Marketization Costs and Household Specialization: Evidence from Indian Farmers

Nicholas Li*

University of Toronto

October 4, 2018 – Version 1.1 PRELIMINARY

Abstract

Farmers in developing countries produce a large fraction of the goods they consume. "Marketization costs" that drive a wedge between the value of producing for home consumption versus market-based sales and purchases provide one explanation. By linking production and consumption decisions, these costs can affect the gains from specialization through occupational choice and crop choice. I test this explanation using the expansion of India's Public Distribution System in the 1990s and 2000s as a natural experiment that lowers the incentive to produce staples for home consumption. The results confirm the quantitative importance of marketization costs as a source of production misallocation, and provide novel evidence that in-kind transfers can have substantial production-side effects.

*Thanks to Mu-Jeung Yang, Dan Trefler, Loren Brandt, Xiaodong Zhu, Diego Restuccia, Gustavo Bobonis, David Atkin, Andres Rodriguez-Clare, Kala Krishna, Lorenzo Caliendo, Trevor Tombe, Palermo Penano, and seminar participants at the University of Toronto, University of Calgary, and Stanford CGID 2014 conferences. Jan Victor Dee and David Walker-Jones provided outstanding research assistance. All errors are my own. Email: nick.li@utoronto.ca

1. Introduction

While virtually all households engage in some non-market production of services, many do not produce any of the goods they consume. Farmers, particularly in the developing world, are an important exception. Household consumption surveys – collected routinely by national statistical agencies to measure nutrition and poverty or construct expenditure shares for consumer price indexes – often collect data on home produced goods in addition to goods purchased from the market. Different survey methods, reporting periods, classification of goods, and methods of imputing the value of home produced goods make international comparisons difficult, but they demonstrate a clear pattern. Table 1 presents data from the World Bank LSMS for Uganda, Timor, Peru, Guatemala and India, ordered from lowest to highest GDP per capita. As countries get richer, the fraction of food consumed out of home production falls for the average household, value added per worker in agriculture rises, and the fraction of households that own or lease-in agricultural land declines. Looking within the subset of "farming" households that consume at least one percent of food out of home production, the home share tends to decline as countries get richer. Although the number of food categories varies across surveys, the home share of food expenditures is generally higher than the home share of food varieties, suggesting that the goods that are produced for home consumption tend to be staples with large expenditure shares.

In this paper I examine the link between farming and home consumption and the implications for occupational choice and crop specialization. Agricultural production has certain technological features (geographically dispersed production with high trade costs) and preference features (low income elasticity) that suggest an especially important role for trade costs in influencing decisions about whether to farm, and what to farm, for poor households. Recent work in trade has highlighted the importance of trade costs and comparative advantage for agricultural productivity within countries (Donaldson (2012), Costinot and Donaldson (2014), Sotelo (2014)) and across countries (Tombe (2014), Swiecki (2014)). In this paper I consider the role of "marketization costs" – broadly interpreted as the relative cost of acquiring a good through own-production versus market exchange – in agricultural specialization decisions for households (Fafchamps (2012)). In the presence of high marketization costs, the consumption advantage offered by subsistence staple production may distort household labor towards farming and distorts crop choices away from crops with higher market value towards those with higher home consumption value. The "food problem" identified by Schultz (1953) linking low productivity and low incomes through production/consumption non-separabilities may thus operate at a household scale and not just at the regional or country level.

I begin by describing a simple Ricardian model to motivate the empirical analysis. The model captures the role of marketization cost in decisions about whether or not to farm and whether to produce staples or other marketed agricultural goods (e.g. cash crops and those with limited home consumption value). While recent work in the macro development literature has highlighted wage gaps across sectors (Restuccia et al. (2008), Adamopolous (2011), Lagakos and Waugh (2013), Gollin and Rogerson (2014), Gollin et al. (2014)) the potential for a direct consumption advantage from farming has not received much focus, with the consumption side often modeled by a representative agent that lumps together farmers with agricultural laborers and/or other rural households. This ignores the fact that it is farmers that make the actual cropping decisions and that often consume a large share of their crop output, and neglects the large micro development literature on non-separable household models (see de Janvry et al. (1991) and de Janvry and Sadoulet (2006) for a recent review).

I next lay out some descriptive facts for India, taking advantage of the fact that the National Sample Survey (NSS) records home consumption shares and can thus be used to examine the extent to which individual households are integrated into the market economy. I focus on the period from 1993 to 2010, during which the aggregate home share of consumption declined from 14% to 7%. This decline is accompanied both by an exit from farming and by growing specialization by farmers. These household-level changes are also reflected in modest increases in crop specialization at the district-level. Turning to prices, I show that there are large differences between market and harvest or farm gate prices for major agricultural commodities that imply a potentially important role for marketization costs. However, these gaps have been fairly stable over time, and the price gaps across markets that would be expected to shrink if trade costs fell appear to be quite stable. The main development in this period is the reform and expansion of the Public Distribution System (PDS) over the 1990s and 2000s, resulting in an approximate doubling of the quantity of rice and wheat purchased through the PDS per capita and doubling of the implicit subsidy value of the PDS.

I use variation in the generosity of India's PDS subsidy during the 1993-2009 period as a natural experiment that reduces marketization costs for rice and wheat. By providing poor households cheaper and easier access to these staples, the PDS could potentially lead to significant changes in household decisions about whether and what to farm under non-separability of production and consumption. This exercise is thus a test of the existence and quantitative importance of non-separabilities on a national scale. My analysis also provides the first evidence on production-side effects of in-kind transfer/subsidy programs to my knowledge, as the literature has focused on nutrition, poverty, and consumption effects (Tarozzi (2005),Dreze and Khera (2013), Kochar (2005), Krishnamurthy et al. (2014), Kaushal and Muchomba (2015), Basu and Das (2014)).[1] I find that the PDS expansion in India has contributed substantially to the changes in production for home consumption, occupational choice, and district-level output of rice and wheat over this period. Somewhat surprisingly, the effect of PDS reforms on output seems to be entirely driven by the marketization cost channel rather than any effect on local output prices.[2] Even though prices for agricultural products have not converged across Indian districts over this period, I show that the the PDS expansion appears to increase allocative efficiency by shifting district land allocations towards crops with higher comparative advantage.

The rest of the paper is organized as follows. Section 2 presents a simple model that motivates the empirical analysis and helps make sense of some of the results. Section 3 discusses the data and patterns of home production, specialization, and prices in India including the PDS reforms I use for the empirical analysis. Section 4 presents the main results concerning the effect of PDS expansion on agricultural production and specialization, and section 5 concludes.

2. Model

This section analyzes a simple model of occupational and crop choice to help frame the subsequent empirical analysis. The model abstracts from factor input markets

[1]Some of the more recent literature also focuses on general equilibrium market price effects of in-kind transfers when villages are less "open" to external trade (Cunha et al. (2014)), similar to debates about food aid and farmer income. Uncovering substantial production-side may require more permanent shocks and longer-term data like I use here.

[2]The important role for "marketization" costs I identify is consistent with misallocation and distortions arising from poor infrastructure, low productivity in distribution/retail (Lagakos (2014)), and high markups by traders and middle-men (Atkin and Donaldson (2013)) but relates these to production decisions.

to focus on the interaction between occupational and crop choice, consumption, and trade and marketization costs.

Consumers have Stone-Geary utility functions defined over a food staple good s and a basket of market goods m. Utility is defined as $U = (q_s)^{\alpha}(q_m + \gamma)^{1-\alpha}$ where $\gamma > 0$ implies that the market good is a luxury and the staple is a necessity. Households with very low utility consume only the staple good. Households can purchase the staple good from the marketplace at price p_s, and the price of market goods p_m is normalized to one. Households that earn cash income Y face standard demand functions given by $q_m = (1-\alpha)Y - \gamma\alpha$ and $q_s = \frac{\alpha Y}{p_s} + \frac{\gamma\alpha}{p_s}$.

Household productivity varies across three sectors: non-farm (nf), market farm (m), and staple farm (s). Productivity is determined by a combination of location-specific shifters A_{nf}, A_m, A_s, which one can think of as land suitability for different crops or occupations, and idiosyncratic productivity in farming (z_i) and non-farming (x_i). Households are free to devote their time between all three sectors.

To generate non-separability of consumption and production at the household level I make the simple assumption that while purchasing the staple good from the market entails a price p_s, the value of selling a unit of the staple good from a farmer's standpoint is only $p_f < p_s$. The wedge between these prices can be expressed as $1 + \tau = p_s/p_f$. The marketization wedge τ can be interpreted narrowly as the gap between farm-gate and retail market prices, but can also be interpreted as encompassing any additional marginal costs of converting staple farm output to marketed staple consumption including the time and travel costs of transacting with the market, along with any implicit insurance value of staple production.[3]

Given these assumptions, household i will select into one of five production types that deliver the following utilities:

1. Specialized non-farmer: $V = \alpha^{\alpha}(1-\alpha)^{1-\alpha}\left(\frac{1}{p_s}\right)^{\alpha}[A_{nf}x_i + \gamma]$

2. Specialized market farmer: $V = \alpha^{\alpha}(1-\alpha)^{1-\alpha}\left(\frac{1}{p_s}\right)^{\alpha}[A_m z_i + \gamma]$

3. Specialized staple farmer: $V = \alpha^{\alpha}(1-\alpha)^{1-\alpha}\left(\frac{1}{p_f}\right)^{\alpha}[A_s z_i p_f + \gamma]$

4. Mixed market/staple farmer: $V = (\lambda_s A_s z_i)^{\alpha}([1-\lambda_s]A_m z_i + \gamma)^{1-\alpha}$ where $\lambda_s = \alpha\left[1 + \frac{\gamma}{z_i A_m}\right]$

[3] For example see Fafchamps (1992) for evidence on volatility and subsistence production.

5. Mixed staple farmer/non-farmer: $V = (\lambda_s A_s z_i)^\alpha \left([1-\lambda_s]A_{nf}x_i + \gamma\right)^{1-\alpha}$ where $\lambda_s = \alpha\left[1 + \frac{\gamma}{x_i A_{nf}}\right]$

When $A_s p_s < A_m$ there is no staple production and all households specialize fully in either market farming or non-farming. When $A_m < A_s p_f$ there is no market farming at all and households are either specialized in staple farming, non-farming, or mixed staple farmer/non-farmer. When $A_s p_f < A_m < A_s p_s$ there will also be mixed market/staple farmers who produce just enough staples for their own consumption and devote the rest of their time to market farming. Because market farming confers no consumption advantage over non-farm production (regardless of τ) there is no reason for a household to undertake both market farming and non-farm production simultaneously given the linear production technology. However, where staple farming is supported, households for whom $A_{nf}x_i/p_s < A_s z_i < A_{nf}x_i/p_f$ will choose to produce their own staple consumption while devoting the rest of their time to non-farm production. For households that engage in some staple production but are not fully specialized, the effort devoted to staples (λ_s) and hence the home share of consumption is decreasing in their market productivity. The only households that are totally autarkic (with a home share of consumption equal to one) are those with low enough productivity in both non-farming and farming that they consume none of the market good.

The wedge τ in this model induces greater home-production of staples through two channels. First, it increases entry into farming by non-farmers, as some households that were previously specialized in non-farming devote some time to staple production. Second, it also causes farmers who were previously specialized in market production to devote some of their effort to staple production. In both cases the total value of output falls as the consumption value of the additional staple output (p_s) is larger than the value at farm-gate prices (p_f), and the home produced share of consumption rises. Welfare also falls for all households that were not previously specialized staple farmers. Note that when overall productivity is very high (such that staples make up a very small share of consumption) the effects of a given change in τ on production choices and specialization is necessarily smaller. The effects of a given change in τ will also vary depending on the initial conditions (e.g. increases in τ for areas with specialized staple-farming have no within-farmer effect, and increases in τ in areas with specialized market farming have no effect on production unless the $A_s p_s < A_m$ threshold is crossed.)

Beyond changes in the marketization wedge τ, this framework is helpful for thinking through the effects of different policies. First, consider an in-kind transfer of the staple good (or some combination of in-kind transfer/subsidy that leads to a $\tau < 0$ with quantity limits, like India's Public Distribution System). This will automatically displace home production of staples for households and potentially induce outright exit from staple production if the transfer is marginal rather than infra-marginal (i.e. greater than the desired staple consumption taking into account the income effects of the transfer). Even if the transfer is infra-marginal it will lower the effort devoted to home staple production for mixed households given that staples are a necessary good. The transfer would have no production effects on specialized staple farmers (although there will be displacement of consumption out of home production), specialized market farmers or specialized non-farmers. Note that a cash (market good) transfer would have a very different effect than an in-kind transfer in this context, potentially increasing staple production even if the cost of acquiring staples in the market p_s is unaffected.

A second policy intervention to consider is a support price for staple farmers that raises p_f but without changing the cost of acquiring staples in the market (p_s). This type of policy could increase staple production by leading more households to become specialized staple farmers, but otherwise it will have no effect on the incompletely specialized households that engage in staple farming only for home consumption. India's price supports for rice and wheat farmers through the Food Corporation of India resemble this type of policy.

Third, consider how policies that lower trade costs across locations with different sectoral productivities (A_{nf}, A_s, A_m) affect production in this environment. In general equilibrium, the relative price of staples must adjust to clear the market. In autarky with low enough productivity all households may specialize in staple production, but if productivity is high enough to generate positive demand for market goods a division into non-farmers, staple farmers, and mixed non-farm/staple farmers can arise. Specialization in market farming will not occur in autarky because the price of staples must be high enough to induce the most productive farmers to create a surplus of staples for the non-farming population. With lower trade costs, regions with high productivity in staples may see experience an increase in farming, while regions with low staple productivity will see a decrease in staple production as households exit farming or switch to specialized market or mixed market/staple farming. As trade costs further more of the mixed market/staple farming regions will transition to market-only farming leading to increased "specialization" in agricultural production across regions.

Finally, note that the effects of in-kind transfers can be different under autarky versus trade. In autarky, in-kind staple transfers can lower staple buying and selling prices, inducing increases in staple consumption and potential exit from staple farming for non-recipients. For a small open economy with fixed relative prices (p_f, p_s), the effects of in-kind staple transfers on staple farming and consumption are restricted to the recipients only and there is no amplification of the transfers direct effect on exit from staple farming. Cash (market) transfers would have the opposite effects. Decreases in the marketization wedge τ lead to exit from staple farming under autarky or trade, although under autarky the effects are muted because the home-staple production that is displaced requires an increase in the price p_f to increase surplus (marketed) staple output.

3. Data and descriptive evidence

In this section I describe the Indian data I will later use to test and quantify the importance of marketization costs for agricultural production and specialization. The discussion is based around four themes and some associated descriptive facts: home consumption, production and specialization, prices, and infrastructure.

3.1. Home consumption

My main measure of household consumption comes from India's National Sample Survey. The survey is based on a 30 day recall period [4] and collects basic household characteristics and detailed consumption data for about 100,000 households in the years 1987-88, 1993-94, 1999-00, 2004-05, and 2009-10. Sampling is based on two-stage stratification with first-stage units (villages and city blocks) randomly sampled within a state, and 10-12 households sampled within each first-stage unit. The most disaggregated geographic unit that can be tracked over time, geocoded and matched

[4] In 1999-00 the survey used both a 7 and 30 day recall period for the same households, while in 2009-2010 the survey used the 7 day period for some food items and households and 30 day period for others. Consequently measured consumption and expenditures for 1999-00 may not be exactly comparable to other survey rounds.

to other data sets is the district. District boundaries change over time as districts are often split up, but using consistent 1961 boundaries there are about 300 districts in India with a median district area of 7500 sq.km (equivalent to the 116th ranked Metropolitan Statistical Area in the United States).

The consumption survey consists of a list of individual items such as rice, wheat, flour, milk, chicken, chick peas, spinach, etc. Households are asked to report the quantity and value of goods purchased from the market, received as gifts or in exchange, or home produced. While goods purchased through the market are valued at the actual transaction price, the value of goods received as gifts or in exchange are valued at "average local retail prices" and home produced goods are valued at "ex farm or ex factory rate" not including "any element of distributive service charges."[5] In addition to market purchases, beginning in 1993-94 the survey records quantities and values of rice and wheat purchased through the Public Distribution System. This coincides with the beginning of a period where the PDS transitioned from a universal entitlement scheme to a more targeted scheme offering fixed quantities of staples at prices well below market rate to households with ration cards.

I aggregate up to a consistent set of 134 food categories for comparisons over time. Out of these, 102 are potentially home-producible and together account for a fairly steady 78% of all food expenditures for the average household between 1987 and 2010. Goods like flour and bread are not considered home-producible as any processing of foods by the household subsequent to purchase is not recorded; these goods are only recorded if they are purchased from the market but not if they are consumed at home. By restricting analysis of home shares of food to a consistent set of 102 home-producible foods I aim to capture the decision of whether to grow rice or purchase it in the market, abstracting from the decision of whether to make rice-cakes at home or purchase them in the market which is closer to a "service" margin of marketization than the "goods" one I examine here. This also lowers the importance of quality and product heterogeneity when analyzing prices although there is still lots of scope for quality variation across unprocessed agricultural goods. Food goods that are not home-producible are thus treated like non-food goods for the rest of the

[5]In earlier years of the survey cash, home-produced, and total were recorded separately but in more recent years home production and total are the two categories recorded. In the 1999-00 survey round home and market consumption are not recorded separately – instead households were asked whether consumption was out of cash, home, or both. We treat "both" as home production for reasons discussed below but this, along with the recall period issue mentioned earlier, suggests caution when comparing the 1999-00 data with other years.

analysis.[6]

Between 1987-88 and 2009-10 the aggregate share of consumption out of home production fell in half, from 14% to 7%. Restricting to food consumption only, the decline was of a similar magnitude from 22% to 14%. Table 2 presents a breakdown of these aggregate figures across three types of rural households – farmers (self-employed in agriculture), agricultural laborers, other rural residents – and urban households, along with average real expenditures and land holdings.[7] Farmers make up about a quarter of all households but own the vast majority of the land and have substantially higher home consumption shares for food and for all goods. Agricultural laborers are poorer than "other rural households" but otherwise are very similar in terms of land holdings and consumption out of home production, while urban households are the richest but have the least land and trivial levels of consumption out of home production. The table makes it clear that the aggregate decrease in consumption out of home production is coming from three sources: a decrease in the share of households that are classified as farmers, a large decrease in the home share for farmers, and a much smaller decrease in the home share for other household types. While part of the overall decline in home share is clearly driven by the decline in food share over time, even within food there is a substantial decrease in the home produced share of consumption.

While table 2 makes evident the broad patterns of home consumption over time and across household types, it masks the heterogeneity in home shares across locations. Appendix figure 7 shows that there is considerable heterogeneity in the distribution of household-level and district-level mean home shares of food, with a spike at zero capturing farmers entirely engaged in non-food (particularly cotton) and cash crop production (e.g. oil seeds). The entire distribution of home shares shifted to the left between 1987 and 2010. Appendix figure 8 presents a map of India for 2009-2010

[6]Note that there are some home-producible goods outside of food recorded in the survey – particularly fuel in the form of dung cakes and firewood, but also clothing – but their expenditure shares are very small and they are less related to occupational choice.

[7]There are always issues involved with classifying rural households as farmers or non-farmers. I use the NSS classification of household type "self-employed in agriculture" which is based on the source of the majority of household income in the preceding year, but farmers could also be defined based on household industry (which would include agricultural laborers), land ownership or share of food out of home production. The aggregate patterns are similar qualitatively but different definitions of farmer will obviously lead to different quantitative conclusions. There is also an issue with urban classification, as urban areas are classified partly on the basis of a decennial population census and areas that were previously classified as rural sometimes become urban. See appendix figures 5 and 6 for the entire distribution of food home shares by household type or using a 0.1 hectare land threshold.

where district-level food home shares for farmers are color coded (see legend) and the fraction of households that are farmers corresponds to the size of the dots. The map indicates that home shares of food are lowest in Southern India and coastal areas and highest in more mountainous, remote and landlocked regions in the North, North-East and North-West. The map also shows a positive spatial correlation (0.37, significant at the 1% level) between the fraction of households that are farmers and the mean home share of farmers across districts.

3.2. Production Specialization

While declining shares of consumption out of home production are consistent with increased specialization in the model, it is also possible that that the changes in home shares are driven by changes in consumption patterns rather than production patterns.[8] Table 3 provides more direct evidence of increasing specialization over time using different data sets. Panel A uses International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) micro-data for the periods 1975-84 and 2005-2012, which consists of detailed crop output quantity and value data at the farmer-plot level for a few villages. In the early period farmers are fairly diversified, producing on average 6.4 crops, but this declines to 3.4 by the later period. This specialization is also reflected in a Herfindahl index constructed using value shares for each crop in total output (equal to one for households producing a single crop). The ICRISAT micro-data also provides further corroboration of the decline in the home share of food for farmers using an alternative source.

Panel B returns to the NSS household consumption data, which does not measure production explicitly but allows a proxy for specialization in food production based on the assumption that any food products produced by the farmer are also consumed in positive quantities. The number of distinct crops produced can be constructed in this way for a district, village or household. In all three cases there is a clear downward trend in the number of varieties produced, and the number of crops for farmers in the 2009 NSS round is similar to the ICRISAT from the same period. Looking at all rural households and not just farmers, there is a drop in the number of crops produced but the median rural household is still producing a food crop in the later

[8]For example Li (2013) finds large increases in the number of food varieties consumed by households in India that could lead to declining home share with no change in the number of crops produced.

period, consistent with the idea that many rural households that derive a majority of their income from agricultural labor and/or non-agricultural activities still engage in production of food crops.

Finally, in panel C I use the ICRISAT Village Dynamics in South Asia (VDSA) district-level data which contains crop areas, yields and prices for 16 major crops[9]. For these 16 crops I construct district-level Herfindahl indexes for land shares and value shares, both of which indicate growing district-level specialization over time.[10] Appendix figure 11 presents the trend towards specialization over the longer 1966-2009 period.

Specialization in this case may mean a smaller number of crops produced or concentration of land on specific crops, but such specialization need not reflect comparative advantage or efficiency gains. To measure "allocative efficiency' due to specialization, I use predicted yields for 13 major crops (from the 16 above minus castor seed, linseed and sesamum) from the FAO/GAEZ Agro-Ecological model. The model provides high resolution predicted yields based on climate, soil and terrain conditions for each crop. I first calculate the average potential yield for each Indian district by crop, assuming intermediate input usage.[11] I then calculate an allocative efficiency index based on the ratio of the district's potential yield relative to the Indian average potential yield for each crop, weighting each ratio with the actual land allocation for the 13 crops.[12] I normalize this index for cross-sectional comparisons by constructing an index that weights the district-level relative productivities using a common (all India) land allocation. Thus in the cross-section, variation in the index still reflects "comparative advantage" in the sense that districts that allocate more land to crops with higher relative potential productivities have higher values of the index, even when their aggregate productivity is low. Over time variation in the index comes only from changes in district land allocations towards (or away) from crops with higher or lower than average productivity relative to the Indian average. This measure shows a modest increase for the median district over the sample pe-

[9]These are rice, wheat, sorghum (jowar), pearl millet, maize, finger millet, barley, chick pea, pigeon pea, ground nut, sugar, cotton, sesamum, rapeseed/mustard seed, castor seed and linseed

[10]Value shares are calculated using the harvest prices multiplied by the output quantities in the data set.

[11]To be consistent with the ICRISAT VDSA I use a 1961district boundary shapefile for the calculation.

[12]Note that land allocated to the 13 major crops covers over 75% of the cultivated area in the sample districts, a share that changes little over time.

riod but the mean actually falls slightly, indicating that the small increase in district specialization based on Herfindahl measures does not necessarily imply an increase in allocative efficiency. An advantage of this index is that unlike actual yields that reflect input intensity and output prices (which are themselves endogenous with respect to trade costs, marketization costs and policy variables), the index is based on purely exogenous (land/climate-based) differences in relative productivity.

As with home shares, there is large dispersion in the degree of specialization across districts. Appendix Panel A of figure 9 shows the dispersion in specialization (Herfindahl index in area) across districts for various years, while Panel B shows the fraction of land used to grow rice and wheat, which shows a clear pattern of polarization over time with some areas becoming more specialized in these staples and others decreasing the share of land allocated to them. Appendix figure 10 presents the mean crops per farmer or total crops per district estimated from the NSS data, and shows a similarly wide dispersion of specialization by the average farmer across space. On average allocation of land to rice and wheat is highly correlated with the relative agro-ecological productivity across districts but there are also substantial deviations.

While patterns of specialization across districts are likely related to heterogeneity in land-based productivity and market access (e.g. the A terms and p_f in the model), the model also highlights a role for marketization wedges in generating differences across households in the same location. Specifically, more productive farmers may allocate more of their production to non-staple products with a higher market value and smaller "farmers" may only be farming for home consumption purposes. While the NSS data does not measure agricultural production, the ICRISAT farmer-level data offers some additional insight into how farm production and consumption patterns vary across farmers in the same location. The first panel of figure 1 measures the fraction of each crop by farmer cell that is sold (as opposed to consumed/stored) during the 2009 year. I plot the kernel density across all farmers, which reveals a distinct bimodal pattern with a large number of crop-farmer cells that are almost entirely home consumed (sale fraction below 20%) and a decent number that are almost entirely sold (sale fraction above 80%), with a smaller number of crop-farmer cells that are evenly split between home consumption and sale to the market. The selling crops are often, but not always, non-food cash crops like cotton and oil seeds. Comparing large farmers (the top 10%, with over 14.125 hectares of land) to small farmers (the bottom 10%, with under 1.125 hectare) in the ICRISAT villages, we see that the small farmers are more skewed towards home production and large farmers are more skewed towards market production.

The second panel of figure 1 provides evidence that the subsistence motivation of smaller farmers results not only in different crop choices but also "less efficient" crop choices from the point of view of cash income. For each ICRISAT village and crop I measure the average market income per hectare by multiplying total yield by the harvest price and dividing by the total area under cultivation. I then measure the fraction of land for each farmer that is devoted to crops that are in the bottom 50% and bottom 25% of crops in terms of revenue per hectare. Using either measure, there is large, sharp decline in the fraction of land devoted to these "marginal" crops going from low to modest size farms. This shows that larger/richer farmers differ not only in how much they sell to the market but also in which crops they produce – they produce crops that are more "efficient" if one ignores consumption motives.

3.3. Prices

Price data also provide insight into the potential of marketization cost and trade costs to explain the production patterns described above. The NSS consumption survey allows for the derivation of "unit value" prices by dividing total value/expenditure by quantity for many items. For food there are potentially three types of prices – market prices, farm-gate prices, and prices for goods purchased through the Public Distribution System (PDS). Calculating market prices is straightforward. Farm-gate prices are reported by farmers who consume home-produced goods, and are supposed to capture the value of the product "net of distribution costs."[13] PDS prices and quantities are recorded separately starting in the 50th NSS round and reflect the actual price paid.[14] I estimate the "price" within a district taking the median across households, which minimizes the impact of outliers and measurement error. Quality differences between market, home, and PDS products may not be that large in recent periods and for such disaggregated items but they do suggest caution in interpreting simple

[13]In practice home produced quantities are measured, and then the "value of consumption" for home produced products is estimated by multiplying that quantity by the self-reported farm-gate price. I derive the farm-gate prices by dividing the imputed value by quantity.

[14]In practice there are some non-food goods purchased through the PDS as well as some other food goods besides rice and wheat in some states, but these amount to a small fraction of PDS sales and are not recorded consistently across years.

unit value differences as price differences. An alternative source of price data comes from the ICRISAT VDSA district data set which has compiled harvest prices (based on a sample of "mandis" or wholesale agricultural markets) at the district level over the 1966-2010 period.

The price data reveal a potentially large role for marketization costs and the extent of government intervention in Indian agricultural markets. Figure 2 plots the ratio of market to harvest prices for each district (rural areas only), based on taking the median ratio across 10 major food products (excluding rice and wheat). The typical wedge is almost always greater than one and market prices are often over 50% higher than harvest prices. Interestingly the wedge has changed very little between 1993-2010. Although this wedge may be capturing differences in apparently similar products (e.g. quality, packaging, convenience and retail amenities) it also suggests a major incentive to produce for home consumption. It also potentially understates the advantages of home consumption if transacting with the market entails additional fixed or marginal costs.

An important exception to the large market/harvest price gaps are the goods that feature heavy government intervention – rice and wheat. Before the expansion of the targeted PDS system in the 1990s, the gap between market (pooling PDS and non-PDS) and farm-gate prices (within a district) for rice and wheat was typically zero (see Appendix table 9) whereas pulses featured market prices about 10-20% above farm-gate prices and potatoes and onions featured market prices 20-40% above farm-gate prices. This was a direct consequence of the government PDS which until this period consisted of some combination of support prices for farmers and large subsidies on distribution costs benefiting consumers. On the farm-gate/harvest price side, the Food Corporation of India sets national minimum support prices for farmers and implements them through procurement. This policy theoretically applies to 25 crops but is dominated by rice and wheat, with the procured rice and wheat allocated to states for use in the Public Distribution System.

The public distribution system until 1992 was considered a universal entitlement, but due to different implementation by the individual states that set prices and run the system of fair price/ration shops where PDS goods were sold, the quality, availability, and actual prices charged for PDS rice and wheat varied significantly. Beginning in 1992 the "Revamped" PDS program specifically targeted tribal/remote areas and offered prices at 50% of the cost of procurement to below poverty line (BPL) households in some areas. In 1997 the nationwide transition to the Targeted PDS system theoretically led to universal entitlements for BPL households who could purchase a fixed quantity at similarly (50%) subsidized price. Due to differences in the identification and targeting of households for BPL cards, pre-existing differences in PDS systems across states, differences in the implicit subsidy (through different PDS prices, market prices, relative qualities, and quantity limits), and differences in the actual availability of goods at the official prices to nominally entitled households (e.g. due to diversion/theft which in some cases exceeded 50%), experiences with the targeted PDS differed greatly across states. Appendix figures 14, 16, 15 provide evidence of the heterogeneity in the reach (fraction of households), conditional quantity (for recipient households) and conditional implicit subsidy value (for recipient households) in 1993 and 2009. Despite this heterogeneity in implementation at the state-level, the national-level reforms led to a clear increase in the quantity of rice and wheat purchased through the PDS, particularly since 2005. Appendix figure 13 shows that the PDS quantity per person rose by 1KG from 1993-2009. Appendix figure 12 shows that the implicit subsidy value of the PDS rose from below 1% of aggregate consumption in 1993 to over 2% in 2009, while the size and increase was even larger for receiving households (from 3% to 7% of their expenditures). Appendix table 9 shows how this effectively lowered the average price of rice and wheat purchased from the "market" (PDS + open market) relative to farm-gate prices.

The increased attractiveness of procuring rice and wheat through a combination of PDS and market purchases effectively lowered the marketization cost of rice and wheat and potentially reversed the cost advantage of home production for many households. No such trends were observed for other major food products, for which large increases in government intervention were lacking. Somewhat surprisingly, there is also little evidence of price convergence across markets in India during this period. While there have historically been numerous barriers to agricultural trade across states and even districts, there is little evidence in the data between 1987 and 2009 of any price convergence trends due to relaxation of these policy barriers or reductions in transport costs due to expansion of the highway network.[15] Figure 3 documents this using the district harvest prices from the ICRISAT VDSA which is available in every year. I plot the coefficient of variation (standard deviation divided by mean) across districts for each year in the price of several major agricultural goods.

[15]See Atkin (2013) for a discussion of these trade barriers and reforms.

Neither the PDS goods (rice and wheat) nor the less regulated goods show evidence of spatial price convergence over time. Of course this does necessarily imply that districts in India are less autarkic in 2009 than in the 1980s, but it does suggest that the patterns of specialization described above do not seem to be a result of conventional trade forces across districts that would lead to price convergence and specialization through arbitrage.[16]

3.4. Infrastructure

Although prices did not appear to converge across districts and market/harvest prices did not appear to change much for goods not subject to the massive PDS expansion, there were still major development in terms of road infrastructure. One such measure is the all-weather district road density compiled from state yearbooks in the ICRISAT VDSA, which shows both a substantial increase in road density over the the 1987-2009 period as well as heterogeneity across districts. Another measure of infrastructure that may be a better proxy for the degree of trade integration across districts is national market access. I construct this measure combining the national highway maps collected by Allen and Atkin (2015) with data on the value of agricultural output in each district.[17]

Figure 4 plots three measures of agricultural specialization against the local district roads measure and the national market access measure in the 1993 cross-section. The first measure is the fraction of the rural population engaged in cultivation from the census, which captures potential incentives for agricultural production.[18] The data reveal that districts with higher road density tend to have smaller fraction of the population engaged in agriculture. There is no similar effect from national market ac-

[16]See also Mallory and Baylis (2012) for evidence of low spatial integration across Indian agricultural markets.

[17]Specifically, I follow Allen and Atkin (2015) and construct a weighted market access measure for each district j given by $Mktaccess_{jt} = \sum_{i\ neqj} \frac{val_{it}}{hours_{ijt}^{1.5}}$ where val_{it} measures the value of agricultural output in each district by combining output and prices from the ICRISAT VDSA for major crops. Allen and Atkin (2015) calculate a highway travel time in hours that varies as the Indian highway system expands during select years of the sample period. I linearly interpolate the travel times for missing years but use actual annual output values. Note that one could use population weights instead for a market access measure, but from the point of view of agricultural markets it is the distance to other sources of agricultural output that are likely more relevant for specialization than proximity to large urban areas.

[18]Note that non-farmers may still engage in some production in both the NSS and census definitions of farming, which are based on the main source of income for the household.

cess. The second measure constructs the district aggregate share of consumption out of home production from the NSS data (rural areas only). Based on the data shown earlier, this measure may capture both exit from "farmer" status as well as decreases in the home share within farmers, agricultural laborers and non-agricultural rural households. There is again a strong positive correlation with local roads, while the correlation with national market access is actually positive. These differences highlight that trade costs as conventionally measured – differences in prices or transport costs across **markets** in locations with different land-based comparative advantage – may have a very different effect on agricultural production patterns than more local infrastructure that may capture the "marketization" costs that affect production choices for individual households. Note also that the correlation between district roads and national market access is weakly negative in the cross-section so these variables appear to capture very different aspects of the economic environment. The third measure in figure 4 is based on the index of comparative-advantage based land allocation described earlier ("land efficiency"). The data reveal a strong positive correlation between both local roads and national market access and the extent to which land is allocated to crops with higher land-based relative productivity.

While none of these correlations have a clear causal interpretation, these patterns are suggestive of an important role for both local and national infrastructure on production patterns. In particular, the variation in local roads is correlated with production patterns in a way similar to the marketization cost τ in the model. The effect of national market access, like trade costs in the model, are more ambiguous. Some districts may become less specialized (switching from staple to staple and market), others may become more specialized (switching from staple and market to market only), and districts where staple prices increase due to trade may see an increase in home production and the number of farmers due to subsistence motives and marketization costs.

4. Estimation of the effects of PDS expansion

In this section I test whether household production patterns respond to the PDS expansion, which I interpret as a combined decrease in marketization costs and in-kind transfer that likely lowers incentives for rice/wheat/staple production for affected households. Compared to data on market/harvest price wedges, which suffer from substantial measurement error, selection (because farm-gate/harvest prices and market prices are not always observed in the same location), and endogeneity with respect to production decisions, the policy variation from the PDS provides a cleaner test of the marketization cost mechanism. Absent any marketization cost, the only reason the PDS would affect production patterns are general equilibrium price effects or income effects that I can account for with my data. While I also explore the role of increases in road density and national market access using variation within districts over time, variation in these variables is less plausibly exogenous and is less directly tied to a market vs. home consumption motive for agricultural production.[19] Beyond providing a test for whether production/consumption non-separability drives agricultural production patterns at an aggregate level and in a manner consistent with the model, quantifying the effects of the PDS expansion on production is itself a novel contribution to the literature on in-kind transfers and India's PDS specifically.

4.1. Within-village analysis

I begin the analysis by looking directly at the effects of PDS access for households living in the same village. This has one major advantage over district-level analysis, which is that the households being compared are guaranteed to face very similar output prices and marketization costs (both for purchasing from the market and selling output to the market). Thus any effects I estimate within village cannot be driven by general equilibrium effects of the PDS on prices and reflect the pure effect of PDS access on production outcomes. The first specification I consider uses data from rural areas in 2004-2005. This round of the NSS has the advantage that ownership of BPL cards is recorded in the survey; for other rounds one can observe PDS usage or not on the extensive margin but some households without BPL cards may access the system and some with BPL cards may not be able to purchase from the system. The equation I estimate is:

$$Y_{hv} = \alpha_v + \beta BPLcard_{hv} + \Omega Controls_{hv} + \epsilon_{hv} \tag{1}$$

where h is household, v is village, Y is the household outcome of interest, BPL is a dummy for actual BPL card ownership, and α_v is a village fixed effect. The main household outcomes I consider are farmer status, a dummy for consumption

[19]Measurement error is a concern with all of the variables but may be worse for the road data as it involves collection from multiple sources and/or interpolation.

of rice/wheat out of home production (a proxy for rice/wheat production), a count of the total number of food products consumed out of home production, and the share of consumption out of home production.

The first four columns of table 4 look at these outcomes for all rural households; the last three columns restrict the sample to farmers so the variation is across farmers only. Note that BPL card ownership varies both because of eligibility (which varies across states but includes both expenditure and asset information collected in special BPL surveys) and implementation (including potentially sale of BPL cards to others), both of which vary across states. Households that own BPL cards are obviously different than households that do not own them, so the simple difference cannot be regarded as causal. Since BPL households are poorer on average, through the lens of the model one might expect them to have higher home-production of staples, suggesting that the simple differences I find are a lower bound. I can also control for predetermined household characteristics that capture socio-economic status and BPL eligibility, including household demographics, education, scheduled caste/scheduled tribe status and religion. Finally, because PDS access potentially has an income effect on households, I can include a control for household expenditures that adjusts for the implicit value of the subsidy. While the results are unlikely to capture a pure causal effect of BPL card ownership, they are consistent with a large effect of PDS access on production patterns. Rural households in the same village with BPL cards are 13 percentage points less likely to be farmers, 10 percentage points less likely to consume rice/wheat out of home production, consume 0.5 less products out of home production and see a 5 percentage point decline in the overall home share. The production effects can also be observed within farmer households only, where I also control for a quadratic in land ownership.

Because comparing households with and without BPL cards raises endogeneity concerns, I also consider an adapted within-village specification that uses variation in the generosity of the PDS system within states over time for identification. I use NSS data from 1993, 1999, 2004 and 2009 for this exercise. The specification is given by the following equation:

$$Y_{hvst} = \alpha_v + \beta(PDS_{hvst} * CondPDSvalpp_{st}) + \gamma_1(PDS_{hvst} * \eta_t) + \gamma_2(PDS_{hvst} * \omega_s) + \epsilon_{hvst} \tag{2}$$

where Y is the household outcome of interest (same as above), PDS is a dummy variable for whether the household accesses the PDS or not, and $CondPDSvalpp$ measures the state-level average value of the PDS subsidy for recipients (average quantity for PDS using households multiplied by the difference between the average PDS price and the average market price in the state, deflated to 1993 rupees using the national price index for rural laborers). I include interactions of the PDS dummy with state and time dummies, which controls for any time-invariant differences between PDS and non-PDS households within a state and any common variables affecting PDS and non-PDS households differentially over time. The coefficient of interest is β, the coefficient on the interaction term; identification is based only on variation in the generosity of PDS subsidies for the average PDS recipient within states over time. While there is still some potential for endogeneity (e.g. if there is time-varying selection into/out of the PDS system on the extensive margin when subsidy values go up or down) this "difference-in-difference" approach addresses some of the concerns with the cross-sectional approach. Panel B of table 4 presents these results. While the estimated coefficient on the interaction term is not statistically significant in two cases, the results point clearly in the same direction as the results in Panel A – increases in the state-level PDS transfer lead to larger production effects for PDS recipient households than non-recipient households.

4.2. District-level analysis

While the results from table 4 show that households seem to respond on multiple margins when they have access to a better alternative than home production, they do not show direct effects on production output. A rice farming household may happily take subsidized PDS rice while still devoting all of their land to rice production for market.

The district level specifications I use are based on the following equation:

$$Y_{dt} = \alpha_d + \gamma_t + \beta QPDSpc_{dt} + \Omega Controls_{dt} + \epsilon_{dt} \quad (3)$$

where Y is the district-level outcome. The specification includes district and time (NSS round) fixed effects, and I always include the following controls: the fraction of households in the district that are below the 1993 Indian poverty line in terms of real per capita expenditure, average real per capita expenditure, the rural population, and the district road and market access measures described earlier. My main district-level measure of changes in PDS intensity is the quantity of rice and/or wheat purchased through the PDS per household; this measure captures both increases on the extensive margin (households using the PDS that did not previously) and the intensive margin (either because official quantity limits changed or actual availability changed).[20] Standard errors are always clustered by district. While the number of districts varies based on data availability depending on the variable under consideration, the main sample includes over 270 districts in the 16 major Indian states for the years 1993, 1999, 2004, and 2009.

Because PDS transfers at the district level could affect production through the prices received for rice and wheat, and not just through the wedge between the value of market and home consumption, I first look at the effects of PDS expansion on prices. Table 9 presents the results based on NSS farm-gate or ICRISAT VDSA harvest price data. I also consider an interaction of the PDS variable with a dummy variable for districts with higher than median market access in 1993, motivated by the idea that districts that are more integrated with other districts may exhibit smaller general equilibrium price effects (consistent with the village-level findings from Cunha et al. (2014)). The results point to fairly small price effects overall (recalling that the average increase in PDS quantity per capita is about one over the 1993-2009 expansion), with effects that are sometimes positive (for rice using harvest prices) and with interaction effects that do not suggest a less negative effect on prices for districts with above median market access.

In Table 6 I turn directly to district level output of rice and wheat to see whether these are affected by PDS subsidies in a manner consistent with the model and the within-village evidence presented earlier. I consider both rice and wheat output combined as well as output measured separately, in which case I also control for farm-gate prices directly. The effects of PDS expansion are quite large although PDS rice and wheat do not quite crowd out district production of rice and wheat one for one, consistent with generally positive effects on overall rice and wheat consumption.[21] This direct evidence that a household consumption subsidy has considerable effects

[20]Results are generally similar using value based subsidies (e.g. multiplying the quantities by gaps between market and PDS prices) because the correlation between the two series is high, but value based measures suffer from greater measurement error (because they require price information as well) and may also partly reflect variation in market prices that themselves respond to shifting production patterns.

[21]See appendix tables 11, 12, 13. Some of the crowd-out is likely affecting production of other, inferior staples.

on agricultural output highlights the link between consumption and production. The fact that this effect persists when controlling for prices (and given the limited effect on prices discussed above) suggests that this is not simply due to the conventional trade forces whereby an in-kind transfer lowers local output prices. I also consider an interaction of the PDS variable with the roads and market access variables (measured as above/below 1993 median once again). The interaction effects are generally not significant for output except in one instance, but point towards a larger effect of the PDS in districts that have better access to the national market.

Table 7 examines the broader measures of specialization described previously and shown in figure 4 – district home share, fraction of farmers, and allocative efficiency based on land allocation and land-based comparative advantage. The first three columns look at district home shares for rural areas, with the second and third columns including interactions with above median market access and road density. Increases in the PDS quantity per capita lead to modest reductions in district home shares of about one percentage point, with much smaller (essentially zero) effects in districts with higher road density. Note that this is consistent with the model, which predicts that absent substantial marketization costs (τ) there is no reason a PDS transfer would affect home shares. The next three columns look at the fraction of households that are farmers in rural areas. The consistent finding is that the increase in PDS quantities leads to modest exit from farming (about 0.6 percentage points for a one KG/capita increase in PDS transfers), with the interaction effects again pointing to more limited effect on occupational choice in areas that have lower marketization costs (higher roads) already. Finally, the last three columns look at the allocative efficiency measure that weights changes in land shares devoted to crops by their relative land-based productivity. The effects here are quite large relative to the overall change in allocative efficiency over this period (only a 2 percentage point increase for the median district over the 1993-2005 period). Interestingly the effect here is much larger for districts with above median market access, suggesting that these districts are best able to take advantage of the subsidy to reallocate land to comparative advantage crops. Districts that are less integrated into the national agricultural market do not seem to feature such "efficient" land reallocation.

While I have treated PDS expansion (conditional on year fixed effects and several control variables) as an exogenous policy change, it is possible that increases in PDS quantities were targeted to districts with downward trends in rice/wheat/home production, fraction of farmers, and allocative efficiency for some other reason. While this is difficult to rule out conclusively, with multiple periods of data I can examine whether future (forwarded) values of the PDS transfer predict current outcomes, which in a panel fixed effects regression with continuous treatment is similar to testing for pre-trends in a difference-in-difference framework. The evidence presented in table 8 suggests this is not the case, as the forwarded values are much smaller in magnitude and not statistically significant. The one exception is the fraction of farmers, but the magnitude is still much smaller and this may be partly an artifact of the interpolation of population estimates between decennial census years.

In terms of the overall quantitative effects of the PDS expansion, the within-village specifications point to potentially very large effects on exit from farming, the decline in the home produced share of consumption and the number of home produced varieties that account for anywhere from a quarter to all of the observed changes over the 1993-2009 period. The district-level specification delivers smaller effects that still account for at least 15% of the observed changes, and in the case of allocative efficiency more than half the change for the median district.[22] Effects of this magnitude suggest that any evaluation of the overall effects of the PDS or proposals to reform the system with cash transfers should seriously consider the effects on agricultural specialization and production.[23]

5. Conclusion

Schultz (1953) identified the "food problem" as the combination of low agricultural productivity and high employment in agriculture in poor countries. The presence of high trade costs for agricultural products combined with food being a necessity can generate this pattern, and high agricultural trade costs within countries have been shown to further lower productivity through lower comparative-advantage based specialization across crops. This paper argues that the "food problem" also operates at a smaller household scale in the presence of marketization costs, which affect

[22]The change in district means over 1993-2009 are -0.072, -0.024, and -0.08 for district home share, fraction farmers and allocative efficiency respectively while the predicted changes due to the increase in PDS quantity of approximately 1KG/pp are -0.008, -0.006 and 0.012. Median allocative efficiency rose by 0.02.

[23]Khera (2011) finds that PDS beneficiaries often prefer the convenience of in-kind transfers from more conveniently located PDS outlets to dealing with cash and the associated banking and market transactions.

choices about whether to farm and what to farm. This additional channel matters because food is important for poor households and farming may offer a substantial consumption advantage for food. This advantage, which is excluded from market-based valuations of agricultural output, comes at a potential cost in terms of market-based allocative efficiency across occupations and crops.

In this context, I show that the massive intervention of the Indian government in agricultural markets through the PDS – a policy designed primarily to alleviate poverty and malnutrition through its consumption effects – has important effects on production, consistent with the non-separability of consumption and production decisions for agricultural households. The PDS reforms have led households to exit farming and become more integrated into the market, increasing allocative efficiency. This is all the more surprising given the limited evidence that integration across district markets has done much for price convergence and agricultural specialization along the lines of comparative advantage.

However, it is important to stress that the evidence presented here does not provide a full account of India's interventionist agricultural policy. I have ignored the effects of national support prices, which may be distorting production decisions along other margins and increasing farming of rice and wheat in some regions. The price subsidy and distribution network that underlie the PDS also have large costs in terms of economic resources. Rather than providing a full picture of the costs and benefits, winners and losers of the current intervention into agricultural markets by the Indian government, my results instead caution that evaluation of reforms to India's PDS and considerations of in-kind versus cash transfers more generally cannot focus solely on consumption-side effects. Policies that distort the consumer market along one margin could improve allocative efficiency on the production margin when households are close to the margin of subsistence and marketization costs incentivize subsistence agriculture. Government intervention could then potentially mitigate the effect of other "marketization" frictions like low productivity in retail and distribution that could lower agricultural productivity through misallocation.

References

Adamopolous, Tasso, "Transportation Costs, Agricultural Productivity, and Cross-country Income Differences," *International Economic Review*, 2011.

Allen, Treb and David Atkin, "Volatility, Insurance and Gains from Trade," *Working Paper*, 2015.

Atkin, David, "Trade, Tastes and Nutrition in India," *American Economic Review*, 2013, *103(5).*

_ and Dave Donaldson, "Who's Getting Globalized? The Size and Implications of International Trade Costs," *Working Paper*, 2013.

Basu, Deepankar and Debarshi Das, "Social Hierarchies and Public Distribution of Food in Rural India," *Working Paper*, 2014.

Costinot, Arnaud and Dave Donaldson, "How Large are the Gains from Economic Integration? Theory and Evidence from U.S. Agriculture, 1880-1997," *Working Paper*, 2014.

Cunha, Jesse, Giacomo De Giorgi, and Seema Jayachandran, "The Price Effects of Cash Versus In-Kind Transfers," *Working Paper*, 2014.

de Janvry, Alain and Elisabeth Sadoulet, "Progress in the modeling of rural households' behavior under market failures," in "Essays in Honor of Erik Thorbecke," Kluwer Publishing, 2006.

_ , Marcel Fafchamps, and Elizabeth Sadoulet, "Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained," *Economics Journal*, 1991, *101*, 1400–1417.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *Forthcoming American Economic Review*, 2012.

Dreze, Jean and Reetika Khera, "Rural Poverty and the Public Distribution System," *CDE Working Paper*, 2013.

Fafchamps, Marcel, "Cash Crop Production, Food Price Volatility, and Rural Market Integration in the Third World," *American Journal of Agricultural Economics*, 1992, *74(1)*, 90–99.

__ , "Development, agglomeration, and the organization of work," *Regional Science and Urban Economics*, 2012, *42*, 459–472.

Gollin, Doug, David Lagakos, and Michael Waugh, "The Agricultural Productivity Gap," *Quarterly Journal of Economics*, 2014, *129(2).*

Gollin, Douglas and Richard Rogerson, "Productivity, transport costs and subsistence agriculture," *Journal of Development Economics*, 2014.

Kaushal, Neeraj and Felix M. Muchomba, "How Consumer Price Subsidies affect Nutrition," *World Development*, 2015, *74*, 25–42.

Khera, Reetika, "Revival of the Public Distribution System: Evidence and Explanations," *Economic and Political Weekly*, 2011, *46*, 36–50.

Kochar, Anjini, "Can Targeted Food Programs Improve Nutrition? An Empirical Analysis of Indias Public Distribution System," *Economic Development and Cultural Change*, 2005, pp. 203–235.

Krishnamurthy, Prasad, Vikram Pathania, and Sharad Tandon, "Food Price Subsidies and Nutrition: Evidence from State Reforms to India's Public Distribution System," *Working Paper*, 2014.

Lagakos, David, "Explaining Cross-Country Productivity Differences in Retail Trade," *Journal of Political Economy*, 2014.

__ and Michael Waugh, "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 2013, *103(2)*, 948–980.

Li, Nicholas, "An Engel Curve for Variety," *Working Paper*, 2013.

Mallory, Mindy and Kathy Baylis, "Food Corporation of India and the Public Distribution System: Impacts of Market Integration in Wheat, Rice, and Pearl Millet," *Journal of Agribusiness*, 2012, *30(2)*, 225–250.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu, "Agriculture and aggregate productivity: A quantitative cross-country analysis," *Journal of Monetary Economics*, 2008.

Schultz, T.W., *The Economic Organization of Agriculture*, New York: McGraw-Hill, 1953.

Sotelo, Sebastian, "Trade Frictions and Agricultural Productivity: Theory and Evidence from Peru," *Working Paper*, 2014.

Swiecki, Thomasz, "Intersectoral Distortions and the Welfare Gains from Trade," *Working Paper*, 2014.

Tarozzi, Alessandro, "The Indian Public Distribution System as provide of food security: Evidence from child nutrition in Andhra Pradesh," *European Economic Review*, 2005, *49*, 1305–1330.

Tombe, Trevor, "The Missing Food Problem: Trade, Agriculture, and International Productivity Differences," *American Economic Journal: Macroeconomics*, 2014.

Table 1: Subsistence farmers across countries

	Uganda 2009	Timor 2000	India 2009	Peru 1994	Guatemala 2000
GDP per capita USD2000	384	494	948	2080	2092
Home share food	0.39	0.37	0.12	0.16	0.17
Agric. VA/worker	221	508	600	1095	1945
Households with land	0.67	0.76	0.61	0.39	0.48
Conditional on home food share $\geq 1\%$					
Home share food exp	0.53	0.43	0.31	0.39	0.30
Home share food varieties	0.34	0.36	0.11	0.26	0.12
Survey food varieties	64	129	134	30	90

All data from World Bank LSMS and World Development Indicators except for India.

Table 2: Farmers and home production

	1987-88				2009-10			
	Farmer	Ag.laborer	Rural other	Urban	**Farmer**	Ag.laborer	Rural other	Urban
Share of households	**0.27**	0.23	0.23	0.26	**0.22**	0.19	0.29	0.29
Real p.c. exp.	**179**	133	189	269	**205**	156	223	364
Land (ha)	**2.6**	0.39	0.57	0.24	**1.7**	0.24	0.27	0.12
Home share of all	**0.29**	0.06	0.07	0.02	**0.2**	0.04	0.04	0.01
Food share of all	**0.67**	0.67	0.65	0.62	**0.56**	0.57	0.55	0.5
Home share of food	**0.43**	0.09	0.11	0.03	**0.36**	0.07	0.08	0.01

All data from the Indian NSS.

Table 3: Agricultural specialization

	Mean	Median	Mean	Median
	1975-1984		2005-2012	
A. ICRISAT micro (farmer-level) data				
Farmer count of crops	6.4	6	3.4	3
Crop value Herfindahl index	0.51	0.44	0.6	0.53
Home share of food	0.53	0.55	0.28	0.25
	1987		2009	
B. NSS data				
District count of crops	35	35	27	26
Village count of crops	9	8	6.5	5
Farmer count of crops	3.9	3	3.57	3
Rural household count of crops	2.09	1	1.76	1
C. ICRISAT VDSA (district-level) data				
16 crop area Herfindahl index	0.41	0.32	0.43	0.38
16 crop value Herfindahl index	0.41	0.33	0.46	0.41
Allocative efficiency index	1.35	1.29	1.34	1.31

See text for description.

Table 4: Within-village effects of PDS on agricultural activity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. variable	Farmer	Home rice/wheat	Number of home foods	Home share	Home rice/wheat	Number of home foods	Home share
	All rural households				Farmers only		
	Panel A: BPL card vs. non-BPL card rural households within-village (2004-2005)						
BPL card	-0.127***	-0.107***	-0.522***	-0.047***	-0.087***	-0.287***	-0.028***
	(0.005)	(0.004)	(0.023)	(0.002)	(0.009)	(0.056)	(0.004)
Observations	66,857	66,857	66,857	66,857	22,643	22,643	22,643
Adj R-squared	0.181	0.386	0.538	0.424	0.573	0.670	0.528
	Panel B: PDS vs. non-PDS households and state-level variation in PDS generosity (1993-2005)						
PDS usage	-0.074***	-0.076**	-0.098	-0.043**	-0.102***	-0.026	-0.070***
	(0.021)	(0.031)	(0.169)	(0.016)	(0.027)	(0.105)	(0.022)
PDS x subsidy value	-0.006***	-0.004	-0.020**	-0.002**	-0.005*	-0.031**	-0.002
	(0.001)	(0.002)	(0.009)	(0.001)	(0.003)	(0.011)	(0.002)
Observations	239,549	239,549	239,549	239,549	84,710	84,710	84,710
Adj R-squared	0.257	0.405	0.553	0.462	0.558	0.669	0.546

Heteroskedasticity robust standard errors in parentheses, clustered by state for Panel B.
Controls as described in the text.
*** p<0.01, ** p<0.05, * p<0.10

Table 5: Effects of PDS subsidies on district farm-gate/harvest prices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var	Rice price				Wheat price			
Source	NSS	NSS	Harvest	Harvest	NSS	NSS	Harvest	Harvest
PDS rice quant p.c.	0.010	0.013	0.022***	0.018**				
	(0.009)	(0.009)	(0.008)	(0.009)				
PDS wheat quant p.c.					0.004	0.005	0.000	0.001
					(0.009)	(0.009)	(0.007)	(0.007)
Quant x mkt. access		-0.017		0.032		-0.041**		-0.013
		(0.019)		(0.020)		(0.021)		(0.015)
Observations	900	900	834	834	805	805	824	824
Adj R-squared	0.877	0.877	0.858	0.859	0.923	0.923	0.943	0.943

Robust standard errors in parentheses clustered by district.
Additional controls include district and year fixed effects, poverty rate, real per capita expenditure, population, roads and market access.
*** p<0.01, ** p<0.05, * p<0.10

Table 6: Effects of PDS on district Rice/Wheat Output

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.	Output of rice and wheat p.c.				Output of rice p.c.		Output of wheat p.c.	
PDS quant p.c.	-0.432***	-0.661***	-0.647***	-0.728***				
	(0.157)	(0.203)	(0.224)	(0.277)				
PDS rice quant p.c.					-0.914***	-0.768***		
					(0.278)	(0.286)		
PDS wheat quant p.c.							-0.794**	-0.495
							(0.353)	(0.379)
Interaction			-0.051	0.134		-0.940*		-0.684
			(0.413)	(0.362)		(0.569)		(0.668)
Mkt. access		4.644**	4.628**	4.622**	4.203***	3.785**	0.690	0.709
		(2.216)	(2.257)	(2.224)	(1.578)	(1.586)	(1.494)	(1.511)
Road density		1.112**	1.116**	1.127**	0.887***	0.945***	0.250	0.283
		(0.438)	(0.438)	(0.442)	(0.301)	(0.304)	(0.272)	(0.272)
Rice farm-gate price					0.748	0.687		
					(0.946)	(0.950)		
Wheat farm-gate price							-0.802	-0.774
							(1.043)	(1.045)
Int. variable			Mkt. access	Roads		Mkt.access		Mkt.access
Observations	1,213	1,100	1,100	1,100	900	900	805	805
Adj R-squared	0.977	0.978	0.978	0.978	0.954	0.954	0.971	0.971

Robust standard errors in parentheses clustered by district.
Additional controls include district and year fixed effects, poverty rate, real per capita expenditure, and population.
*** p<0.01, ** p<0.05, * p<0.10

Table 7: Effects of PDS on district-level agricultural specialization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep.var	District home share			Fraction farmers			Allocative efficiency		
PDS quant p.c.	-0.008***	-0.008***	-0.012***	-0.006***	-0.005***	-0.008***	0.012**	0.006	0.016***
	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)	(0.005)	(0.006)	(0.004)
Interaction		0.002	0.008*		-0.002	0.005***		0.018**	-0.008
		(0.005)	(0.004)		(0.002)	(0.002)		(0.007)	(0.007)
Mkt.access	0.031	0.032	0.030	0.017**	0.016**	0.016**	0.028	0.035	0.029
	(0.022)	(0.022)	(0.022)	(0.008)	(0.008)	(0.008)	(0.024)	(0.023)	(0.024)
Road density	-0.009	-0.010	-0.008	-0.001	-0.001	-0.001	-0.015***	-0.016***	-0.016***
	(0.006)	(0.006)	(0.006)	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)
Int.variable		Mkt.access	Roads		Mkt.access	Roads		Mkt.access	Roads
Observations	1,113	1,113	1,113	1,113	1,113	1,113	997	997	997
Adj R-squared	0.800	0.800	0.800	0.945	0.945	0.946	0.992	0.992	0.992

Robust standard errors in parentheses clustered by district.
Additional controls include district and year fixed effects, poverty rate, real per capita expenditure, and population.
*** p<0.01, ** p<0.05, * p<0.10

Table 8: Robustness to pre-trends: Effects of PDS on district-level agricultural specialization

	(1)	(2)	(3)	(4)	(5)	(6)
Dep.var.	District home share		Fraction farmers		Allocative efficiency	
PDS quant p.c.(t)	-0.008***		-0.006***		0.012**	
	(0.003)		(0.001)		(0.005)	
PDS quant p.c.(t+1)		-0.003		-0.002**		0.005
		(0.004)		(0.001)		(0.005)
Observations	1,113	828	1,113	828	997	741
Adj R-squared	0.800	0.821	0.945	0.939	0.992	0.994

Robust standard errors in parentheses clustered by district.
PDS quant p.c. (t+1) is the district PDS quantity from the next (future) period.
Additional controls include district and year fixed effects, poverty rate,
real per capita expenditure, and population.
*** $p<0.01$, ** $p<0.05$, * $p<0.10$

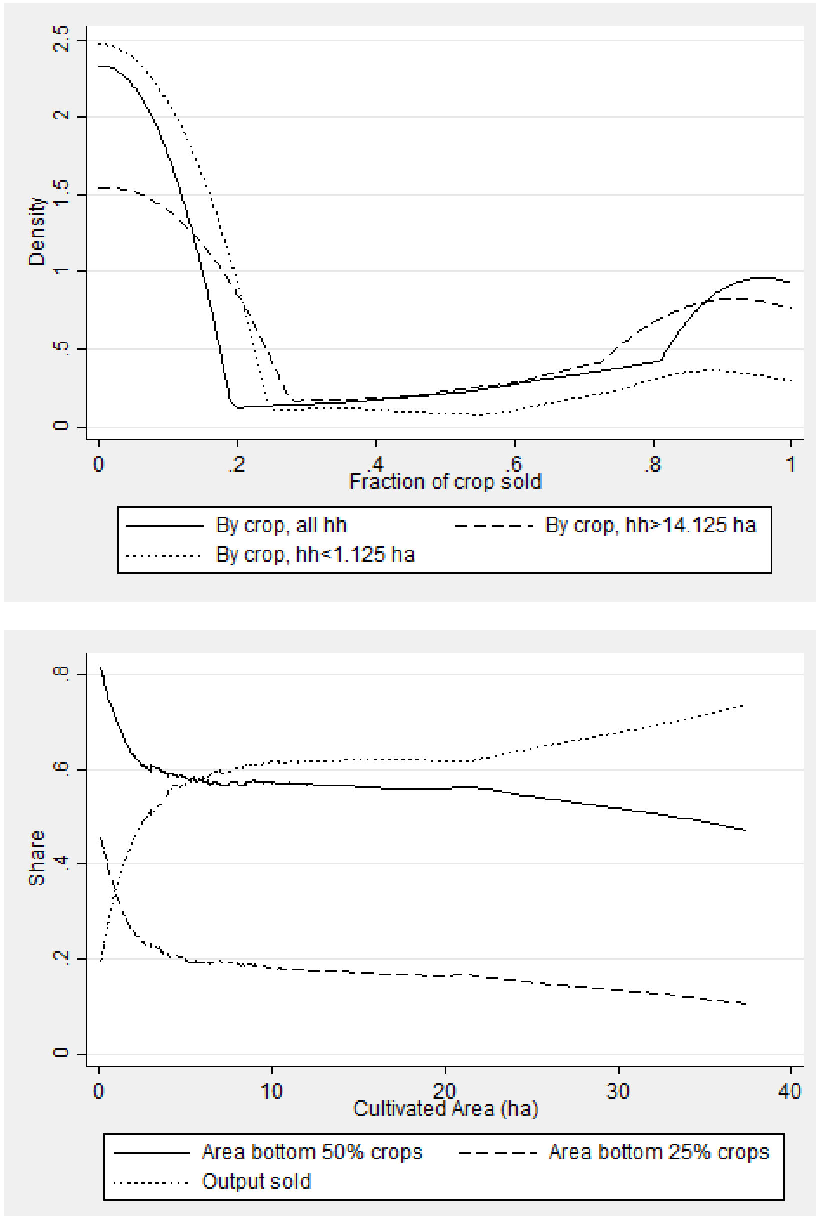

Figure 1: ICRISAT: Farmer characteristics and production

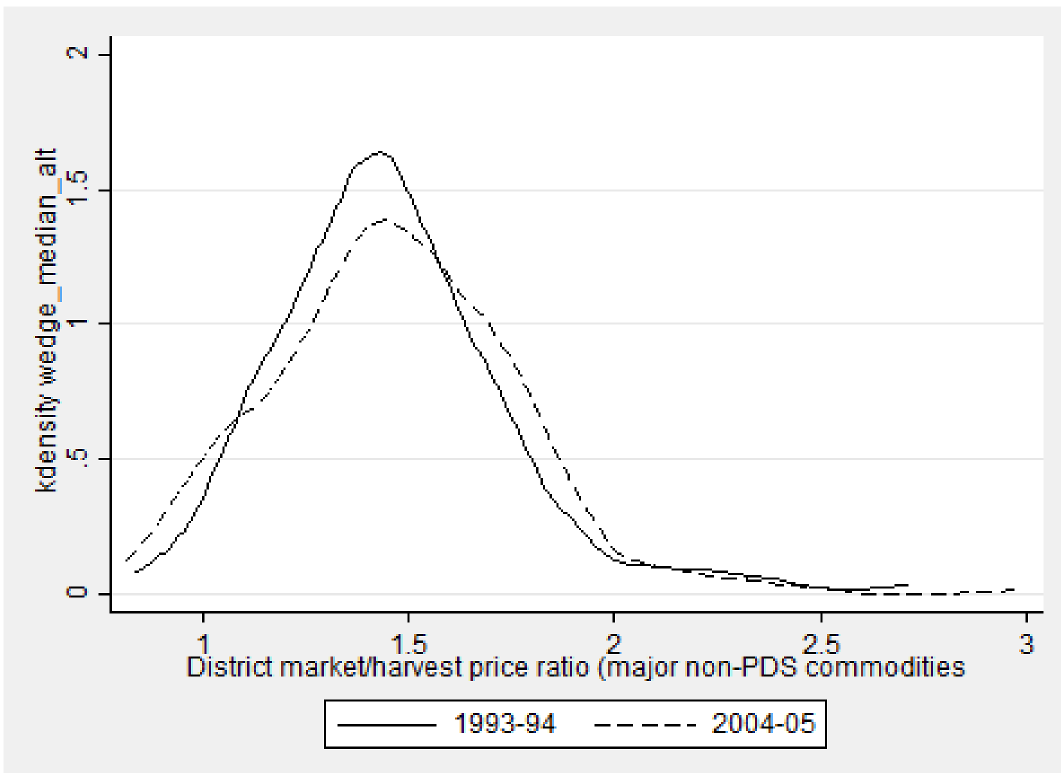

Figure 2: Distribution of market/harvest price wedges across districts (non-PDS major food items)

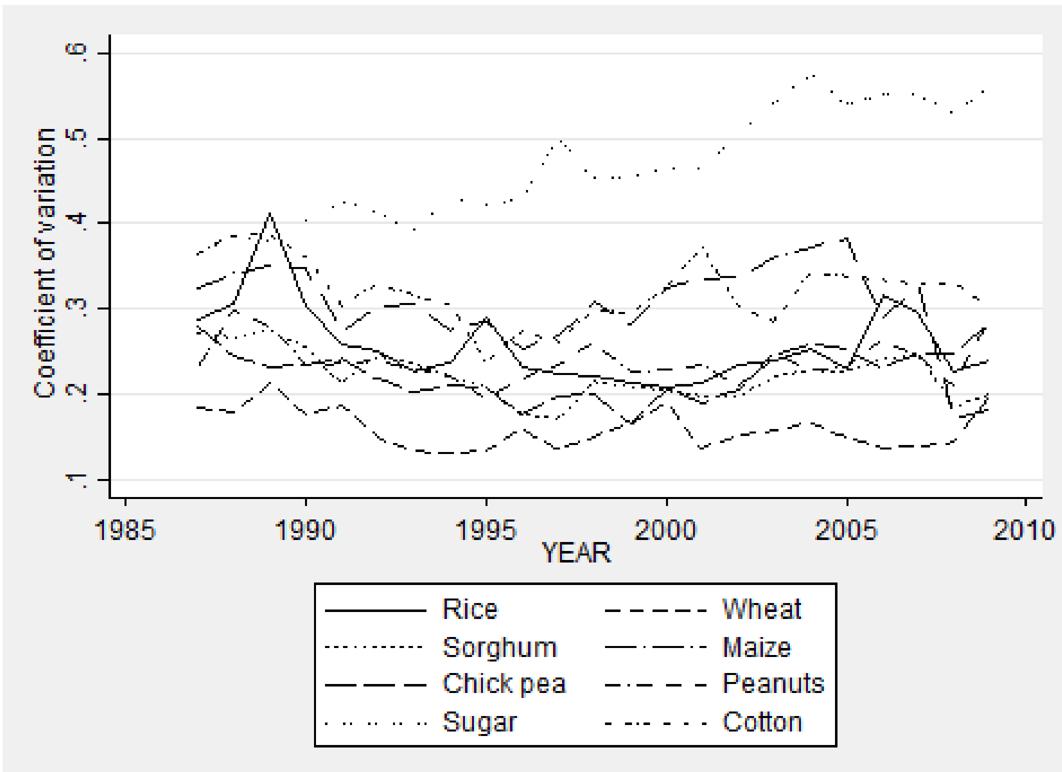

Figure 3: Price dispersion across districts by year for major agricultural commodities

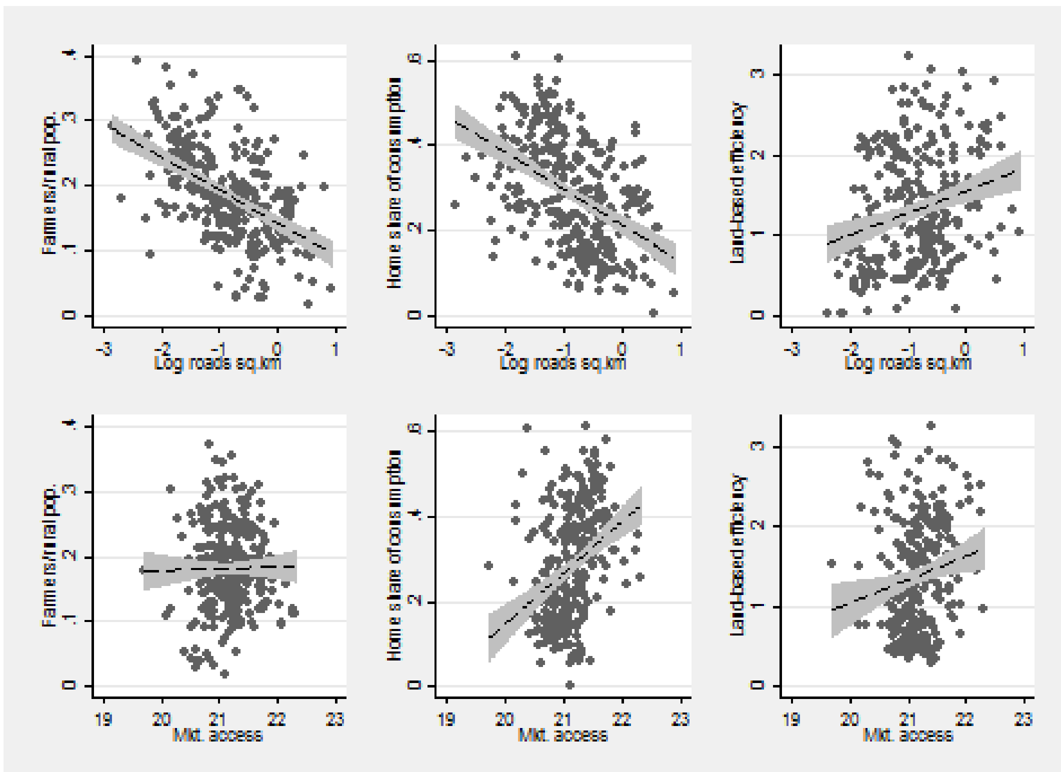

Figure 4: Correlations of internal and external market access with farmer specialization

A Additional Graphs and Figures

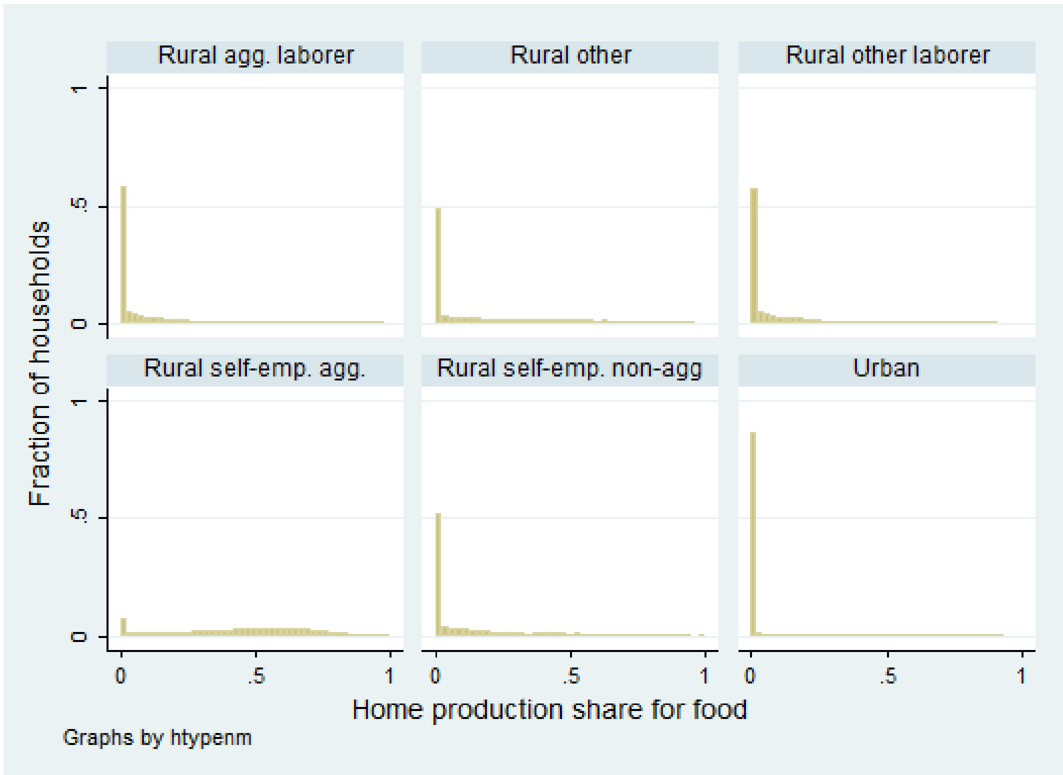

Figure 5: 1987-88: Distribution of home shares of food by household type

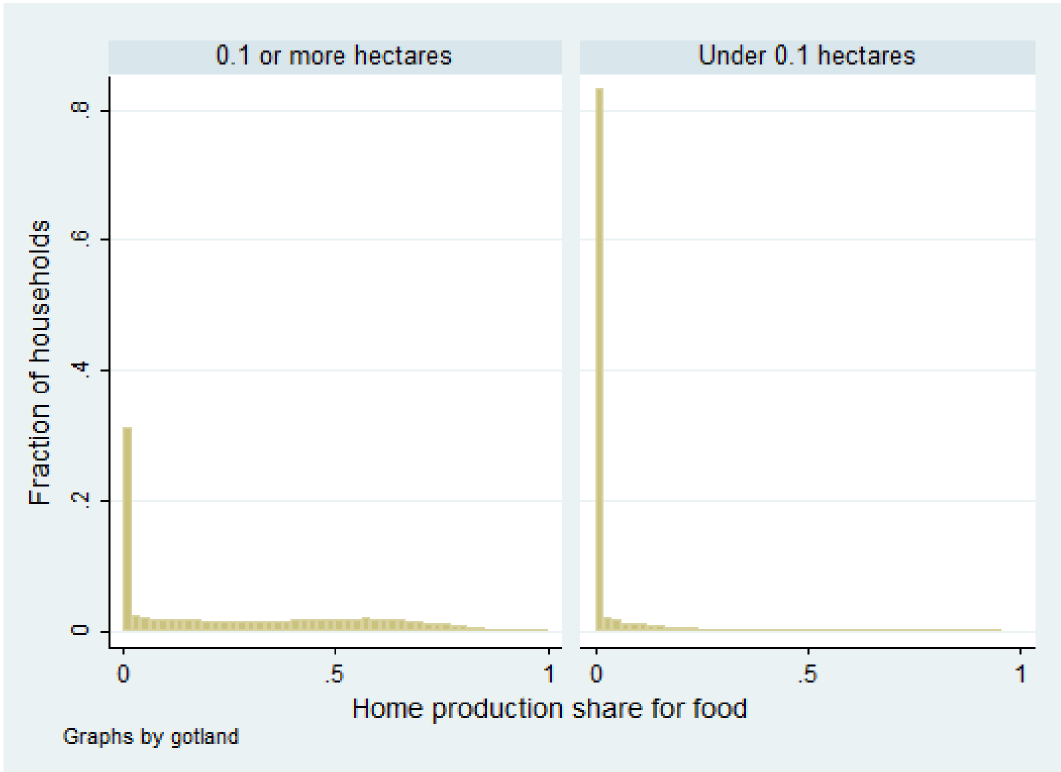

Figure 6: 1987-88: Distribution of home shares by land holdings

Table 9: District average price paid vs. farm-gate price

	1987-88				2009-10			
	Farmer	Ag.laborer	Rural other	Urban	Farmer	Ag.laborer	Rural other	Urban
Rice	1.01	0.98	1.04	1.14	0.95	0.76	0.91	1.11
Wheat	1.02	1.03	1.05	1.12	0.96	0.87	0.96	1.09
Chickpea	1.1	1.16	1.16	1.2	1.14	1.18	1.18	1.21
Pigeon Pea	1.1	1.12	1.16	1.22	1.13	1.15	1.17	1.21
Potato	1.3	1.29	1.33	1.36	1.29	1.3	1.3	1.33
Onions	1.26	1.27	1.28	1.26	1.38	1.42	1.4	1.39

All data from the Indian NSS.

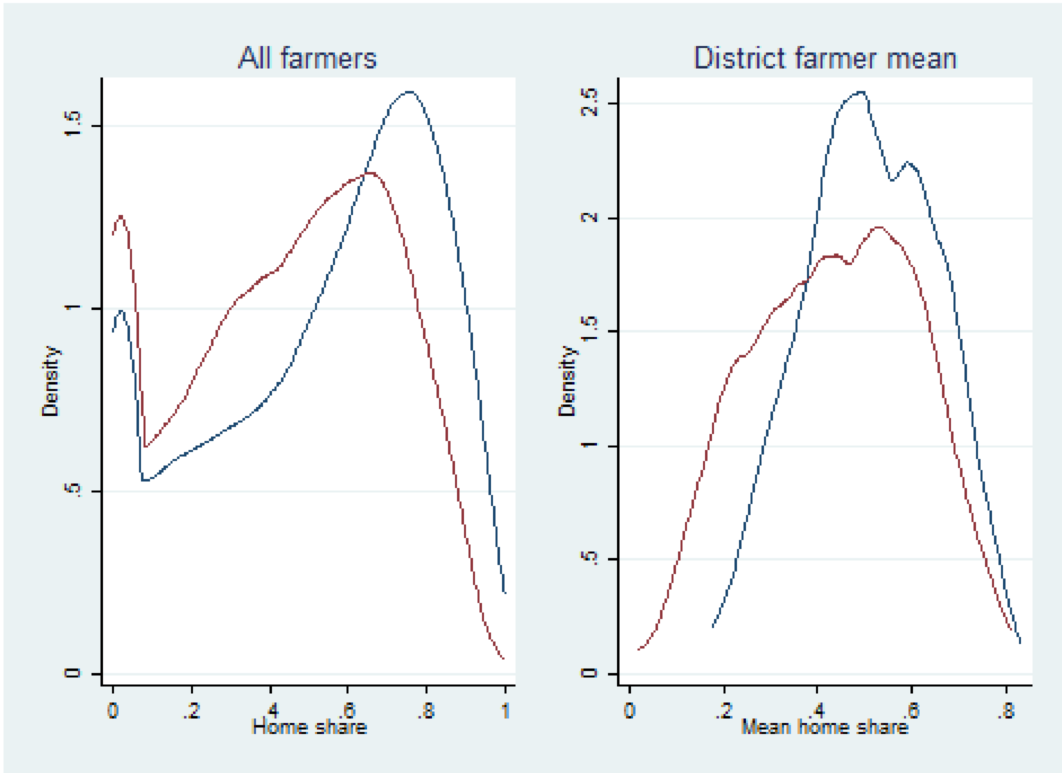

Figure 7: Dispersion and decline in food home shares. Left panel is individuals, right panel is district-level means. Blue line represents 1987, red line represents 2009.

Table 10: Sources of rice and wheat for households consuming an PDS or not

	Non-PDS households			PDS households		
Year	Mkt	Home	Neither	Mkt	Home	Neither
Additional sources of rice						
1987	0.72	0.18	0.10			
1993	0.65	0.24	0.11	0.69	0.06	0.25
1999	0.69	0.20	0.10	0.75	0.09	0.16
2004	0.72	0.19	0.08	0.67	0.06	0.27
2009	0.74	0.17	0.09	0.70	0.10	0.20
Additional sources of wheat						
1987	0.52	0.14	0.34			
1993	0.49	0.16	0.36	0.19	0.01	0.80
1999	0.50	0.16	0.34	0.37	0.04	0.60
2004	0.58	0.15	0.27	0.26	0.02	0.71
2009	0.61	0.16	0.23	0.37	0.06	0.57

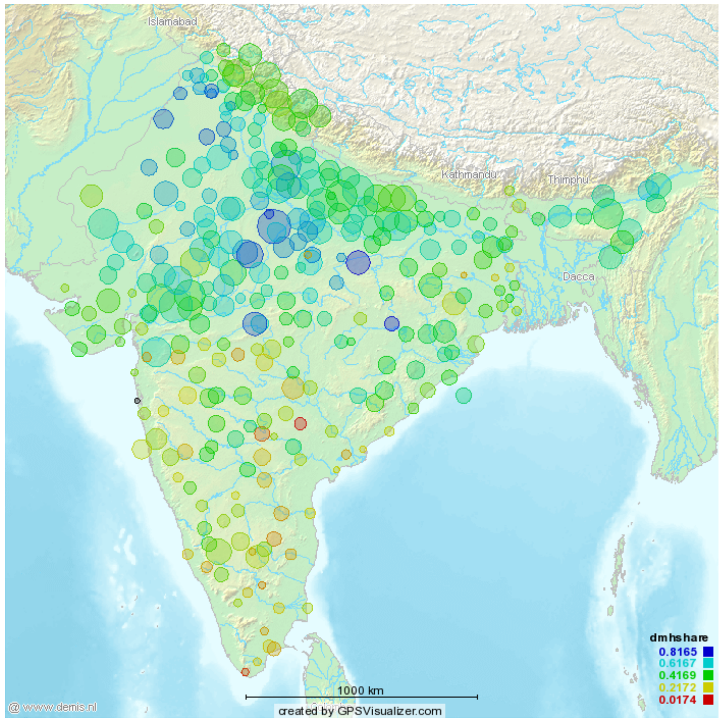

Figure 8: District mean farmer home shares 2009-2010. Size of dot corresponds to fraction of households who are farmers, color legend maps on to home shares.

Table 11: Household consumption (KG/grains per capita)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Grains	Both	Both	Both	Both	Both	Both	Both	Rice	Rice	Wheat	Wheat
PDS usage	0.463***	0.533***	1.018***								
	(0.042)	(0.040)	(0.042)								
BPL card				0.150***	0.201***	0.690***					
				(0.037)	(0.035)	(0.035)					
PDS quant. p.c.							0.265***				
							(0.007)				
PDS rice usage								0.654***			
								(0.037)			
PDS rice quant. p.c.									0.269***		
									(0.009)		
PDS wheat usage										0.766***	
										(0.034)	
PDS wheat quant. p.c.											0.297***
											(0.012)
Controls	None	Demog.	Inc.	None	Demog.	Inc.	Inc.	Inc.	Inc.	Inc.	Inc.
Observations	67,055	66,857	66,857	67,055	66,857	66,857	66,857	66,857	66,857	66,857	66,857
Adj R-squared	0.474	0.520	0.565	0.473	0.519	0.562	0.577	0.801	0.808	0.789	0.793

Robust standard errors in parentheses

$***p < 0.01, **p < 0.05, *p < 0.10$

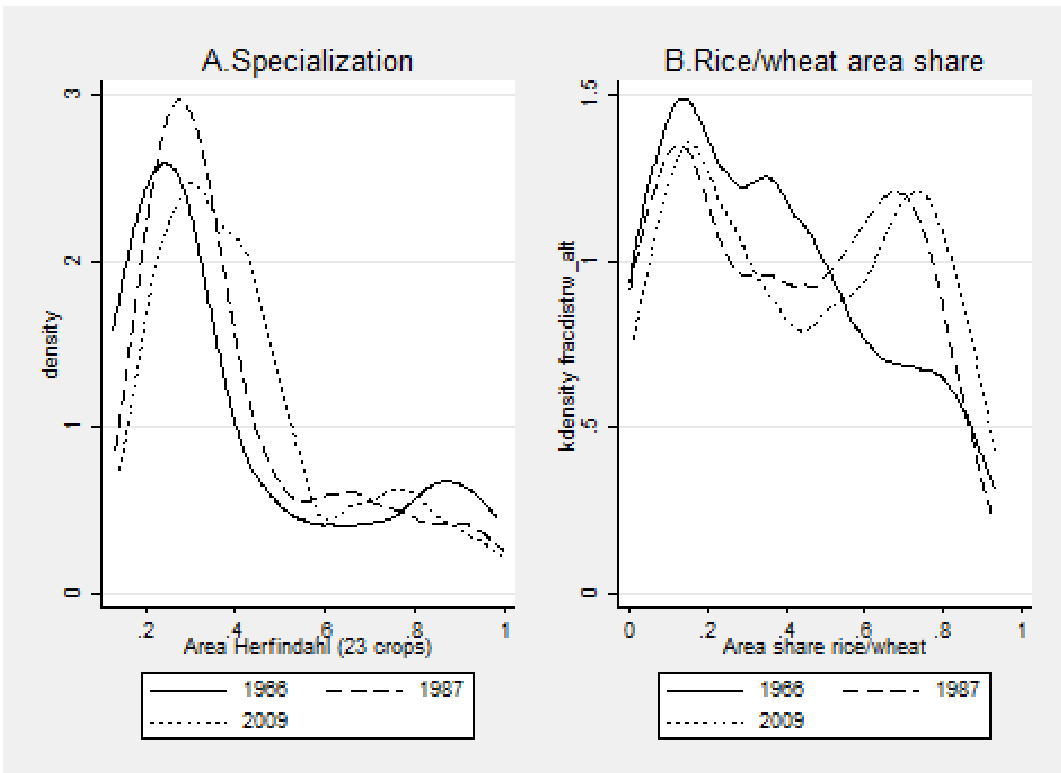

Figure 9: ICRISAT VDSA: A. Herfindahl index for land allocation and B. specialization in rice/wheat

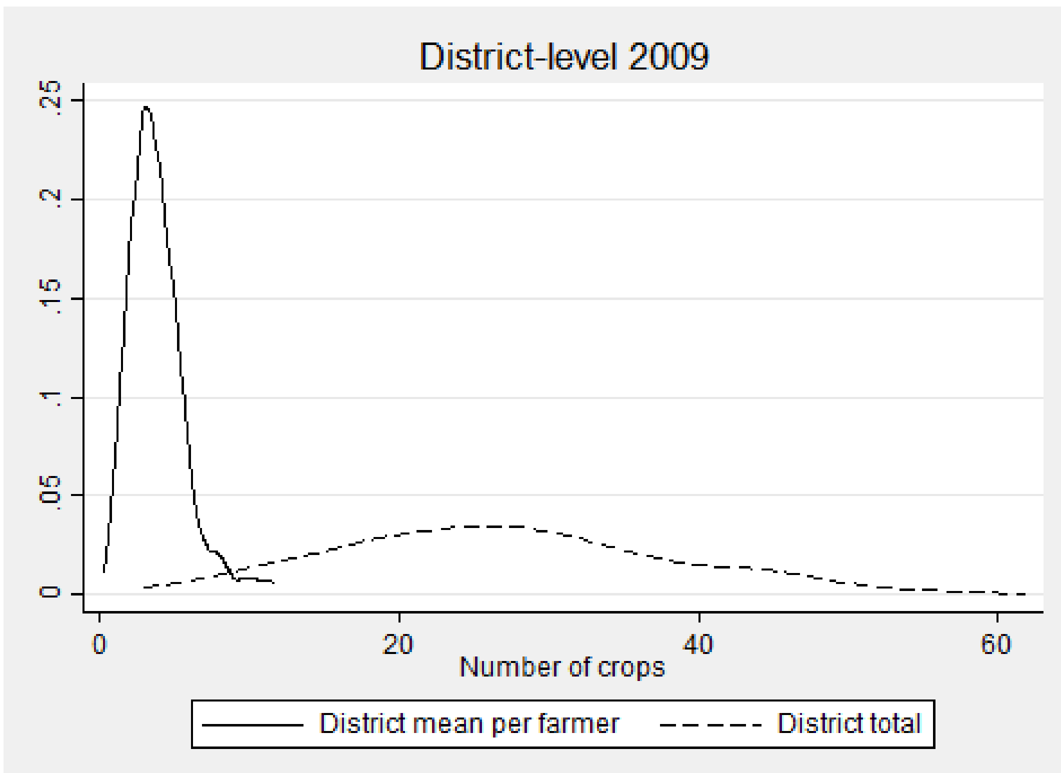

Figure 10: NSS 2009: District-level mean number of crops per farmer and total number of crops

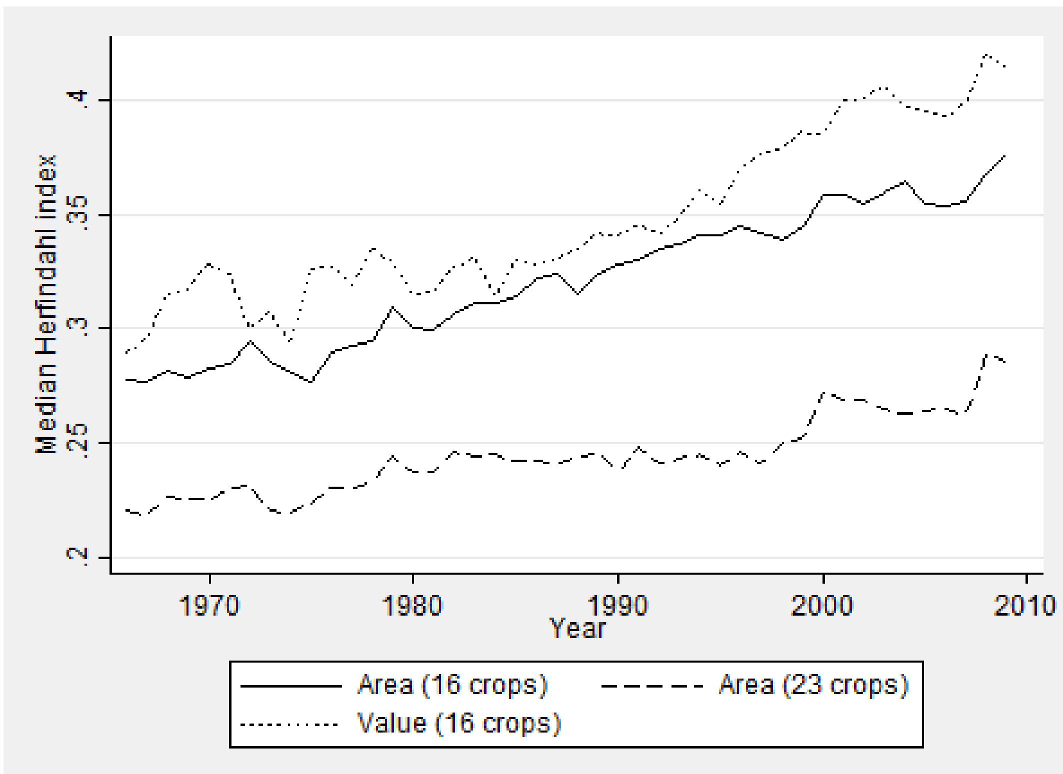

Figure 11: District-level specialization over time

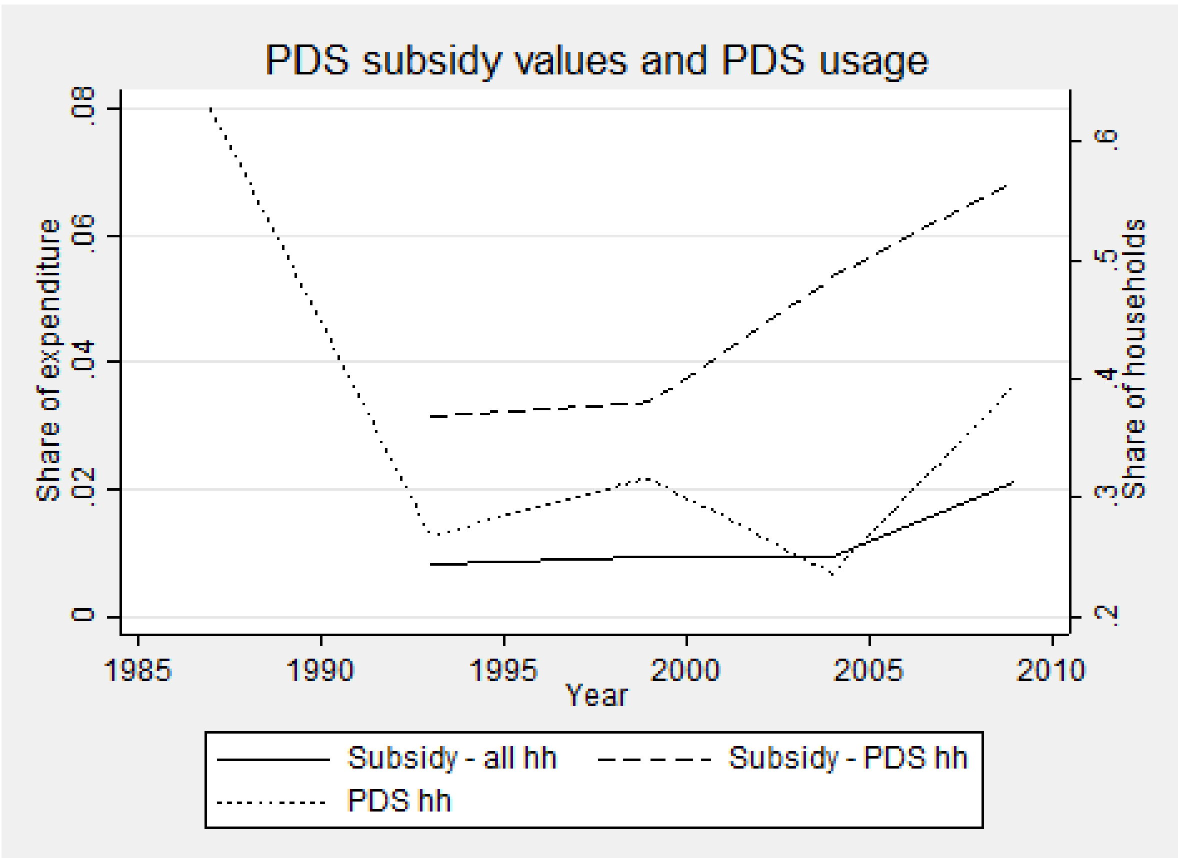

Figure 12: Value of PDS subsidy as share of consumption

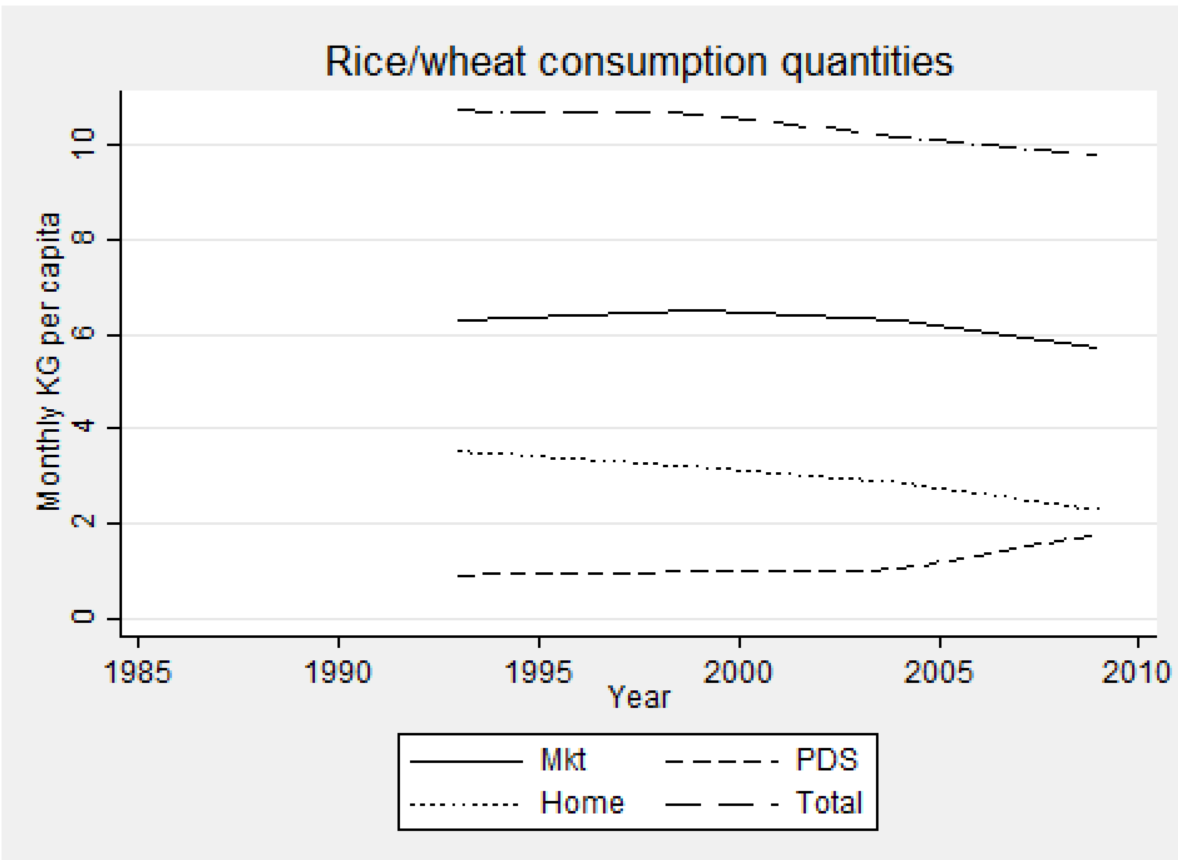

Figure 13: PDS quantities

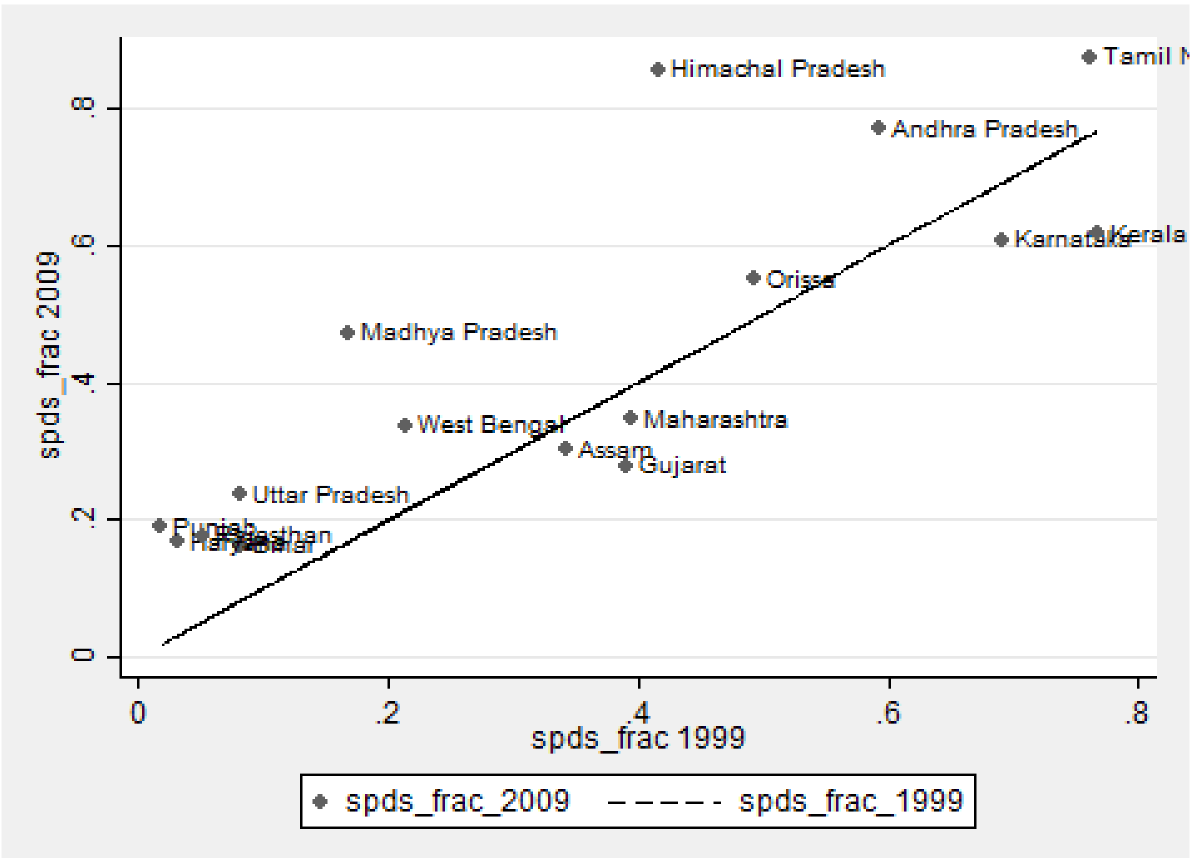

Figure 14: State-level variation in fraction of households served by PDS

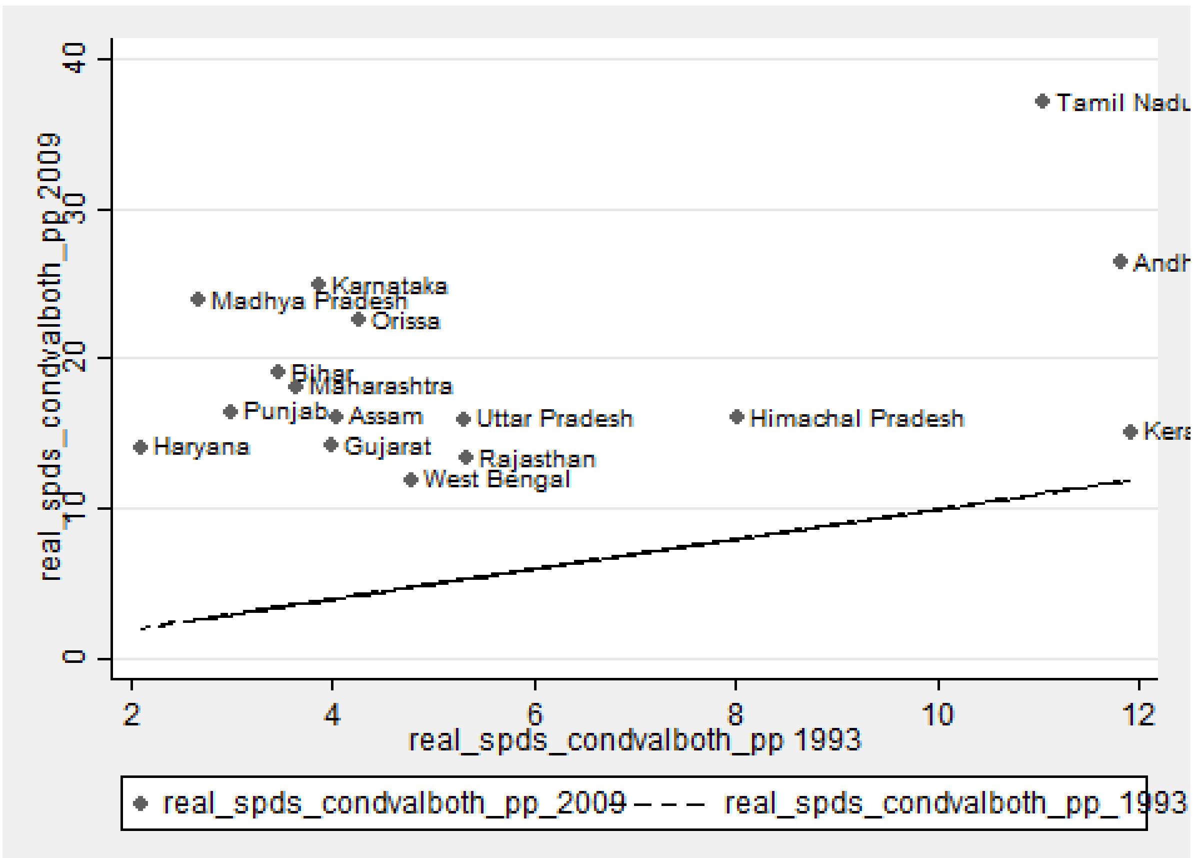

Figure 15: State-level variation in value of PDS subsidy for recipients

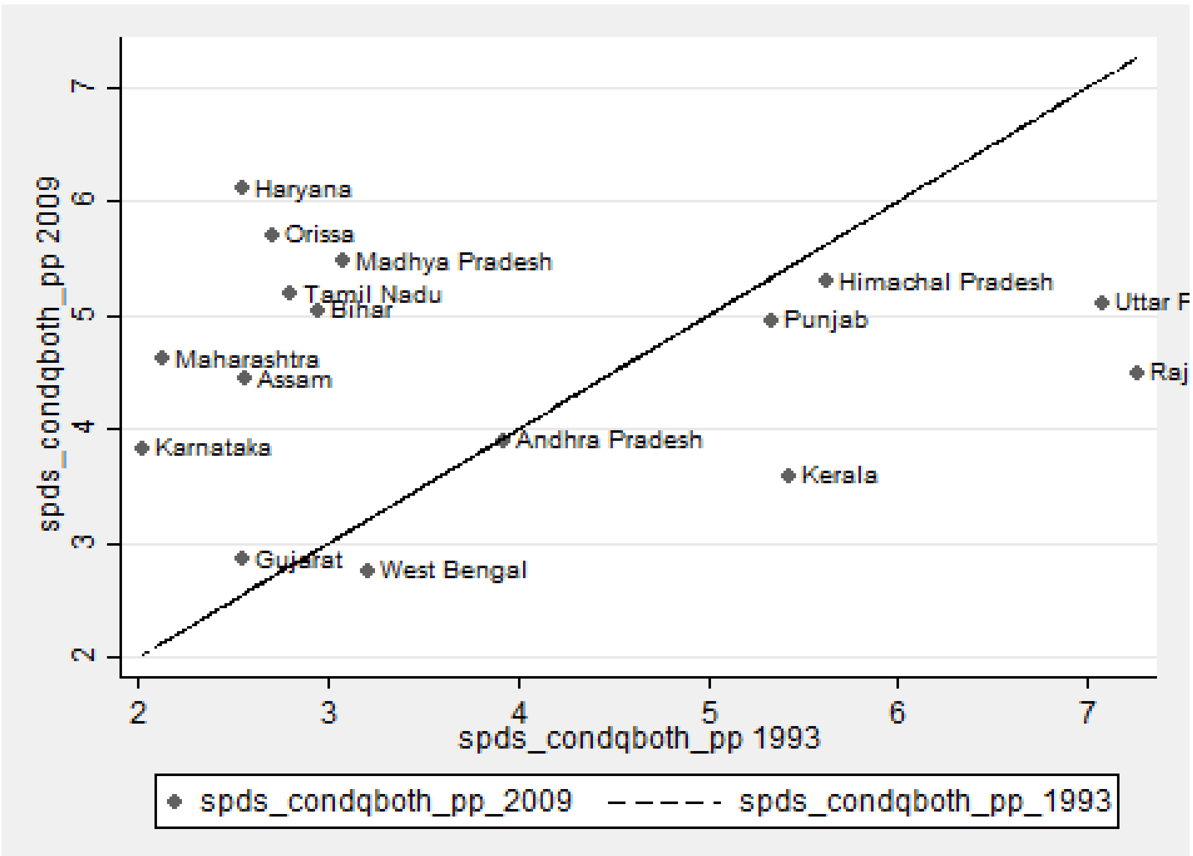

Figure 16: State-level variation in PDS quantity for recipients

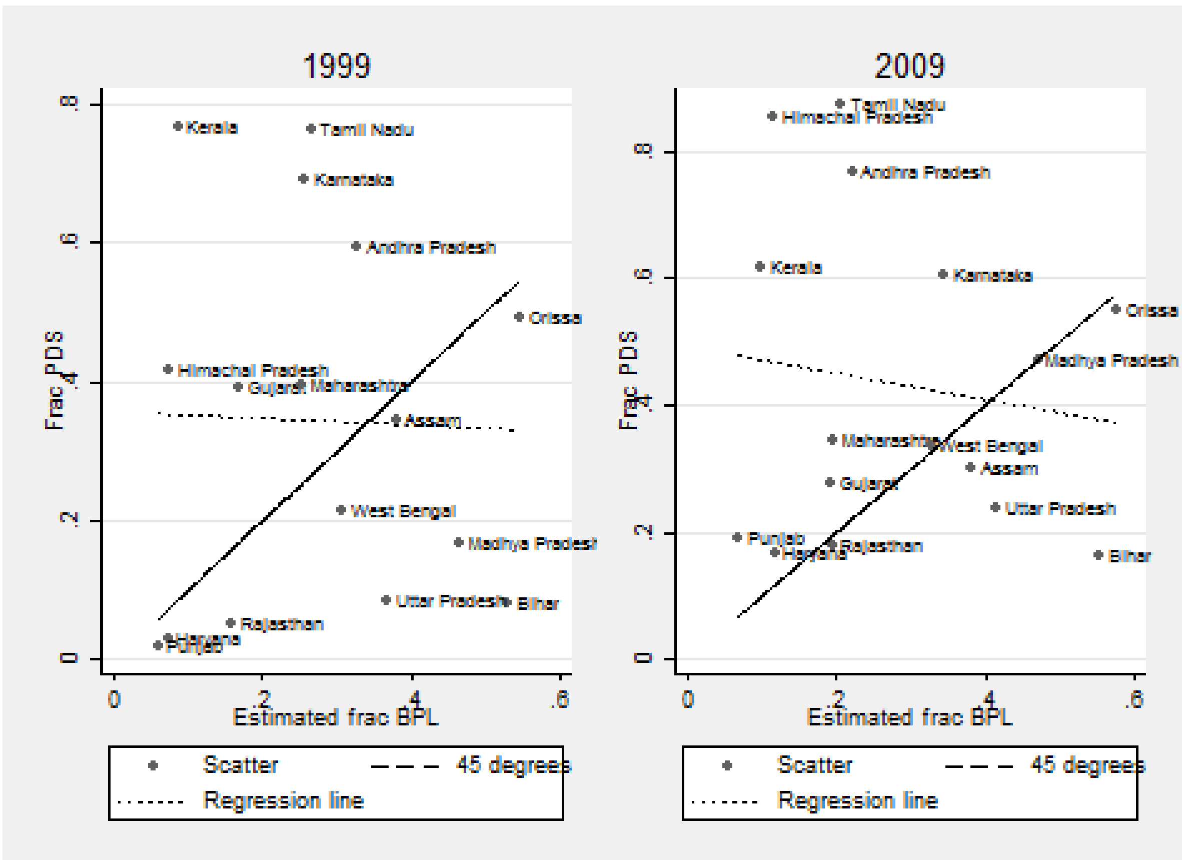

Figure 17: PDS targeting based on BPL households across states

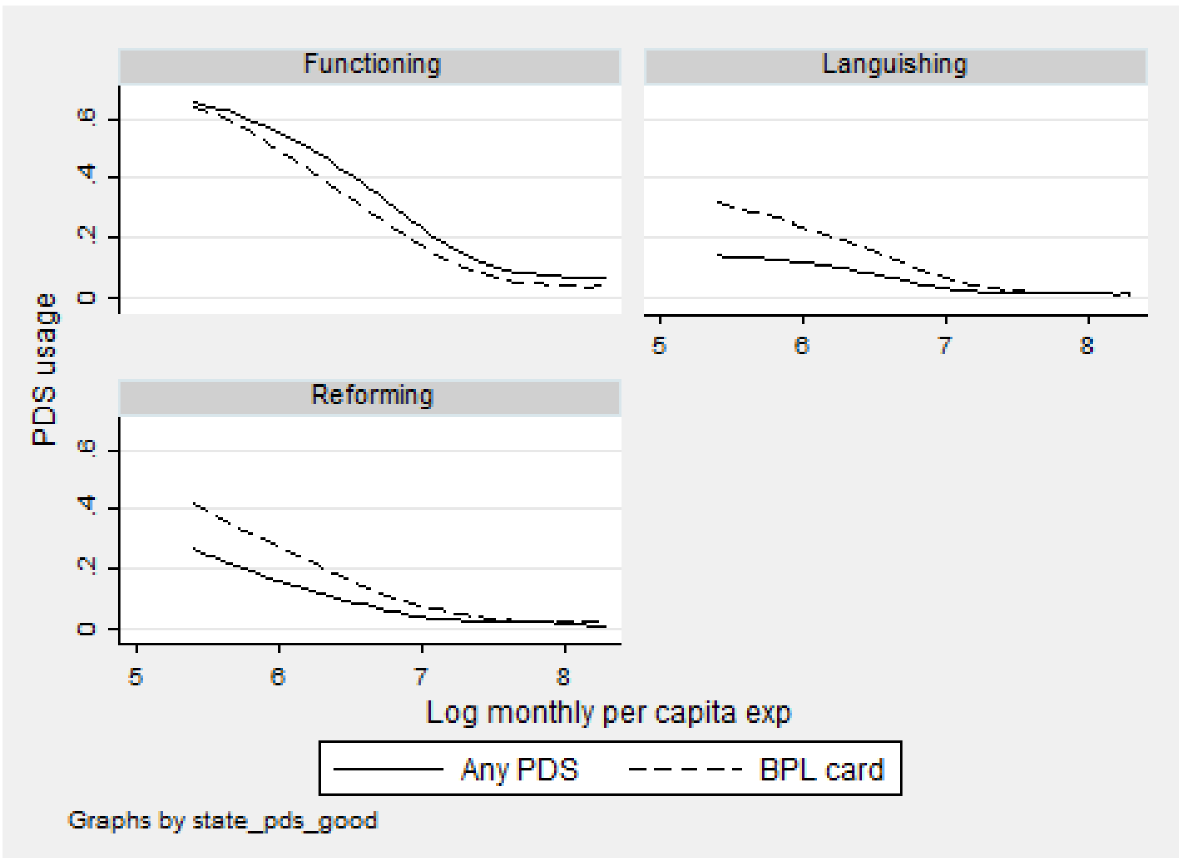

Figure 18: PDS usage and BPL card ownership by income for grouped states

Table 12: Household quantity of grains per capita

Grains	(1) Both	(2) Both	(3) Both	(4) Rice	(5) Rice	(6) Rice	(7) Wheat	(8) Wheat	(9) Wheat
PDS usage	-0.467***	-0.361***	0.068						
	(0.085)	(0.076)	(0.077)						
PDS x subsidy value	0.049***	0.048***	0.045***						
	(0.010)	(0.011)	(0.009)						
PDS rice usage				-0.402***	-0.433***	-0.124			
				(0.092)	(0.100)	(0.079)			
PDS rice x subsidy value				0.044***	0.042***	0.040***			
				(0.010)	(0.012)	(0.010)			
PDS wheat usage							0.802***	0.713***	0.569***
							(0.163)	(0.164)	(0.172)
PDS wheat x subsidy value							0.030**	0.031**	0.027**
							(0.011)	(0.012)	(0.011)
Observations	240,301	239,549	239,549	240,301	239,549	239,549	240,301	239,549	239,549
Adj R-squared	0.479	0.525	0.574	0.766	0.783	0.796	0.761	0.771	0.782

Robust standard errors in parentheses clustered by state
$***p < 0.01, **p < 0.05, *p < 0.10$

Table 13: District quantity of grains consumed per capita

Dep. var.	(1) All grains	(2) Home grains	(3) Market grains	(4) All rice	(5) Home rice	(6) Market rice	(7) All wheat	(8) Home wheat	(9) Market wheat
Real pds val pp	-0.010	0.000	-0.010						
	(0.008)	(0.007)	(0.009)						
PDS quant. p.c.	0.217***	-0.282***	-0.502***						
	(0.053)	(0.063)	(0.060)						
Real pds val rice p.c.				-0.010	0.004	-0.014			
				(0.009)	(0.007)	(0.009)			
PDS quant. rice p.c.				0.033	-0.478***	-0.489***			
				(0.065)	(0.089)	(0.077)			
Real pds val wheat p.c.							0.026	-0.001	0.027
							(0.031)	(0.025)	(0.027)
PDS quant. wheat p.c.							0.285***	-0.226**	-0.490***
							(0.108)	(0.094)	(0.106)
Observations	977	977	977	1,069	1,069	1,069	1,146	1,146	1,146
Adj R-squared	0.857	0.800	0.701	0.960	0.839	0.898	0.935	0.867	0.813

Robust standard errors in parentheses clustered by district
$***p < 0.01, **p < 0.05, *p < 0.10$