky price will be higher for the products where a country or region has a comparative advantage. With free trade, prices will be equalized across nations or states. The more barriers there are to trade, the more important will domestic production be for the domestic price. For food prices this may be particularly urgent since there are seasonal harvests and if products cannot be sold to other states or internationally, the farmers' only option is to sell them at a low price in the domestic market.

Statewise production for each product is therefore a key variable in the analysis. As shown in Appendix Table 7, a problem is that production volume data is incomplete but we have data on the value of production in current as well as fixed prices. By combining these three sources and extrapolating we are able to construct complete time series from 2004/05 to 2011/12 for $PRODVOL_{ikt}$ $PRODVAL_{ikt}$ – the quantity and value (in fixed prices) of production for each product, state and year.[18]

In the regressions, we use the shares of all-India production $PRODSHAR_{ikt}$ as variables. I.e.

$$PRODSHAR_{ikt} = \frac{PRODVOL_{ikt}}{\sum_i PRODVOL_{ikt}}$$

Large producers will have high values and we expect this to be inversely related to price levels.

Small states may however be net exporters or "abundant" for a product even if their share of national production is modest. Hence it would be appropriate to measure the size of production related to the size of

[17] Modern gravity also includes "multilateral resistance terms" (Head and Mayer 2014, 150) capturing the average distance to other trade partners for the exporting and importing region/countries. Since DISTPROD is multilateral and not bilateral, we do not include such a term in its definition.

[18] A problem is that the Indian Horticultura Database for relevant years only reports the production of main suppliers and we have complete volume data for all states only for some years. Using value data in fixed prices we could (i) verify for which states production was actually zero and not only unreported; and (ii) extrapolate between years to complete the production volume time series for small suppliers. E.g. if we know that production volume in year t was 100 and the the value in fixed prices increased by 12% from this year to the next, we extrapolate the value of 112 for the production volume in year t+1.

domestic consumption. From the NSSO household survey data we can derive product- and statewise values of consumption and we therefore form the ratio:

$$PRODCONS_{ikt} = \frac{PRODVAL_{ikt}}{CONSVAL_{ikt}}$$

Also in this case, we expect an inverse relationship to prices; the higher is PRODCONS, the lower is the price. This measure is available only for the three years with household survey data; i.e. 2004/5, 2009/10 and 2011/12. There is an issue of endogeneity since consumption values are affected by the price; but for our purpose this will be acceptable, as discussed in the context of identification earlier. PRODCONS and PRODSHAR are correlated (correlation coefficient = 0.21) but capture different aspects and there is not a collinearity problem in the regressions.

In order to examine the price-income link, we use the statewise income level, in the form of statewise per capita GDP relative to the all-India average. We have data in current and fixed prices and generally use current values, $GDPCCURR_{it}$. From a value chain perspective or other explanations suggested in the international literature, we expect a positive relationship so that states with high income levels also have higher prices.[19]

The variables PRODSHAR, DISTPROD, PRODCONS and GDPCCURR are the key variables we use to examine the drivers of inter-state price variation. In addition, we include some other variables based on econometric and other scientific reasons.

An econometric issue is about spatial correlation; i.e. that prices may be correlated between states that are closer to one another and this may influence residuals in the regressions unless it is accounted for. Our variable DISTPROD addresses spatial correlation related to the location of production. In addition, there may be other forms of spatial interaction that affect price levels. We therefore introduce the variable

$$PERIPHER_{it} = \sum_{j \neq i} \frac{r_{jt}}{\left(b + d_{ij}\right)}$$

where r_{jt} is the share of state j in India's GDP in year t and d_{ij} is distance, as before. Also here, the constant b is necessary for handling zero distances. Using a value for the constant at b=50, we obtain a variable with 35 (states and territories) * 10 (years) = 350 observations ranging from 0.52 to 2.75. Peripheral states with large distances to the economic mass

19 GDPCCURR could also be relevant from an economic geography perspective since geography could be a driver of income level differences. Also in this case, there would be a positive correlation with prices.

of India have low values of this variable.[20] The PERIPHER variable will be correlated with DISTPROD but captures broader aspects of spatial interaction that is not linked to individual products and that may influence cost levels. This variable is linked to economic geography and thereby "trade-related" but is also relevant in the value chain perspective. We generally expect an inverse relationship between PERIPHER and price levels, but the patterns could be more complex. According to Melchior (2010), India is not "monocentric" in terms of income level variation and core-periphery patterns, but "polycentric" with different areas with higher economic concentration and income levels. In this case, a state between two such centres may be backward and still have a high value for the PERIPHER index.[21]

An issue for the analysis is whether consumption patterns are similar across states. If this is not the case, the higher demand for a product in some states may drive up prices. An interesting contribution in the field is Atkin (2013), who show that there is a regional "consumption bias" in Indian agriculture so that people have stronger preferences for locally produced goods, due to habit formation. In the regressions based on NSSO data, it is possible to control for differences in consumption patterns. In order to control for differences in consumption patterns, we therefore include the variable

$$CONSSHAR_{ikt} = \frac{c_{ikt}}{c_{kt}}$$

where c_{ikt} is the value share of product k in total food consumption for state in year t, and c_{kt} is the corresponding share for All-India. We generally expect a positive relationship to prices, since higher demand tends to bid up prices. Regarding the issue of endogeneity, we refer to Diagram 9 and the subsequent discussion.

In the regressions, we also include POPSHARE; a state's share of India's population. This is a measure of size that we include since state size may affect prices in various ways. From a trade and economic geography perspective, there may be issues of economies of scale and competition. In this case, larger states could have more firms and therefore more competition, and be better able to exploit economies of scale. There would then be a negative relationship so that large states have – other things

20 In gravity regressions for international trade, it has become standard to include a variable measuring "multilateral resistance". The PERIPHER variable is partly similar (see Head and Mayer 2014, 150) but may have different interpretations here.

21 As a third attempt to capture spatial correlation, we also constructed a variable reflecting the average income levels of the 10 closest states. The idea was to check for other spillovers or "contagion" effects that might affect prices. This variable had however little explanatory power and was therefore dropped from the analysis.

equal – lower prices. There could also be other competition issues involved so that e.g. limited competition has a stronger impact on prices in smaller states, due to a smaller number of competitors.

As a supplementary variable we also include NATROAD, a measure of national roads per square kilometre by state. Ideally we should also have data on roads between states, but we do not have appropriate data at hand. NATROAD intends to capture the development of infrastructure at the state level. We expect a negative relationship to prices since better roads will increase economic efficiency and lower costs and prices.

In regressions using mandi price data we also include $ARRIVAL_{mkt}$; the share of each market area m in total arrivals for product k in period t. Since larger supply tends to reduce prices we expect a negative relationship. For the analysis of wholesale prices, the arrival variable represents supply shifts, along with PRODSHAR.

In the estimations, all variables are expressed in relative terms; either relative to the all-India average or as shares of the all-India total. We thereby focus exclusively in inter-state variation and not how quantities and prices change over time.

4.2. Results from regression analysis

We run the two regressions shown in Table 5.

Table 5: Variables in regressions on the determinants of inter-state price differences
Note: All variables are expressed in logs except for PRODSHAR.

Independent variable	Abbrevi-ation	Expected sign	Dependent variable:	
			I	II
			RELPRICENSSO	RELPRICENHB
Prod./cons. ratio	PRODCONS	Negative	X	
Production share	PRODSHAR	Negative	X	X
Distance to other suppliers	DISTPROD	Negative	X	X
Peripherality	PERIPHER	Negative	X	X
Per capita income (relative)	GSDPCURR	Positive	X	X
Population share	POPSHARE	Negative	X	X
National roads	NATROAD	Negative	X	X
Consumption share (relative)	CONSSHAR	Positive	X	
Share of mandi arrivals	ARRIVAL	Negative		X
Time intervals			Annual	Annual
Years covered			2004/05, 2009/10 2011/12	2004/05 to 2011/12

PRODSHAR has often a value of zero and is therefore not expressed in logs. All other variables are expressed in logs. Consumption-related variables are not available for all the years covered by NHB data, so PRODCONS and CONSSHAR are not included in (II). ARRIVAL is not relevant for equation (I) based on NSSO data.

The results are shown in Tables 6a (NSSO consumption prices) and 6b (NHB wholesale prices). Estimates that are statistically different from zero at a level of 5% or better are marked with shading. We run each equation in three forms; with or without dummies for states (equations Ia and Ic) or mandis/ market areas (equations IIa and IIc), and with clustered standard errors (Ib – clustering by state, and IIb – clustering by market area). We always include product and time dummies.[22]

[22] Results for dummies are not reported but may be obtained from the author upon request.

Table 6a: Results from pooled regressions on the determinants of inter-state relative price differences
Note: In the regressions, all variables except "Production share" are expressed in logs. Statistically significant estimates in shading.

Variable	Expected sign	Dependent variable: Relative consumption price								
		(Ia)			(Ib)			(Ic)		
		Standard OLS			Clustered standard errors			With state dummies		
		Estimate	t stat.	P value	Estimate	Z stat.	P > |Z|	Estimate	t stat.	P value
Prod./cons. ratio	Negative	-0.159	-2.02	0.0435	-0.155	-3.31	0.0009	-0.179	-2.12	0.0339
Production share	Negative	-1.069	-6.31	<.0001	-1.299	-5.62	<.0001	-1.031	-6.58	<.0001
Dist. to other suppliers	Negative	-0.344	-13.47	<.0001	-0.339	-10.25	<.0001	-0.315	-11.86	<.0001
Peripherality	Negative	-0.109	-4.22	<.0001	-0.110	-1.74	0.0819	-1.172	-1.11	0.2654
Per capita income	Positive	0.159	9.11	<.0001	0.159	4.72	<.0001	-0.127	-1.50	0.1329
Population	Negative	-0.047	-8.01	<.0001	-0.045	-4.44	<.0001	-0.095	-0.29	0.7683
National roads	Negative	-0.040	-3.44	0.0006	-0.039	-1.44	0.1493	-0.044	-0.73	0.4635
Consumption share	Positive	0.032	1.38	0.1693	0.032	0.98	0.3252	0.035	1.48	0.1397
Adjusted R^2		0.495			n.a.			0.537		
State dummies/ clustering by state		No			Clustering			State dummies		
Time dummies		Yes								
Product dummies		Yes								
Observations		1056								
Years covered		2004/5, 2009/10, 2011/12								

Note: t statistics and P values are heteroskedasticity-corrected. Date sources are described in Appendix Table 7.

Table 6b: Results from pooled regressions on the determinants of inter-state relative price differences

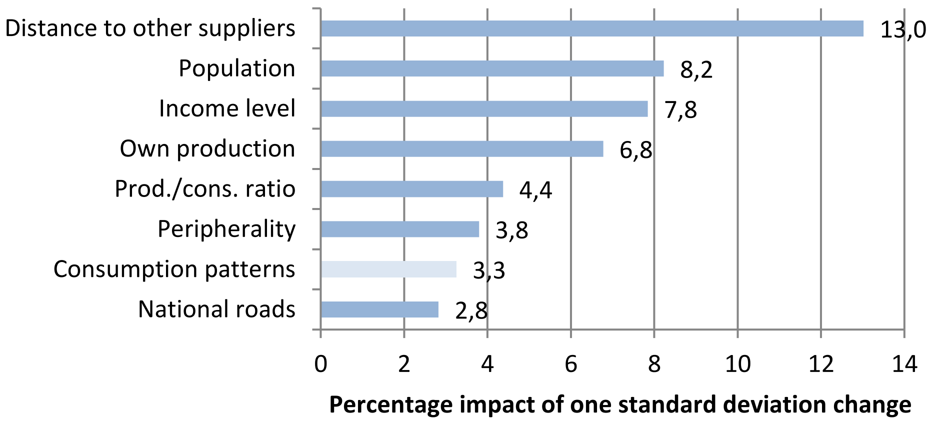
Note: In the regressions, all variables except "Production share" are expressed in logs. Statistically significant estimates <5% in shading.

Variable	Expected sign	Dependent variable: Relative wholesale price								
		(Ia) Standard OLS			(Ib) Clustered standard errors			(Ic) With market area dummies		
		Estimate	t stat.	P value	Estimate	Z stat.	P > |Z|	Estimate	t stat.	P value
Production share	Negative	-0.401	-8.89	<.0001	-0.4012	-4.07	<.0001	-0.391	-9.26	<.0001
Dist. to other suppliers	Negative	-0.189	-9.79	<.0001	-0.1890	-4.31	<.0001	-0.163	-8.85	<.0001
Peripherality	Negative	-0.189	-7.78	<.0001	-0.1887	-1.68	0.0933	0.925	1.38	0.1692
Per capita income	Positive	0.020	1.54	0.1248	0.0202	0.42	0.6718	0.105	1.47	0.1427
Population	Negative	-0.042	-9.18	<.0001	-0.0418	-2.14	0.0326	-0.393	-1.51	0.1315
National roads	Negative	0.007	0.9	0.3663	0.0073	0.21	0.8376	-0.031	-0.38	0.7027
Share of mandi arrivals	Negative	-0.020	-4.62	<.0001	-0.0197	-1.38	0.1691	-0.007	-1.21	0.2269
Adjusted R^2		0.191			n.a.			0.3597		
Market area dummies/ clustering		No			Clustering by market area			Market area dummies		
Time dummies		Yes								
Product dummies		Yes								
Observations		2694								
Years covered		2004/5 to 2011/12								

Note: t statistics and P values are heteroskedasticity-corrected. Date sources are described in Appendix Table 7.

In equation (Ia) with relative prices from NSSO household survey data as the dependent variable and without state-level dummies, all variables have the expected signs and are statistically highly significant, except for PRODCONS that has a lower significance level and CONSSHAR which is not statistically significant at all. Measured by adjusted R^2 the regression explains about half the inter-state variation in prices. This is a good fit, showing that a large share of the inter-price variation in consumer prices is explained by the regression. In Diagram 11, we show the impact of a one standard deviation change for each variable (with brighter shading for the consumption pattern variable since it is not statistically significant), using results from (Ia).

Diagram 11: Ranking of the determinants of inter-state consumer price differences for 11 food products



The diagram clearly shows the importance of the trade and economic geography variables, with distance to other suppliers (DISTPROD) on top and the production share (PRODSHAR) and the production/consumption ratio (PRODCONS) in the mid-range. The highly significant estimate for the income level variable (GSDPCURR) suggests that the income-related explanations also play a role. Peripherality and national roads also contribute significantly to the explanation, but population/size is even more important and indicates that other things equal, large states have lower prices. Whether this is due to economic efficiency (e.g. better exploitation of scale economies) or the degree of competition (more traders, more suppliers) is not possible to say from the results.

Some of the independent variables vary only by state/year; sometimes with only small variation over time. In equation (Ia) we capture all the cross-section variation but there is a risk that residuals are correlated for each state. In equations (Ib) (clustered errors) and (Ic) (state dummies) we address the problem. State dummies absorb the cross-section variation across states and there is some risk that the state/year variables may lose significance. when we include state dummies in equation (Ib), these variables will lose significance. This is indeed what happens; the variables POPSHARE, GSDPCURR, PERIPHER and NATROAD are no longer significant in (Ic). The explained variation increases slightly to 53.7 per cent – nevertheless indicating that equation (Ia) captures most if the important aspects of inter-state variation. In the "milder test" using clustered errors, the POPSHARE and GSDPCURR variables remain significant; thus supporting their explanatory role.

In (Ib) as well as (Ic), the variables that also vary by product remain significant, with almost identical parameter estimates. Hence equation (Ib) and (Ic) add support to the trade and economic geography explanation, with own production and distance to other suppliers as key determinants.

Turning to equation (II) using relative wholesale prices as the dependent variable, observe that while the prices and the ARRIVAL variable are for the specific market area, the other variables are at the state or state/product level and assigned to each market area based on its location/ state affiliation. While equation (I) covers only three years, we now have data for eight years. There are however mandis only in 23 states/union territories in the data set so the smaller states and union territories drop out.

The explained variation for equation (IIa), without market area dummies, is much lower than for equation (I); only 19 per cent of the variation in wholesale prices is explained by the regression; increasing to 36% when we include market area dummies (equation IIc). For the individual variables, results for the economic geography variables are similar to the results from equation (I); especially the key variables DISTPROD and PRODSHAR are highly significant; however, with somewhat lower estimates than in equations (I).

For the state-level variables, results in equations (II) are mixed; with peripherality (PERIPHER) and state size (POPSHARE) as the only ones that have a consistent pattern. So the price-income link is not confirmed for wholesale prices even if the sign is as expected and it is not too far from being significant.

The magnitude of arrivals is inversely related to the price, but the estimate is no longer significant with clustering or market area dummies. Hence the market area dummies capture most of the variation in this variable.

Diagram 12 shows a similar ranking as above, for the impact of a one standard deviation change for each variable (non-significant variables, P>10% in two equations or more, brighter), using results from (IIa).

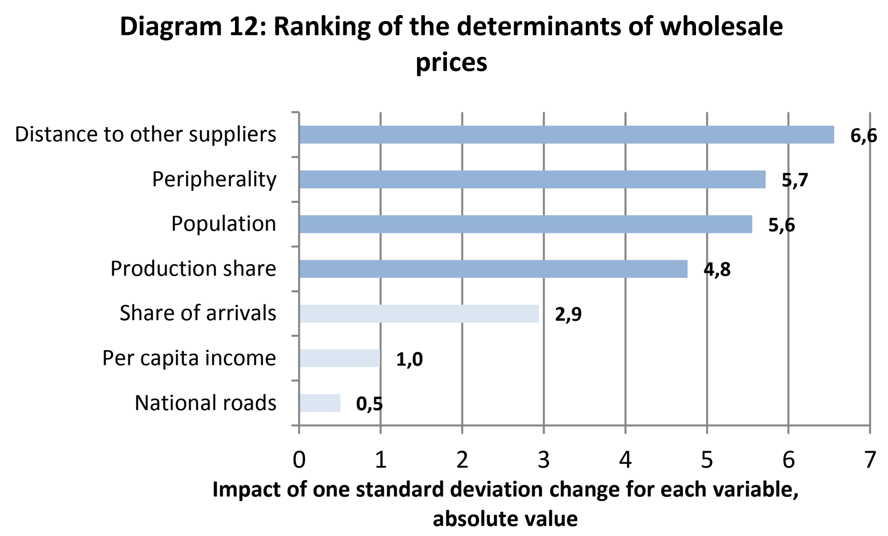


Diagram 12 demonstrates that the supply-side variables (DISTPROD and PRODSHAR) are still key explanations, but peripherality (PERIPHER) and market size (POPSHARE) are also important determinants of wholesale prices. The income-price related explanation fares worse since per capita income does not have a significant impact on wholesale price levels. From a value chain perspective this is perhaps plausible since consumer prices capture costs along the whole value chain (and therefore also reflect costs in the consuming state) whereas wholesale prices are "closer to the farm" with a more limited services component.

We may conclude that the regressions strongly support the "trade and economic geography explanation", with domestic production and distance to supply as key drivers of price differences – most strongly for consumer prices but also for wholesale prices. Explanations related to the income-price link are supported only by equations (I); indicating that income levels are important for consumer prices but to a lesser extent for wholesale prices. A reservation is the impact of state-level variables may change little over time so these variables may matter even if they drop out with clustering or state/market area dummies.

4.3. Regression analysis of changes over time

The preceding analysis has provided new results on the determination of price level differences. Based on these results, price levels could change over time as a result of changes in these underlying variables. We are interested in examining this explicitly and therefore run the same regressions with changes from one year to the next as the variables, still including product and time dummies. We prefer this to panel regressions since we have three levels of aggregation (product; state/mandi; and year) and there is limited variation over time within each cross section (product & state or product & mandi).

In these regressions for the change over time we also include the relative price in the preceding year as an explanatory variable, to check for price convergence. This is an approach similar to the one used in economic growth regressions (see e.g. Durlauf 2009). If there is price convergence, states with high prices initially will have lower price growth so there will be a negative relationship between price change and the initial price level. This is mainly a descriptive measure since the issue of causality is not so clear. This issue is debated for growth regressions (ibid.) and similar concerns apply here.

The results are shown in Table 7. We report simple OLS as well as regressions with clustered errors; clustered with respect to states in equation (III) and market areas in equation (IV).

Table 7: Results from pooled regressions on the determinants of changes in inter-state relative price differences
Note: In the regressions, all variables except "Production share" are expressed in logs. All variables except the initial price are of the form VAR_t-VAR_{t-1} i.e. the change from the preceding year in the data. Statistically significant estimates in both regressions in shading.

Variable	Expected sign	Dependent variable									
		(III) Change in relative consumption price					(IV) Change in relative wholesale price				
		Estimate	OLS		Clustered errors		Estimate	OLS		Clustered errors	
			t stat.	P value	Z stat.	P>|Z1|		t stat.	P value	Z stat.	P>|Z1|
Relative price in preceding year	Negative	-0.362	-16.34	‹.0001	-3.71	0.0002	-0.288	-9.76	‹.0001	-8.99	‹.0001
Prod./cons. ratio	Negative	0.118	2.62	0.0090	0.75	0.4513	n.a.				
Production share	Negative	-0.044	-0.13	0.8990	-0.17	0.8662	0.1241	0.50	0.6194	0.49	0.6228
Distance to other suppliers	Negative	0.085	0.77	0.4438	0.98	0.3290	-0.0286	-0.38	0.7064	-0.47	0.6404
Peripherality	Negative	-1.252	-2.36	0.0187	-1.11	0.2667	-0.4169	-0.59	0.5549	-0.54	0.5879
Per capita income	Positive	-0.084	-1.22	0.2226	-0.95	0.3423	0.1006	1.38	0.1686	0.98	0.3285
Population	Negative	0.384	2.00	0.0464	1.31	0.1913	-2.2203	-4.21	‹.0001	-2.02	0.0438
National roads	Negative	0.058	1.49	0.1376	2.66	0.0077	0.0219	0.28	0.7759	0.36	0.7175
Consumption share	Positive	0.082	3.66	0.0003	1.61	0.1067	n.a.				
Share of mandi arrivals	Negative	n.a.					-0.019	-3.04	0.0024	-2.04	0.0412
Time dummies		Yes									
Product dummies		Yes									
Adjusted R^2			0.300		n.a.			0.154		n.a.	
Observations		704					2336				
Years covered		2004/5, 2009/10, 2011/12					2004/5 to 2011/12				

Note: Date sources are described in Appendix Table 7.

In general, the only robust result is that there is price convergence; with a negative estimate on the initial price. As noted, this is a descriptive result which does not showing any particular causality except it may be interpreted in the light of the somewhat abstract concept of “the law of one price”.

For the variables addressed in the preceding analysis, the results are generally mixed. For the NSSO consumer prices, some variables are significant with OLS, but this no longer the case with clustered errors. For wholesale prices, state size (POPSHARE) and the importance of the market areas (share of mandi arrivals) are significant and with the expected signs both with OLS and clustered errors. Except for this, the explanations that have shown to be key drivers of inter-state differences in price levels have relatively little to contribute concerning the price changes over time.

Hence the results suggest that there was price convergence over time but casts doubt as to whether this change over time was caused by the trade/economic geography or price-income-related explanations. A possible reason is that there was little variation over time for the relevant variables over the relatively few years covered by the analysis. The drivers of change in price differentials over time should therefore be examined further. The strong impact of the lagged price variable could to some extent be a statistical artefact due to price outliers – extremely high prices in one year tend to be followed by lower prices during the next year. As noted, the price convergence regression is no explanation of causality since the lagged price variable is in itself not a cause, except perhaps for the issues of stochastic variation and outliers.

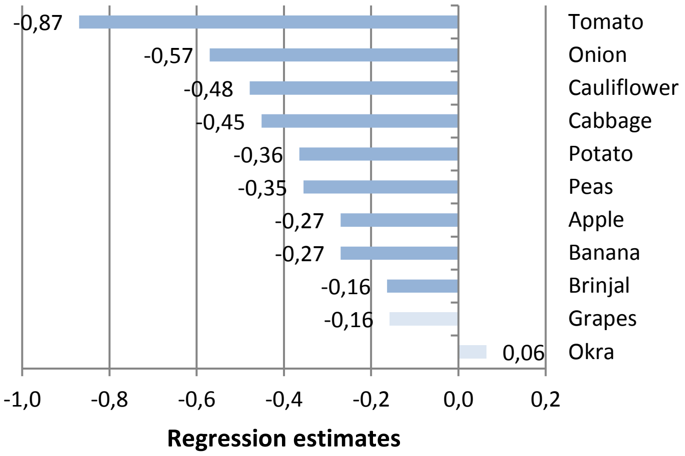
4.4. Heterogeneity across products

The regressions have been undertaken for all products combined, using dummy variables that allow the intercepts to differ across products. There could also be heterogeneity in slopes, and we check this by replicating equation (I) using NSSO consumer prices for each product. The main results are shown in Appendix Table A7 and we provide a brief summary of some aspects here.[23]

The productwise regressions also have high explanatory power, with adjusted R^2 ranging from 0.23 to 0.78 but 9 out of 11 cases above 50%. Estimates on individual variables are statistically significant at the 5% level or better in 46 out of 88 cases (8 variables, 11 products). In line with the former results from (I), the most “persistent” variables were the distance to other suppliers (DISTPROD, significant and negative for 9 of 11 products) and state size (POPSHARE, significant in 8 cases). As an illustration, Diagram 13 shows the variation across products for the variable DISTPROD, the index of distance to other supplying states. As before, non-significant estimates have brighter shading.

[23] Alternatively, we may have interaction terms allowing the DISTPROD estimate in equation (Ia) to vary across products. Results in this case are similar to the results in Diagram 13 and are not reported here. These results can be provided by the author upon request.

Diagram 13: The impact of distance to other suppliers: Estimates for individual products



One might interpret this as an indicator of the importance of inter-state trade frictions and it is perhaps plausible that the easily perishable product tomatoes are on top. On the other hand, we find similar products such as grapes and banana low on the ranking so there is no easy cross-cutting explanation. High weight/price ratios may also be a reason that distance to suppliers is important for products such as onion, cauliflowers, cabbage and potato.

We may conclude that the explanations we have found for inter-state variation in price levels largely survive in regressions at the product level; however, with considerable variation across products in the relative impact of individual variables.

5. Implications and policy issues

The analysis has shed new light on the importance of food price differentials across Indian states and their determinants. There are large and persistent price differentials across Indian states, and this applies to consumer as well as wholesale prices. Consumer or retail prices are on average 50-60 per cent higher than wholesale prices. Consumer prices capture the whole food value chain whereas wholesale prices represent an intermediate level of the value chain. It is therefore likely that the cost level of the consuming state affects consumer prices more strongly than wholesale prices. This expectation is largely confirmed by the econometric analysis, since per capita income is a determinant of consumer but not wholesale prices. We do however not have data at hand to explain what sub-explanation is underlying this price-income link.

The strongest driver of inter-state price differences is what we have called the "trade and economic geography explanation". The econometric analysis has confirmed that states with larger domestic production or better roads have lower prices, whereas states with a larger average distance to other supplying states have higher prices. The latter variable, our index DISTPROD, is the variable having the strongest impact on inter-state price differences. We interpret this as saying that there is a considerable "cost of distance" in India's food markets and that the cost of inter-state trade is a key driver of food price gaps.

The descriptive analysis of prices based on the NSSO household survey data as well as the later regression analysis suggests that there has been a decline in inter-state price differences over time. This is supported by regression analysis but we have not been able to map the exact drivers of change over time. A possibility is that better inter-state integration has facilitated trade and stimulated price convergence, but we have not been able to prove that this is the case. Our study covers a limited time span and this may be a reason why it has been difficult to obtain very clear results on the determinants of change.

In the analysis, we have focused on the "price-income link" and the "trade and economic geography explanation" of price differences. There could however also be other drivers such as changes in competition due to institutional reforms that affect price formation. For example, it would be of interest to examine whether inter-state differences in the implementation of the 2003 Agriculture Produce Marketing Act affect prices. We did not have such data at hand for our purpose here, and it is an aspect that should be taken into account in later work.

The analysis shows that inter-regional price differences ate the product level are large and therefore of key importance for a large nation like India, where food represents 40-50 percent of private consumption and there is large heterogeneity across states. How large is this India's interstate food price dispersion in a global or international perspective? While a comparative study is beyond the scope of this paper, some illustrations may be useful and we will therefore add some evidence. Recall from section 2 that the Gini for average food price levels across states was at 0.052. In this light, it is of interest that for aggregate food prices in the EU-28 in 2014, we obtain a population-weighted Gini of 0.08, i.e. slightly higher than for India.[24] Based on ICP/World Bank data (International Comparison Programme, icp.worldbank.org), we obtain a population-weighted Gini for food price differences across the whole world at 0.197; i.e. much higher than within India. An even more relevant comparison is obtained if we select the 64 countries around the world that are in the same income span as India's states in 2011/12. In this case we obtain a Gini coefficient for food price dispersion at 0.147; i.e. also considerably higher than within India. Hence this evidence suggests that even if food price differences within India are high, Indian states are much better integrated than countries within the same income range, and slightly more integrated than EU countries.[25]

For policy, an issue is to what extent price dispersion is bad and should be eliminated. The answer is generally affirmative; with the reservation that price differences are persistent and do not disappear even if tariffs and other formal trade barriers are removed. In this respect it is useful to distinguish between the trade and economic geography drivers of price gaps, and the drivers related to the income-price correlation. The income-related drivers do not disappear with development; on the contrary they may become more important; e.g. it may be the case that logistics and services components of the final price could increase over time. One should not get rid of the middlemen but strive for efficient trade and distribution services, also creating good jobs with decent pay. For the trade and economic geography drivers of price gaps, it is more unambiguously the case that costs should be reduced and trade should be facilitated, even if distance will not disappear and trade costs will remain even when the policy-affected barriers have been eliminated. It is interesting that our spatial variable on proximity to other suppliers turned out to be the strongest driver of inter-state price differences. An aim for India should therefore be to promote further intra-Indian trade integration, raise efficiency in trading and distribution, and thereby

---

[24] Based on data from Eurostat on prices and population for EU countries in 2014, retrieved 12 January 2016. Price data for "Food and non-alcoholic beverages" has been used.

[25] For the calculations, international population and income level data (GDP per capita at current prices) were taken from the World Development Indicators (WDI). The income span of Indian regions was defined from the data on GSDP fom NITI Ayyog, se Appendix Table A6.

lower the prices of goods imported from other states. This is also an aim that has been a key factor in the European internal market (Goldberg and Verboven 2005).

Even from the trade cost perspective, it is not certain that all prices will be reduced because of lower trade costs. For the supplying states, better trade integration may actually increase prices, when they no longer have to sell all their produce locally due to trade impediments. Hence lower trade barriers facilitate exports to other states, so the domestic prices in supplying states may rise. This is good news for the producers and for the supplying states at large, even if there is – as always – an income distribution component and some consumers may face higher prices.

As noted, India has the lowest food price level in the world. With growth and development, India should be prepared to lose this lead, since food prices may rise along with the development of a stronger food distribution system providing new jobs beyond agriculture. India should promote trade and efficiency and work to keep food prices down, but at the same time expect that they actually may rise in the process of development.

For policy, an issue is also about "statistics policy" – what efforts should India make to supply statewise data on a systematic and regular basis? While the availability of such data has increased greatly and should be commended, there are still many gaps and more to be done. In particular, for large emerging nations such as India, it is hard to measure and analyse regional development appropriately without better data on regional price level differences. The current practice of setting all prices equal to 100 every now and then, with price indexes diverging until the next change of basis, could be supplemented with a more systematic supply of statewise price level data. For research on price differences, India is an interesting case and with a growing supply of data, new knowledge may then be provided in the future.

In the article, we have covered several aspects but certainly not all. We did not have data to explore what is underlying the income-price relationship observed for consumer prices. The analysis of determinants has covered some products but far from all, and more research is needed to say whether our results are general. We have not addressed urban-rural differences, which are shown by Deaton and Dupriez (2011) to matter. As noted, issues of competition could also affect prices and should be taken into account. For some products, international trade plays an important role that has not been addressed. Our analysis of determinants has covered only some products and could be extended to more. The mandi price data is a rich source of information that could be used for analysis of the nature of markets. Hence research on inter-state food price variation in India should continue and shed light on these and other unresolved issues.

References

Alessandria, G. and J. P. Kaboski, 2011, Pricing-to-Market and the Failure of Absolute PPP, *American Economic Journal: Macroeconomics* (3): 91–127.

Atkin, D., 2013, Trade, Tastes, and Nutrition in India, *American Economic Review* 103(5): 1629-63.

Balassa, B., 1964, The Purchasing-Power Parity Doctrine: A Reappraisal, *Journal of Political Economy* 72: 584-596.

Bekker, E., J. F. Francois and M. Manchin, 2012, Import Prices, Income, and Inequality, Vienna: Vienna Institute for International Economic Studies, WIIW Working Paper No. 82, June 2012.

Brandt, L. and C. A. Holz, 2006, Spatial Price Differences in China: Estimates and Implications. *Economic Development and Cultural Change* 55(1): 43-86.

Chand, R., 2006, Agricultural Markets in India: Implications for Competition. Chapter 11, pp. 139-151 in Pradeep S. Mehta (ed.). Towards a Functional Competition Policy for India. Delhi/Jaipur: Academic Foundation/CUTS (Consumer Unity and Trust Society). 2006

CSO, 2013, Statewise estimates of value of output from agriculture and allied activities with new base year 2004-2005 (2004-05 to 2010-11). New Delhi: Central Statistics Office, Ministry of Statistics and Programme Implementation, Government of India, available at www.mospi.gov.in.

Deaton, A., 2003, Prices and Poverty in India, 1987-2000. *Economic and Political Weekly*, January 25, 2003: 362-368.

Deaton, A. and O. Dupriez, 2011, Spatial Price Differences within Large Countries, Working Paper No. 1321, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies, Princeton University, Princeton NJ, and World Bank, Washington DC.

Durlauf, S. N., 2009, The Rise and Fall of Cross-Country Growth Regressions, History of Political Economy 41 (annual suppl.): 315-333. DOI 10.1215/00182702-2009-030.

Feenstra, R. and J. Romalis, 2014, International Prices and Endogenous Quality, *The Quarterly Journal of Economics* 129(2): 477–527. doi:10.1093/qje/qju001.

Gluschenko, K. and A. Khimich, 2007, Food Markets in Russia: Dynamics of Their Integration. Vienna: Institute for Advanced Studies, Economics Series No. 221, November 2007.

Goldberg, P. K. and F. Verboven., 2005, Market Integration And Convergence To The Law Of One Price: Evidence From The European Car Market, *Journal of International Economics* 65(1): 49-73.

Gong, C. H. and X. Meng, 2008, Regional Price Differences in Urban China 1986–2001: Estimation and Implication. Bonn: The Institute for the Study of Labor (IZA), IZA Discussion Paper No. 3621, July 2008.

Gorodnichenko, Y. and L. L. Tesar, 2009, Border Effect or Country Effect? Seattle May Not Be so Far from Vancouver After All. *American Economic Journal: Macroeconomics* (1): 219–241.

Gulati, A. & S. Saini, 2013, Taming Food Inflation in India. New Delhi: Ministry of Agriculture, Commission for Agricultural Costs and Prices, Discussion Paper No. 4.

Hajzler, C. and J. MacGee, 2014, Retail Price Differences across U.S. and Canadian Cities during the Interwar Period. Mimeo, Bank of Canada.

Hallak, J. C. and P. K. Schott, 2011, Estimating Cross-Country Differences in Product Quality, *Quarterly Journal of Economics* 126: 417–474.

Hallak, J. C., 2006, Product Quality and the Direction of Trade, *Journal of International Economics* 68: 238–265.

Head, K. og T. Mayer (2014). Gravity Equations: Workhorse, Toolkit, and Cookbook. I: G. Gopinath, E. Helpman og K. Rogoff (red.), *Handbook of International Economics,* Volume 4, kap. 3, s. 131–195. Amsterdam: Elsevier.

Li, C. And J. Gibson, 2014, Spatial Price Differences and Inequality in the People's Republic of China: Housing Market Evidence, *Asian Development Review* 31(1): 92–120.

Melchior, A. and K. Telle, 2001, Global Income Distribution 1965-98: Convergence and Marginalisation. *Forum for Development Studies* 1/2001.

Melchior, A., 2008, Regional Inequality and Convergence in Europe, 1995-2005, Warsaw: CASE Studies and Analyses No. 374/ 2008.

Melchior, A., 2010, Globalisation, Domestic Market Integration, and the Regional Disparities of India, Oslo: Norwegian Institute of International Affairs, Working Paper No. 780, available at www.nupi.no.

Milanovic, B., 2005, Worlds Apart. Measuring International and Global Inequality. Princeton/Oxford: Princeton University Press.

Ministry of Agriculture, 2016, Horticulture Statistice at a Glance 2015. New Delhi: Oxford University Press/ Ministry of Agriculture & Farmers Welfare, Government of India.

NHB, various years, Indian Horticulture Database. Gurgaon: National Horticulture Board, Ministry of Agriculture, Government of India. Editions for 2008, 2011 and 2014 were used.

NSSO, 2007, Household Consumption of Various Goods and Services in India, 2004-05. NSS 61st Round (July 2004 - June 2005) Vol. I: Major States and All-India and Vol. II: Minor States and Union Territories (Detailed Tables 1-4 only). New Delhi: National Sample Survey Organisation, Ministry of Statistics and Programme Implementation, Government of India, NSS Report No. 509(61/1.0/2), April 2007, available on mospi.nic.in.

NSSO, 2012, Household Consumption of Various Goods and Services in India, 2004-05. NSS 66st Round (July 2009 - June 2010). New Delhi: National Sample Survey Organisation, Ministry of Statistics and Programme Implementation, Government of India, NSS Report No. 541(66/1.0/3), February 2012, available on mospi.nic.in.

NSSO, 2013, Key Indicators of Household Consumer Expenditure in India. NSS 66st Round (July 2010 - June 2011). New Delhi: National Sample Survey Organisation, Ministry of Statistics and Programme Implementation, Government of India, NSS Report KI (68/1.0), June 2013, available on mospi.nic.in.

NSSO, 2014a, Level and Pattern of Consumer Expenditure 2011-12. NSS 68th Round (July 2011 - June 2012). New Delhi: National Sample Survey Organisation, Ministry of Statistics and Programme Implementation, Government of India, NSS Report No. 555 (68/1.0/1), February 2014, available on mospi.nic.in.

NSSO, 2014b, Level and Pattern of Consumer Expenditure 2011-12. NSS 68th Round (July 2011 - June 2012). New Delhi: National Sample Survey Organisation, Ministry of Statistics and Programme Implementation, Government of India, NSS Report No. 558 (68/1.0/2), June 2014, available on mospi.nic.in.

O'Connell, P. G. and S-J. Wei, 1997, "The Bigger They Are, The Harder They Fall": How Price Differences Across U.S. Cities Are Arbitraged. NBER Working Paper No. 6089.

Samuelson, P.A., 1964, Theoretical Notes on Trade Problems, *The Review of Economics and Statistics* 46(2): 145-54.

Timmer, , M. P., A. A. Erumban, B. Los, R. Stehrer and G. J. de Vries, 2014, Slicing Up Global Value Chains, *Journal of Economic Perspectives* 28(2): 99–118.

Appendix Table A1: Population-weighted Gini coefficients for food price differences across Indian states and union territories

NSSO code			Description	Gini coefficients		
2004-2005	2009-2010	2011-2012		2004-2005	2009-2010	2011-2012
101	101	101	rice - P.D.S.	0.171	0.337	0.323
102	102	102	rice - other sources	0.096	0.108	0.100
103	103	103	chira	0.109	0.094	0.093
104	104	104	khoi, lawa	0.264	0.143	0.146
105	105	105	muri	0.149	0.089	0.152
106	106	106	other rice products	0.235	0.107	0.122
107	107	107	wheat/atta - P.D.S.	0.227	0.185	0.277
108	108	108	wheat/atta - other sources	0.149	0.145	0.180
110	110	110	maida	0.090	0.076	0.087
111	111	111	suji, rawa	0.065	0.058	0.080
112	112	112	sewai, noodles	0.169	0.137	0.109
113	113	113	bread (bakery)	0.065	0.064	0.080
114	114	114	other wheat products	0.248	0.282	0.224
115	115	115	jowar & products	0.174	0.121	0.247
116	116	116	bajra & products	0.132	0.140	0.176
117	117	117	maize & products	0.196	0.155	0.222
118	118	118	barley & products	0.267	0.211	0.397
120	120	120	small millets & products	0.360	0.323	0.338
121	121	121	ragi & products	0.182	0.114	0.120
122	122	122	other cereals	0.317	0.277	0.229
139	139	139	cereal substitutes	0.346	0.309	0.497
140	140	140	arhar, tur	0.027	0.027	0.041
141	141	141	gram (split)	0.053	0.055	0.044
142	142	142	gram (whole)	0.085	0.087	0.083
143	143	143	moong	0.063	0.049	0.062
144	144	144	masur	0.067	0.050	0.063
145	145	145	urd	0.078	0.075	0.080
146	146	146	peas	0.115	0.114	0.107
148	147	147	khesari	0.177	0.172	0.177
150	148	148	other pulses	0.130	0.091	0.120
151	150	150	gram products	0.174	0.099	0.107
152	151	151	besan	0.059	0.043	0.053
153	152	152	other pulse products	0.159	0.109	0.095
160	160	160	milk: liquid (litre)	0.057	0.056	0.071
161	161	161	baby food	0.400	0.194	0.118
162	162	162	milk: condensed/powder	0.284	0.124	0.091
163	163	163	curd	0.183	0.160	0.149
164	164	164	ghee	0.435	0.048	0.046

NSSO code			Description	Gini coefficients		
2004-2005	2009-2010	2011-2012		2004-2005	2009-2010	2011-2012
165	165	165	butter	0.262	0.217	0.177
279	189	170	salt	0.105	0.065	0.064
260	170	171	sugar - P.D.S.	0.057	0.057	0.136
261	171	172	sugar - other sources	0.020	0.017	0.021
262	172	173	gur	0.086	0.042	0.063
263	173	174	candy, misri	0.305	0.351	0.141
264	174	175	honey	0.393	0.216	0.207
170	190	180	vanaspati/ margarine	0.053	0.087	0.058
171	191	181	mustard oil	0.045	0.047	0.042
172	192	182	groundnut oil	0.046	0.068	0.062
173	193	183	coconut oil	0.124	0.181	0.168
174	194	185	edible oil (others)	0.088	0.057	0.134
180	200	190	eggs (no.)	0.073	0.035	0.061
181	201	191	fish, prawn	0.128	0.099	0.106
182	202	192	goat meat/mutton	0.097	0.085	0.090
183	203	193	beef/ buffalo meat	0.178	0.112	0.116
184	204	194	pork	0.068	0.168	0.110
185	205	195	chicken	0.068	0.087	0.055
186	206	196	others (birds, crab etc.)	0.271	0.152	0.265
190	210	200	potato	0.128	0.128	0.152
191	211	201	onion	0.092	0.069	0.075
214	234	202	tomato	0.093	0.086	0.067
210	230	203	brinjal	0.128	0.122	0.102
192	212	204	radish	0.227	0.231	0.220
193	213	205	carrot	0.239	0.204	0.193
212	232	206	palak/other leafy vegs.	0.200	0.352	0.177
216	236	207	chillis (green)	0.188	0.104	0.122
211	231	208	lady's finger	0.163	0.138	0.121
207	227	211	cauliflower	0.223	0.187	0.185
208	228	212	cabbage	0.200	0.157	0.170
215	235	214	peas	0.225	0.135	0.186
213	233	215	french beans and barbati	0.122	0.150	0.115
221	241	216	lemon (no.)	0.109	0.112	0.103
224	242	217	other vegetables	0.328	0.143	0.142
230	250	220	banana (no.)	0.125	0.175	0.122
231	251	221	jackfruit	0.449	0.457	0.221
232	252	222	watermelon	0.190	0.197	0.180
233	253	223	pineapple (no.)	0.268	0.215	0.176
234	254	224	coconut (no.)	0.371	0.178	0.170
297	255	225	coconut: green (no.)	0.361	0.131	0.126
235	256	226	guava	0.453	0.107	0.122

NSSO code			Description	Gini coefficients		
2004-2005	2009-2010	2011-2012		2004-2005	2009-2010	2011-2012
236	257	227	singara	0.298	0.259	0.276
237	258	228	orange, mausami (no.)	0.121	0.162	0.118
238	260	230	papaya	0.161	0.162	0.088
240	261	231	mango	0.142	0.162	0.142
241	262	232	kharbooza	0.346	0.219	0.243
242	263	233	pears (naspati)	0.323	0.218	0.272
243	264	234	berries	0.253	0.222	0.275
244	265	235	leechi	0.267	0.228	0.210
245	266	236	apple	0.138	0.102	0.100
246	267	237	grapes	0.297	0.078	0.098
250	270	240	coconut (copra)	0.488	0.233	0.160
251	271	241	groundnut	0.290	0.097	0.092
252	272	242	dates	0.360	0.139	0.151
253	273	243	cashewnut	0.467	0.093	0.076
254	274	244	walnut	0.551	0.193	0.290
255	275	245	other nuts	0.394	0.209	0.152
256	276	246	raisin, kishmish etc.	0.487	0.091	0.083
257	277	247	other dry fruits	0.499	0.220	0.169
223	281	250	ginger (gm)	0.059	0.062	0.069
222	280	251	garlic (gm)	0.116	0.055	0.065
280	282	254	turmeric (gm)	0.088	0.041	0.070
281	283	255	black pepper (gm)	0.089	0.057	0.068
282	284	256	dry chillies (gm)	0.127	0.076	0.074
283	285	257	tamarind (gm)	0.228	0.199	0.117
284	286	258	curry powder (gm)	0.117	0.114	0.113
285	287	260	oilseeds (gm)	0.195	0.167	0.143
286	288	261	other spices (gm)	0.108	0.089	0.107
290	290	270	tea: cups (no.)	0.169	0.150	0.157
291	291	271	tea leaf (gm)	0.077	0.064	0.079
292	292	272	coffee: cups (no.)	0.515	0.227	0.210
293	293	273	coffee: powder (gm)	0.314	0.239	0.279
295	295	275	cold beverages	0.204	0.095	0.114
296	296	276	fruit juice and shake	0.229	0.113	0.080
305	305	294	pickles (gm)	0.128	0.097	0.086

Source: Own calculations based on NSSO (2007, 2012, 2014b).

Appendix Table A2: Indicators on relative price levels for food across Indian states, 2011/12.
Source: Own calculations based on data from NSSO (2014b).

State/ territory	Average of relative prices	Rank for average of relative prices (1=lowest, 35=highest)	Rank based on fixed effect from regression	Difference between the two rankings
A & N Island	1.58	35	35***	0
Andhra Pradesh	1.04	9	18	-9
Arunachal Pradesh	1.23	26	13	13
Assam	1.09	13	15	-2
Bihar	0.90	1	3**	-2
Chandigarh	1.17	20	11	9
Chhattisgarh	0.95	4	4**	0
Dadra & N. Haveli	1.24	27	9	18
Daman & Diu	1.22	24	8	16
Delhi	1.28	29	33***	-4
Goa	1.29	30	28**	2
Gujarat	1.11	14	30**	-16
Haryana	1.13	17	26*	-9
Himachal Pradesh	1.15	19	23	-4
Jammu & Kashmir	1.15	18	22	-4
Jharkhand	0.99	6	7*	-1
Karnataka	1.03	8	20	-12
Kerala	1.11	16	27*	-11
Lakshadweep	1.33	32	10	22
Madhya Pradesh	0.91	2	6**	-4
Maharashtra	1.17	21	34***	-13
Manipur	1.23	25	14	11
Meghalaya	1.27	28	21	7
Mizoram	1.30	31	17	14
Nagaland	1.43	34	32***	2
Odisha	0.93	3	2***	1
Puducherry	1.20	23	19	4
Punjab	1.18	22	24	-2
Rajasthan	1.04	10	16	-6
Sikkim	1.35	33	25	8
Tamil Nadu	1.11	15	31***	-16
Tripura	1.08	12	5**	7
Uttar Pradesh	0.96	5	1***	4
Uttarakhand	1.07	11	29**	-18
West Bengal	1.01	7	12***	-5

Note: A panel regression of the form $p_{ik}=f(v_{ik})$ was run, where $p_{ik}$is the relative price of state I for product k, and v_{ik} is the value of consumption in state i for product k. In the regression, the variables were expressed in logs, with fixed effects for the states. Including a constant term, one of the fixed effects is automatically dropped (West Bengal, which is represented by the constant term). In the table */**/*** indicates whether the fixed effect is significantly different from zero at the 10/5/1 per cent level. For the variable v_{ik}, the coefficient estimate was -0.0386; highly significant with t value at -14.32 and P value <0.0001. The constant term estimate was 0.5475, with t value at 11.04 and p value $<.0001$. For the regression, R^2 was 0.20, and the number of observations 3053.

Appendix Table A3: Products selected for analysing wholesale prices and the determinants of price differences

NSSO cat.	Product in NHB data	% of food consumption in India	Gini for NSSO interstate price variation
200	Potato -fresh	2.14	0.152
	Potato - store		
201	Onion	1.40	0.075
202	Tomato - hybrid	1.23	0.067
	Tomato - local		
203	Brinjal - long	0.79	0.102
	Brinjal - round		
208	Lady's finger (okra)	0.61	0.121
211	Cauliflower	0.54	0.185
212	Cabbage	0.39	0.170
214	Peas	0.25	0.186
220	Banana(no.)	1.18	0.122
236	Apple – Jammu & Kashmir	0.99	0.100
	Apple – Royal Del. (H.P.)		
237	Grapes	0.32	0.098
	Sum	9.84	

Note: Consumption shares and Ginis were calculated from NSSO data for 2011/12.

Appendix Table A4: Market areas covered by mandi price data from the National Horticulture Board

State	Market area
Andhra Pradesh/Telangana	Hyderabad
Assam	Guwahati
Bihar	Patna
Chandigarh	Chandigarh
Chhattisgarh	Raipur
Delhi	Delhi
Gujarat	Ahmedabad
Gujarat	Surat
Himachal Pradesh	Shimla
Jammu & Kashmir	Jammu
	Srinagar
Jharkhand	Ranchi
Karnataka	Bangalore
Kerala	Trivendrum
Madhya Pradesh	Bhopal
Maharashtra	Mumbai
	Nagpur
	Nasik
	Pune
Orissa	Bhubaneshwar
Punjab	Abohar
	Amritsar
Rajasthan	Jaipur
Sikkim	Gangatok
Tamil Nadu	Chennai
Uttar Pradesh	Agra
	Baraut
	Kanpur
	Lucknow
Uttarakhand	Dehradun
West Bengal	Kolkata

Note: The market areas of Lasalgaon and Pimpalgaon were in the data but were dropped due to a low number of observations.

Appendix Table A5: Price levels in major market areas, 2004-2014. Based on price data from NHB.

Market	Average relative price	Rank (1=low) according to: Relative price average	Rank (1=low) according to: Regression fixed effects	Difference between rankings	Number of significant fixed effects (max.=10)
KANPUR	0.74339	1	1	0	9
AMRITSAR	0.74907	2	2	0	10
NAGPUR	0.81958	3	5	-2	10
LUCKNOW	0.82420	4	4	0	9
JAIPUR	0.82677	5	3	2	10
AGRA	0.83122	6	6	0	10
BHOPAL	0.86098	7	7	0	10
DEHRADUN	0.86665	8	9	-1	10
CHANDIGARH	0.87620	9	8	1	10
RANCHI	0.90473	10	10	0	9
AHMEDABAD	0.93820	11	11	0	10
PATNA	0.94058	12	12	0	10
RAIPUR	0.94894	13	19	-6	10
JAMMU	0.97490	14	16	-2	10
DELHI	0.97868	15	13	2	10
BHUBANESHWAR	0.99927	16	14	2	9
HYDERABAD	1.00073	17	17	0	9
BANGALORE	1.04301	18	20	-2	9
ABOHAR	1.05625	19	18	1	6
PUNE	1.06904	20	22	-2	8
MUMBAI	1.08456	21	15	6	7
SURAT	1.10823	22	21	1	7
KOLKATA	1.11202	23	23	0	8
SRINAGAR	1.13280	24	24	0	8
BARAUT	1.14754	25	26	-1	7
GUWAHATI	1.15860	26	25	1	7
CHENNAI	1.20251	27	27	0	8
SHIMLA	1.24809	28	28	0	5
GANGATOK	1.46123	29	29	0	0
TRIVENDRUM	1.52424	30	30	0	4
Markets not in regressions/ranking					
LASALGAON	1.03900	17-18			
NASIK	1.08628	21-22			
PIMPALGAON	0.84411	6-7			

Appendix Table A6: Data sources and years covered
Note: x = covered by data; (x) = covered by data but incomplete

Variable	Data source	Years 2004/5	2005/6	2007/8	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15
Per capita value and volume of consumption by state	NSSO (2007, 2012, 2014)	x			x		x			
Wholesale and retail prices for 11/15 products	NHB (www.nhb.gov.in)	x	x	x	x	x	x	x	x	(x)
Value of production	CSO (2013); Ministry of Agriculture (2016)	x	x	x	x	x	x			
Volume of production	NHB (2008, 2011, 2013), Ministry of Agriculture (2016); www.nhb.gov.in.	(x)	(x)	(x)	(x)	(x)	x	x	x	x
Per capita value and volume of consumption by state	NSSO (2007, 2012, 2014)	x			x		x			
Population by state, rural and urban	NSSO (2007, 2012, 2014) and NITI Ayyog	x			x		x			
Population by state, total	Niti Ayyog (National Institution for Transforming India); niti.gov.ind	x	x	x	x	x	x	x	x	(x)
Per capita income		x	x	x	x	x	x	x	x	(x)
Length of national highways per sq. km.		x	x	x	x	x	x	x		
Geographical data	Global Cities Database; http://www.diva-gis.org/Data; and web sources such as www.gadm.org (on administrative areas, from University of California Berkeley)	No time variation								

Appendix Table A7: Productwise regressions on the determinants of inter-state food price differences in India.
Note: The regressions are similar to regression (Ia) in the main text, except that product dummies are not included.

		Apple	Banana	Brinjal	Cabbage	Cauliflower	Grapes	Okra	Onion	Peas	Potato	Tomato
Consumption share	Estimate	0.144	0.226	0.084	0.160	0.012	-0.025	0.045	-0.142	-0.034	-0.036	0.109
	P value	0.307	0.001	0.003	0.003	0.736	0.851	0.341	0.029	0.442	0.400	0.027
Peripherality	Estimate	-0.117	0.210	-0.255	0.018	-0.287	-0.303	-0.230	-0.058	0.173	-0.070	0.026
	P value	0.128	0.015	<.0001	0.827	0.000	0.008	0.000	0.298	0.195	0.296	0.681
Distance to other suppliers	Estimate	-0.270	-0.270	-0.164	-0.451	-0.479	-0.157	0.064	-0.570	-0.355	-0.365	-0.869
	P value	<.0001	0.001	0.040	0.000	0.000	0.179	0.546	<.0001	0.003	<.0001	<.0001
Per capita income	Estimate	0.083	0.264	0.300	0.125	0.170	0.017	0.318	0.072	0.134	-0.032	-0.028
	P value	0.373	<.0001	<.0001	0.050	0.007	0.908	<.0001	0.033	0.024	0.323	0.509
Population	Estimate	-0.024	-0.009	-0.046	-0.042	-0.069	-0.095	-0.037	-0.032	-0.082	-0.056	-0.081
	P value	0.074	0.669	<.0001	0.066	0.000	0.001	0.011	0.019	<.0001	<.0001	<.0001
Prod./cons. ratio	Estimate	0.011	-0.155	-0.420	-0.340	-0.405	-0.207	-0.243	0.514	-0.078	-0.271	-0.169
	P value	0.860	0.293	0.016	0.013	0.010	0.438	0.107	0.051	0.205	0.227	0.136
Production share	Estimate	-1.874	-1.285	0.459	-1.331	-0.897	-0.112	0.047	-1.766	-0.850	-1.109	-1.381
	P value	<.0001	0.023	0.316	0.003	0.085	0.762	0.936	<.0001	0.013	<.0001	0.002
National roads	Estimate	-0.023	-0.114	-0.053	-0.046	-0.026	-0.020	-0.086	0.037	-0.061	0.008	-0.013
	P value	0.392	0.009	0.011	0.162	0.468	0.682	0.007	0.055	0.120	0.733	0.625
Adjusted R^2		0.506	0.484	0.777	0.699	0.752	0.236	0.508	0.784	0.561	0.753	0.713
Observations		96	96	96	96	96	96	96	96	96	96	96

Note: P values are corrected for heteroskedasticity.

Norwegian Institute of International Affairs

Established in 1959, the Norwegian Institute of International Affairs [NUPI] is a leading independent research institute on international politics and areas of relevance to Norwegian foreign policy. Formally under the Ministry of Education and Research, NUPI nevertheless operates as an independent, non-political instance in all its professional activities. Research undertaken at NUPI ranges from short-term applied research to more long-term basic research.

About the Author

Arne Melchior is Senior Research Fellow and Head of the Research group on internatinal economics at the Norwegian Institute of International Affairs (NUPI).

NUPI
Norwegian Institute of International Affairs
C.J. Hambros plass 2D
PO Box 8159 Dep. NO-0033 Oslo, Norway
www.nupi.no | info@nupi.no