Whither the Pin Factory? Modern Food Supply Chains and Specialization in India*

Rajib Sutradhar† N. Chandrasekhara Rao‡ Marc F. Bellemare§

January 29, 2019

Abstract

What is the relationship between smallholder participation in modern food supply chains and their degree of economic specialization? We investigate this research question by looking at whether there is a systematic relationship between whether a household grows crops for a supermarket chain—participation in the supermarket channel—and how concentrated among fewer crops are that household's total landholdings and its output. Using longitudinal data from India, we find that there is no relationship between whether a household participates in the supermarket channel and that household's Herfindahl-Hirschman index for land, but there is a robust statistically significant and negative relationship between whether a household participates in the supermarket channel and that household's Herfindahl-Hirschman index for output. These findings hold true at both the extensive margin (i.e., when considering whether a household participates in the supermarket channel) and the intensive margin (i.e., when considering how much of a household's marketed output is sold through the supermarket chain). This suggests that instead of fostering a greater degree of specialization, participation in modern food supply chains might lead to greater diversification.

Key words: Agricultural Value Chains, Food Supply Chains, Supermarkets

JEL Classification Codes: L24, O13, O14, Q12

* Sutradhar would like to thank the Agricultural and Applied Economics Association's Uma Lele fund for a fellowship that allowed this collaboration to happen. Bellemare would like to express his gratitude to the National Institute of Food and Agriculture for funding this work through grant MIN-14-061. All remaining errors are ours.

† Assistant Professor, Omeo Kumar Das Institute of Social Change and Development, Guwahati, India.

‡ Professor, Institute of Economic Growth, University of Delhi, New Delhi, India.

§ Northrop Professor, Department of Applied Economics, and Director, Center for International Food and Agricultural Policy, University of Minnesota, Saint Paul, MN, United States.

1. Introduction

In the introduction to his *Inquiry into the Nature and Causes of the Wealth of Nations*, Adam Smith (1776, 1976) famously wrote of a visit he had made to a pin factory, wherein

> "[e]ach person … making a tenth part of forty-eight thousand pins, might be considered as making four thousand eight hundred pins in a day. But if they had all wrought separately and independently, and without any of them having been educated to this peculiar business, they certainly could not each of them have made twenty, perhaps not one pin in a day."

Since then, the notion that economic specialization—the phenomenon whereby economic agents focus on producing more each of a handful of commodities rather than a little bit of everything—is key to economic growth and development has been part and parcel of economics. This is presumably why, for instance, most farmers in OECD countries grow only one or two crops and, conversely, why it is not uncommon for farmers in developing countries to grow up to a dozen crops. For Smith, the level of economic development enjoyed by farmers in OECD countries would be a consequence of their specializing in one or two crops.

It is not unlikely, however, that economic development is not only a consequence but also a cause of specialization. This is perhaps best seen when considering agriculture in developing countries, where the typically low level of economic development is associated with credit, insurance, input, and other market failures which end up constraining one's ability to specialize. Thus, the causality between specialization and development is likely to run both ways.

In this paper, we look at whether participating in modern food supply chains is associated at all with a change in the degree of economic specialization. Specifically, we use longitudinal data from India to look at whether those farmers who sell some of their crops to a supermarket specialize in those crops which they grow for the supermarket value chain, whether they choose instead to generalize by adding those crops to those they already grow for themselves, or neither.

To do so, we look at the relationship between participation in the supermarket channel and two outcome variables—one a Herfindahl-Hirschman index measuring the concentration of a household's landholdings across the crops it grows, and the other a Herfindahl-Hirschman index measuring the concentration of a household's output[1]—and test whether the former is associated with either of the latter two variables. In both cases, if the hypothesis that higher levels of economic development—proxied by participation in modern food supply chains—leading to more specialization is true, we would expect to see a positive association between participation in the supermarket value chain and the outcome variables. If instead the hypothesis is false, and higher levels of economic development lead to diversification instead of specialization, we would expect to see a negative association between participation in the supermarket value chain and the outcome variables.

Starting in the early 1990s, supermarkets have become more widespread in developing countries (Reardon et al., 2003). Beyond the processed foods they typically sell, because supermarkets strive to offer consumers fresh produce—in our data, tomatoes, several varieties of gourds, cabbage, chilis, carrots, cauliflower, radishes, potatoes, cucumbers, eggplants, onions, spinach, okra, turnip, chickpeas, beans, lettuce, and long melon—of consistent quality, they put in place their own procurement systems, which involve centralized buying and contractual arrangements with farmers and traders. Thus, the rapid diffusion of supermarkets in developing countries has fostered a growing body of literature focused mostly on the welfare implication for smallholders of changes in the agricultural marketing system (Hernandez et al., 2007; Reardon et al., 2009; Rao and Qaim, 2011; Bellemare, 2012; Michelson et al., 2013; Andersson et al., 2015). Most of these studies have largely focused on the impact of supermarket procurement on household income (Rao and Qaim, 2010; Bellemare, 2012; Andersson, 2015), farm productivity (Neven et al., 2009), or household assets (Michelson, 2013). Generally, empirical findings indicate that while supermarkets tend to bypass smallholder farmers, those who have

[1] This refers to the quantity of a household's output, and not its value.

managed to participate in supermarket-driven modern food supply chains benefit through higher incomes, higher productivity, and more assets. A more recent strand of the literature has also noted positive effects of participation in modern food supply chains on the nutritional status or on the food security of grower households (Chege et al., 2015; Bellemare and Novak, 2017; Montalbano et al., 2018). Despite these benefits for smallholders of participating in modern food supply chains, and even though the notion of economic specialization is so fundamental to neoclassical economics, it is striking to note that there is little evidence on the link between participation in the supermarket channel and economic specialization. Recent research on market size and agricultural specialization shows that the degree of agricultural specialization in rural areas follows a U-shaped relationship with the degree of nearby urban market expansion (Emran and Shilpi, 2012).[2]

We contribute to the literature by looking at the relationship between participation in modern food supply chains and economic specialization. Theoretically, in a developing-country setting, it is not entirely obvious how participation in modern food supply chains affects economic specialization at the farm level. Typically, farmers chose a crop mix that maximizes profits evaluated at farm gate prices adjusted for transaction costs rather than profits evaluated at market prices (Omamo, 1998). When a supermarket chain enters a new region, farmers who otherwise grow subsistence crops are incentivized to diversify toward high-value cash crops. Thus, participation in modern food supply chain could lead to less specialization as farmers merely add those crops to an already diverse portfolio of crops. As a supermarket chain grows in size, however, its quality requirements get increasingly stringent, which means that the supermarket is likely to source a higher volume of each crop from fewer suppliers capable of meeting those higher standards. Thus, participation in modern food supply chains could lead to more specialization. Whether participation in modern food supply chains leads to more or less

[2] Relatedly, Qin and Zhang (2016) find evidence that road construction leads to specialization in China. Earlier, Kurosaki (2003) had found evidence that specialization obtains as a result of improvements in land productivity.

specialization may well be dependent on the context. In earlier stages of development, farmers might be more likely to diversify their cropping patterns in response to participation in the supermarket channel. After some time, however, they might be more likely to concentrate on fewer crops.

Rather than support the idea that modern food supply chains lead to economic specialization, our empirical results support the opposite. We find a negative association between a household's Herfindahl-Hirschman index for output and a household's participation in the supermarket channel at both the extensive margin (i.e., whether the household participates in the supermarket channel) and the intensive margin (i.e., how much the household participates in the supermarket channel), but no robust association between a household's Herfindahl-Hirschman index for land and a household's participation in the supermarket channel. This suggests that instead of leading to more economic specialization, participation in modern food supply chains appears to lead to diversification in our Indian data, at least when it comes to a household's crop portfolio.

The remainder of this paper is organized as follows. In Section 2, we present the data and some descriptive statistics. Section 3 discusses our empirical framework, discussing in turn our estimation and identification strategies. In section 4, we present and discuss our empirical results. Section 5 concludes with directions for future research.

2. Data and Descriptive Statistics

The data used in this study were collected in two rounds from villages near the city of Hyderabad, the state capital of Telangana in South India. The first phase of the survey, commissioned by the Centre for Economic and Social Studies (CESS), was implemented in 2009 in four *mandals*[3] spread over the two adjoining erstwhile districts of Rangareddy and Medak. The selected *mandals* cover most of the villages that supply vegetables to supermarkets in the city. We conducted a census in all major vegetable-

[3] A *mandal* is an administrative division below the district level in Telangana and Andhra Pradesh. Outside of those two states, it is known as a *tehsil*.

producing villages in each *mandal* to identify villages with high and low intensity of participation in the supermarket channel. In each *mandal*, we randomly selected four high-intensity villages and one low-intensity village. In each high-intensity village, we randomly selected 10 supermarket farmers (i.e., farmers who participate in the supermarket channel) and five traditional farmers (i.e., farmers who do not participate in the supermarket channel, and who sell at traditional markets instead) from the census list. Likewise, in each low-intensity village, we randomly selected 10 traditional farmers from the census list. The household survey, however, could only be implemented in 19 villages. After dropping incomplete questionnaires, we ended up with a sample of 254 farmers, which included 150 supermarket farmers and 104 traditional market farmers.

The second phase of the survey, commissioned by Indian Council of Social Science Research and implemented by a team of researchers from Institute of Economic Growth, was conducted in 2014. The idea behind the second-round survey was to resample the households interviewed in the first round so as to generate evidence on the dynamics of household participation in the supermarket channel. The second-round census results show a high attrition rate, as only 39 out of 150 supermarket farmers interviewed in the first round were still selling vegetables to the supermarkets. Further, as many as 21 farmer households from the previous sample quit vegetable cultivation altogether and thus can no longer be part of the sample for the second-round survey.[4] To ensure an adequate sample of supermarket farmers in the second round of survey, the survey team randomly selected additional supermarket farmers from census in high-intensity villages. Additionally, the team randomly selected six

[4] This known in the literature on agricultural value chains as "survivorship bias" (Ton et al., 2018; Bellemare and Bloem, 2018) whereby the observations included in a sample of households that participate in agricultural value chains are those successful households, i.e., those who chose to keep participating in value chains. By introducing an issue of sample selection, the high number of attritors in our study seriously limits the external validity of our findings in that they only apply to similar households, i.e., households that would remain participants in supermarket value chains. That said, the high number of attritors does not limit the external validity of our findings. If our identification strategy is valid, however, the internal validity of our findings is not affected by the high number of attritors.

additional high-intensity villages. In each of these newly selected villages, a census of households was conducted to draw a random sample of 10 supermarket farmers and five traditional market farmers to interview.

Ultimately, we have a sample of 359 farmers, which includes 222 traditional market farmers and 137 supermarket farmers. In each round, the survey team implemented a pre-tested questionnaire to collect information on household characteristics, farm assets, and crop-specific information on inputs, output, and marketing for the agricultural year that ended immediately before the survey was implemented. The data used in this paper thus consist of an unbalanced panel of 613 observations across two rounds, with only 233 farmers appearing in both rounds of surveys.

A brief discussion on the nature of contractual arrangement is in order to set the background for the study. In India, the diffusion of supermarket is a relatively recent phenomenon that started only in the late 1990s (Reardon et al 2003). Quite early into their diffusion, however, most supermarket chains have already started sourcing directly from farmers (see for example Rao et al., 2017, Pritchard et al., 2010, and Chen et al. 2005). Many supermarket chains, both national and regional, have been operating in the survey region. Among them, Reliance Fresh, the largest supermarket chain in India, set up its first retail outlet in the country in the city of Hyderabad in 2006. The supermarket chain started procuring directly from the farmers in the region early on. There are other national level supermarket chains such as Big Bazaar, ITC and More that figure prominently in the region. Heritage Fresh, a regional chain, is also present in the region. Most of these supermarket chains followed the example of Reliance Fresh, setting up collection centers, typically managed by a local agent and his assistant, to procure fresh produce directly from farmers. In a slightly different variant, Big Bazaar procures fresh produce from farmers through a lead farmer who gets 1% of the total volume of transaction conducted through him as commission for his services. Most of the supermarket chains follow “listing,” which means inclusion of farmers in the company vendor list, and payment is settled through a farmer’s bank accounts upon

delivery of the contracted produce. There is no prior commitment on either party to buy or sell produce. Supermarket chains make no provision of resources for their supplying farmers nor do they offer any extension advices to the listed farmers. Prices are set based on the prevailing prices on the day of the transaction, either at par or with a premium for better-quality produce.

The contract is an informal arrangement, with no commitment from either party to buy or sell their produce. Such arrangement seems to work to the advantage of both farmers and supermarket managers. With prices set early in the day based on prevailing prices in the wholesale market, farmers can decide how much to sell to the supermarket collection center. Moreover, when farmers dispose of their produce at the farm gate, it reduces the risk that they face when traveling long distance to sell their produce on the wholesale market (Michelson et al 2012).

The quality criteria used by supermarket manager are still at an early stage. A case study of Delhi shows that the quality of produce sold in supermarkets is only slightly better or comparable to that of produce sold through traditional marketing systems (Minten et al., 2010). Such quality criteria may not necessitate resource provision contracts from the supermarket manager.

Moreover, in the present Indian context, a formal contract is difficult to enforce legally, which means parties tend to sign informal contracts, with fewer obligation on both sides (Narayanan, 2012).

Table 1 presents descriptive statistics for the variables we retain for analysis. Our two dependent variables are Herfindahl-Hirschman indices (HHIs)—one for land, one for output—which respectively aim at measuring the degree of crop concentration for the lands cultivated by the household and the degree of crop concentration for the household's total output. We use both measures as proxies for how much a household's lands are specialized in growing fewer crops (in the limit, only one crop) as well as for how much of a household's output is specialized in fewer crops (in the limit, only one crop).

We compute those HHIs as follows. The cultivation of each crop $j \in \{1, \dots, N_i\}$ grown by household *i* occupies a share $s_{\ell j}$ of the household's total cultivated landholdings. Thus, household *i*'s HHI for land $H_{\ell i}$ is such that

$$H_{\ell i} = \sum_{j=1}^{N_i} s_{\ell j}^2. \quad (1)$$

Likewise, the cultivation of each crop $j \in \{1, \dots, N_i\}$ grown by household *i* occupies a share s_{oj} of the household's output. Thus, household *i*'s HHI for output H_{oi} is such that

$$H_{oi} = \sum_{j=1}^{N_i} s_{oj}^2. \quad (2)$$

Both $H_{\ell i}$ and H_{oi} are in the $(0,1]$ interval, with $H_{\cdot i} \to 0$ meaning many crops with small shares, and with $H_{\cdot i} = 1$ meaning a single crop occupying the entirety of a household's cultivated landholdings or of its output.

Table 1 shows that the households that participate in the supermarket channel have a lower HHI for land than traditional households, but that there are no differences between the two types of households when it comes to the HHI for output.

Table 1 also shows the number of friends and relatives a respondent has who are working as local agents of a supermarket supply network. This number is substantially higher for those households that participate in the supermarket channel (0.69 such friends and relatives) than for those households that do not (0.08 such friends and relatives). Additionally, households that participate in the supermarket channel tend to have a household head who is slightly younger than households that do not, and they tend to be significantly wealthier as well.

Finally, supermarket farmers have a lower level of off-farm employment compared to their traditional market counterparts; this is consistent with the findings in Bellemare (2018), according to

which households that participate in contract farming derive lower levels of income from labor markets and nonfarm businesses.

3. Empirical Framework

In this section, we first present the equations we use to study the relationship between participation in the supermarket channel and specialization. Because our data are observational, and thus do not allow making causal statements, we next discuss the identification strategy we use to help disentangle the potential causal relationship flowing from supermarket channel participation to specialization from their correlation.

3.1. Estimation Strategy

We estimate two core equations to study the relationship between participation in the supermarket channel and economic specialization. The equation is specified as

$$H_{\ell it} = \alpha_{\ell} + \beta_{\ell} x_{it} + \gamma_{\ell} D_{it} + \delta_{\ell i} + \epsilon_{\ell it}, \quad (3)$$

where $H_{\ell it}$ denotes household *i*'s HHI for land in year *t*, *x* is a vector of control variables, *D* is our variable of interest (i.e., a dummy variable for whether the household participates in the supermarket channel), δ is a vector of household fixed effects, and ϵ is an error term with mean zero.

The second equation is specified as

$$H_{oit} = \alpha_{o} + \beta_{o} x_{it} + \gamma_{o} D_{it} + \delta_{oi} + \epsilon_{oit}, \quad (4)$$

where all variables are defined as in equation 3 except for the dependent variable, which denotes household *i*'s HHI for output.

We estimate four specifications of equations 3 and 4. The first specification (OLS) treats our data like a cross-section and pools all observations together, with no regard as to whether two observations

belong to the same household over time. The second specification (2SLS) treats the data like a cross-section but instruments our variable of interest in an attempt at exogenizing it; we defer our discussion of the instrumental variable to the next sub-section. The third specification (FE) takes into account the longitudinal nature of the data by controlling for household fixed effects, as shown in equations 3 and 4. The last specification (FE-2SLS) both takes into account the longitudinal nature of the data and instruments our variable of interest and is thus our preferred specification.

3.2. Identification Strategy

Any identification strategy should be judged relative to how it fares in the face of (i) unobserved heterogeneity, (ii) reverse causality, and (iii) measurement error. In this section, we explain how our preferred specification (i.e., FE-2SLS) fares in the face of those three sources of statistical endogeneity as we attempt to control for household selection into participating in the supermarket channel. Then, because there is selection on both sides and supermarkets select the households they work with in addition to the households selecting into the supermarket channel, we discuss grower selection by the supermarkets.

Throughout this discussion, it helps to keep in mind that our preferred FE-2SLS specification controls for everything that is time-invariant at the level of the household by incorporating household fixed effects and by instrumenting the supermarket channel participation dummy with a measure of how many people in a farmer's social network work as agents in the supermarket supply network. The logic behind that instrumental variable (IV) is as follows. On whether the IV is correlated with our variable of interest, which is a testable requirement for a valid IV, we expect that as the number of supermarket channel agents in one's social network increases, one is more likely to participate in the supermarket channel since one is likely to have more information about the supermarket channel the more of one's friends and relatives work in the supermarket channel. In addition, we expect the IV to have a

monotonic effect on participation in the supermarket channel, meaning that as the number of supermarket channel agents in one's network grows, one becomes ever more likely to participate in the supermarket channel.[5]

On whether the IV is plausibly exogenous, which is not testable, we argue—but can obviously not prove beyond any reasonable doubt—that the number of supermarket channel agents in one's social network only affects how concentrated one's cultivated landholdings and one's output are across crops via whether one participates in the supermarket channel. Ultimately, the estimate generated by our preferred FE-2SLS specification is a local average treatment effect (LATE). Here, this estimate tells us the effect of participation in the supermarket channel on the HHIs we use as outcome variables for those household-year observations that were induced to participate in the supermarket channel by the number of friends and relatives one has who work as agents in the supply network of a supermarket chain. As with many other instances of LATEs, it is unfortunately not possible to determine which household-year units of observations were induced to take up the supermarket channel treatment by our IV, i.e., the number of friends and relatives they have who work as agents in the supermarket channel. More importantly, this means the coefficients estimated by our 2SLS and FE-2SLS specifications are not directly comparable with their OLS and FE counterparts, given that the former are more limited in their external validity than the latter.

Our preferred FE-2SLS specification largely takes care of *unobserved heterogeneity* by controlling for household fixed effects. Indeed, because several of the households in our data are observed in both rounds, the use of household fixed effects allows controlling for everything that remains constant for a household between the two rounds we consider. This helps purge the error term of a great deal of prospective endogeneity.

[5] Recall that monotonicity of the relationship between the endogenous variable and the IV is required to properly identify the local average treatment effect (LATE).

Regarding *reverse causality*, although it is a priori not impossible that the expectation of a higher or lower degree of crop concentration on one's landholdings or of one's total output might induce some households to participate in the supermarket channel, we deem that possibility to be improbable. Indeed, households are almost surely driven by the prospect of improved welfare outcomes (e.g., higher income, lower income variability, improved food security), but it is unlikely that a desire for specialization (or its opposite) drives participation in the supermarket channel.

Regarding *measurement error*, it is also unlikely that our variable of interest—an indicator for whether the household participates in the supermarket channel—is measured with systematic error that would bias our estimate of the LATE in the FE-2SLS specification. Indeed, respondents have little incentive to lie one way or the other, and participation in the supermarket channel is important enough for the households we study that they would not forget about it. There remains the possibility of classical measurement error—say, because of errors in data entry. If that is the case, it merely introduces attenuation bias, which would make a rejection of the null either way more meaningful.

There remains one potential source of endogeneity—one that is context-specific. In studies such as this one, where one party (i.e., the households in our data) contract with another (i.e., supermarket chains), selection operates both ways, and though households select into the supermarket channel, supermarket exercise a certain amount of discretion about whom they choose to contract with or not. As in Bellemare (2012), however, we make the case that the selection of households by the supermarket chain is accounted for by the observables we include on the right-hand side (RHS) of equations 3 and 4. In other words, we argue that this is accounted for by selection on observables, and that we (as econometricians) have access to more—and more precisely measured—information than the supermarket chains do, and so the RHS variables in equations 3 and 4 fully account for how supermarket chains select the households they contract with.

4. Results and Discussion

Estimation results for OLS, FE, 2SLS, and FE-2SLS specifications of equations 3 and 4 are presented in tables 2 to 5. In tables 2 and 3, we present estimation results for the land HHI, which measures how crop-concentrated a household's landholdings are, at the extensive margin (i.e., using an indicator for whether the household participates in the supermarket channel as our variable of interest) in table 2 and at the intensive margin (i.e., using the share of a household's output that is sold to the supermarket channel as our variable of interest) in table 3. Those results suggest that participation in the supermarket channel is not systematically related to how concentrated a household's landholdings might be among fewer crops.

Tables 4 and 5 present estimation results for the output HHI, which measures how crop-concentrated a household's output is, respectively at the extensive and intensive margins. These results tell a very different story than those in tables 2 and 3. Indeed, tables 4 and 5 show a robust negative relationship between a household's output HHI and whether it participates in the supermarket channel (table 4) and the share of its marketed output it sells to the supermarket channel (table 5). Our preferred specification in table 4—that in column 4, or the FE-2SLS specification—suggests that for those households whose participation in the supermarket channel was induced by their social network, participation in the supermarket channel is associated with a 0.17 decline in a household's HHI for output, or a decline of almost 38 percent in the average household's HHI for output, given the average HHI for output of 0.45 in the sample. Likewise, our preferred specification in table 5—again, the FE-2SLS specification in column 4—suggests that for those households whose participation in the supermarket channel was induced by their social network, a 10-percentage point increase in the share of its output it sells to the supermarket channel is associated with a decrease in its output HHI of 0.05, though the F-statistic on the IV in column 4 of table 4 indicates that the FE-2SLS coefficient estimate for the variable

of interest is biased toward the OLS coefficient, i.e., with a stronger IV, we would likely see a coefficient of larger magnitude.

Instead of specialization, the results in tables 4 and 5 thus suggest that participation in the supermarket channel is associated with *diversification*, as households who participate—or who participate more intensively—in the supermarket channel see their output spread out over more crops than those households that do not participate in the supermarket channel or those who participate less intensively.

The results in tables 4 and 5 may seem paradoxical when contrasted with those in tables 2 and 3. After all, how can participation in the supermarket channel be associated with the concentration of crops in terms of output, but not in terms of landholdings? Here, there are two possibilities. One possibility is that our estimates in tables 4 and 5 capture something which we should in principle be able to capture in tables 2 and 3, but that the dependent variable in the latter tables is noisier than in the former. Given the much lower coefficient magnitudes in tables 2 and 3 relative to tables 4 and 5, we deem this possibility to be unlikely.

Another possibility is that participation in the supermarket channel is indeed associated with less specialization in terms of output, but not in terms of land. Intuitively, this would mean that the crops grown for the supermarket channel are simply added to what the household already does in terms of its output, but not in terms of its landholdings. One way to resolve the paradox here is to consider the following: In table 1, total landholdings are slightly larger (albeit not statistically significantly so) for households that participate in the supermarket channel than for those that do not. As such, it is likely that the crops grown for the supermarket channel are compensated for in the HHI for land by a commensurate increase in a household's landholdings.

5. Summary and Concluding Remarks

Starting from the observation that the lack of economic specialization can constrain economic development, we have looked at the relationship between participation in agricultural value chains and the degree of crop concentration on a household's landholdings as well as in the household's total agricultural output. Specifically, using longitudinal data from India, we have looked at whether selling one's output to a supermarket chain—what we refer to as participation in the supermarket channel—is associated with significant changes in a household's Herfindahl-Hirschman index of crop concentration within (i) a household's total landholdings, and (ii) that household's output, looking at the relationships at both the extensive and intensive margins. In doing so, we rely on both household fixed effects as well as an instrumental variable to identify the local average treatment effect of participation in the supermarket channel on the degree of concentration of a household's crop.

Our results show no statistically significant relationship between the HHI for crop concentration for a household's landholdings and participation in the supermarket channel at either the extensive or the intensive margin, but a robust and statistically significant negative relationship between the HHI for a household's output and participation in the supermarket channel at both the extensive and intensive margins.

Beyond statistical significance, the economic significance of our results is as follows. Given that the local average treatment effect of participation in the supermarket channel on the HHI for a household's output is equal to -0.173 and the average HHI for output in the data is equal to 0.445, participation in the supermarket channel is associated with a 39% (i.e., -0.173/0.445) decrease in how concentrated a household's crops are within that household's output for those households whose participation in the supermarket channel (our treatment variable) was induced by the number of friends and relatives they have who work as agents in the supermarket chain (our instrumental variable). Put differently,

participation in the supermarket channel makes participating households less specialized as the crops grown for the supermarket channel merely add themselves onto the other crops grown by the household.

The fact that this is only true for specialization as measured by how crop-concentrated a household's total output is but not for specialization as measured by how crop-concentrated a household's landholdings are poses a seeming paradox. We suggest a few ways this paradox can be resolved in our discussion of our result.

From a policy perspective our results suggest that, for policy makers interested in fostering economic specialization, encouraging participation in the supermarket channel is likely to backfire, though the relationship between participation in agricultural value chains might change with the level of economic of development. One possible reason why participation in the super market channel does not lead to economic specialization might be that the farmers in our data are risk-averse and that insurance markets fail, but since our data were not collected to test that hypothesis, it must remain speculative. Still, if one were to take that our results at face value and that hypothesis were true, then providing agricultural insurance would lead to greater specialization. Finally, contrasting our findings with the degree of specialization of farmers in OECD countries suggest an interesting area of theoretical investigation.

References

Andersson, C. I., Chege, C. G., Rao, E. J., Qaim, M. (2015), "Following up on smallholder farmers and supermarkets in Kenya," *American Journal of Agricultural Economics* 97(4): 1247–1266.

Bellemare, M.F. (2012), "As you sow, so shall you reap: The welfare impacts of contract farming," *World Development 40*(7): 1418-1434.

Bellemare, M.F. (2018) "Contract farming: Opportunity cost and trade-offs," *Agricultural Economics* 49(3): 279-288.

Bellemare, M.F., and J.R. Bloem (2018), "Does Contract Farming Improve Welfare? A Review," *World Development* 112: 259-271.

Bellemare, M.F., and Novak, L. (2017), "Contract farming and food security," *American Journal of Agricultural Economics* 99(2): 357–378.

Chege, C.GK, Andersson, C.I.M., and Qaim, M. (2015), "Impacts of supermarkets on farm household nutrition in Kenya," *World Development* 72: 394-407.

Chen K., Shepherd, A., and Silva, C. (2005), "Changes in food retailing in Asia: Implications of supermarket procurement practices for farmers and traditional marketing system," Rome: FAO, http://www.fao.org/3/a-a0006e.pdf last accessed on 30th October, 2018

Emran, M.S., and Shilpi, F. (2012), "The extent of the market and stages of agricultural specialization," *Canadian Journal of Economics* 45(3): 1125-1153.

Hernández, R., Reardon, T., and Berdegué, J. (2007), "Supermarkets, wholesalers, and tomato growers in Guatemala," *Agricultural Economics* 36(3): 281-290.

Kurosaki, T. (2003), "Specialization and diversification in agricultural transformation: The case of West Punjab, 1903-92," *American Journal of Agricultural Economics* 85(2): 372-386.

Michelson, H. C. (2013), "Small farmers, NGOs, and a Walmart world: Welfare effects of supermarkets operating in Nicaragua," *American Journal of Agricultural Economics* 95(3): 628-649.

Michelson H.C., Reardon T., and Perez, F. (2012), 'Small Farmers and Big Retail: Trade-offs of Supplying Supermarkets in Nicaragua," *World Development* 40(2): 342-354.

Minten B., Reardon T., and Sutradhar, R. (2010), "Food prices and modern retail: The case of Delhi," *World Development* 38(12): 1775-1787.

Montalbano, P., Petrelli, R., & Salvatici, L. (2018), "Participation in the market chain and food security: The case of the Ugandan maize farmers," *Food Policy*, 76: 81-98.

Narayanan S. (2012), "Notional contracts: The moral economy of contract farming arrangements in India," IGIDR Working Paper http://www.igidr.ac.in/pdf/publication/WP-2012-020.pdf last accessed 30th October, 2018.

Neven, D., and Reardon, T. (2004), "The rise of Kenyan supermarkets and the evolution of their horticulture product procurement systems," *Development Policy Review* 22(6): 669-699.

Omamo, S.W. (1998), “Farm to market transaction and specialization in small-scale agriculture: Explorations with a nonseparable household model,” *Journal of Development Studies* 35: 152-63.

Pritchard, B., Gracy, C. P., and Godwin, M. (2010), “The impacts of supermarket procurement on farming communities in India: Evidence from rural Karnataka,” *Development Policy Review* 28(4): 435-456.

Qin, Y., and Zhang, X. (2016), ”The road to specialization in agricultural production: Evidence from rural China,” *World Development* 77: 1-16.

Rao, E. J., & Qaim, M. (2011), “Supermarkets, farm household income, and poverty: insights from Kenya,” *World Development* 39(5): 784-796.

Rao N.C., Sutradhar, R., and Reardon, T. (2017), “Disruptive Innovations in Food Value Chains and Small Farmers,” *Indian Journal of Agricultural Economics* 72(1): 24-48.

Reardon, T., Barrett, C. B., Berdegué, J. A., and Swinnen, J. F. (2009), “Agrifood industry transformation and small farmers in developing countries,” *World Development* 37(11): 1717-1727.

Reardon, T., Timmer, C. P., Barrett, C. B., and Berdegué, J. (2003), “The rise of supermarkets in Africa, Asia, and Latin America,” *American Journal of Agricultural Economics* 85(5): 1140-1146.

Smith, A. (1994), *An inquiry into the nature and causes of the wealth of nations*, New York: Modern Library.

Ton, G., Vellema, W., Desiere, S., Weituschat, S., and D’Haese, M. (2018), “Contract farming for improving smallholder incomes: What can we learn from effectiveness studies?,” *World Development* 104: 46–64.

Table 1. Descriptive Statistics and Balance Tests

Variables	(1) Whole Sample	(2) Traditional	(3) Supermarket	(4) Diff. (2)-(3)	(5) 2009	(6) 2014
Herfindahl-Hirschman Index for Land (in (0,1] Interval)	0.339	0.347	0.330	*	0.332	0.343
	(0.005)	(0.007)	(0.007)		(0.007)	(0.007)
Herfindahl-Hirschman Index for Output (in (0,1] Interval)	0.445	0.471	0.416		0.356	0.508
	(0.007)	(0.010)	(0.009)		(0.007)	(0.009)
Supermarket Channel Participant (Dummy)	0.468				0.591	0.382
	(0.020)				(0.031)	(0.026)
Supermarket Channel Agents in Social Network (Number)	0.364	0.077	0.690	***	0.413	0.329
	(0.032)	(0.015)	(0.062)		(0.070)	(0.025)
Age of Household Head (Years)	47.011	48.356	45.484	***	45.720	47.925
	(0.467)	(0.648)	(0.664)		(0.692)	(0.626)
Education of Household Head (Years)	3.897	3.399	4.463	***	3.142	4.432
	(0.165)	(0.230)	(0.232)		(0.153)	(0.256)
Household Head Member of Cooperative or Self-Help Group (Dummy)	0.594	0.617	0.568		0.268	0.825
	(0.020)	(0.027)	(0.029)		(0.028)	(0.020)
Landholdings (Hectares)	3.487	3.409	3.575		3.207	3.685
	(0.109)	(0.141)	(0.170)		(0.122)	(0.165)
Household Size (Individuals)	4.817	4.834	4.798		4.772	4.850
	(0.072)	(0.103)	(0.099)		(0.104)	(0.098)
Value of Farm Equipment (1,000 Rupees)	62.953	52.797	74.488	*	66.243	60.624
	(5.442)	(5.921)	(9.448)		(5.974)	(8.282)
Off Farm Work (Dummy)	0.564	0.623	0.498	***	0.531	0.588
	(0.020)	(0.027)	(0.030)		(0.031)	(0.026)
Observations	613	326	287		254	359

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2. Estimation Results for Land and Participation in the Supermarket Value Chain at the Extensive Margin

Variables	(1) OLS	(2) 2SLS	(3) FE	(4) 2SLS-FE
Dependent Variable: Herfindahl-Hirschman Index for Land				
Supermarket Channel Participant	-0.016*	-0.021	-0.019	-0.090
	(0.009)	(0.020)	(0.015)	(0.075)
Age	-0.001	-0.001	-0.002	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Education	0.001	0.001	-0.001	-0.005
	(0.001)	(0.001)	(0.003)	(0.004)
Member of Coop or SHG	0.014	0.014	0.022	0.003
	(0.010)	(0.010)	(0.014)	(0.024)
Landholdings	-0.011***	-0.011***	-0.009***	-0.012***
	(0.002)	(0.002)	(0.004)	(0.005)
Household Size	-0.008***	-0.008***	-0.008*	-0.009*
	(0.002)	(0.002)	(0.005)	(0.005)
Value of Farm Equipment	0.000*	0.000*	0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Off Farm Work	0.022**	0.021**	0.050**	0.048**
	(0.010)	(0.010)	(0.021)	(0.023)
Constant	0.417***	0.420***	0.458***	0.549***
	(0.030)	(0.033)	(0.059)	(0.112)
Observations	613	613	613	613
R-squared	0.084	0.083	0.100	-
F-statistic (Instrumental Variable)	-	22.80	-	13.69

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3. Estimation Results for Land and Participation in the Supermarket Value Chain at the Intensive Margin

Variables	(1) OLS	(2) 2SLS	(3) FE	(4) 2SLS-FE
Dependent Variable: Herfindahl-Hirschman Index for Land				
Share of Marketed Output Sold to Supermarket	-0.000	-0.000	-0.000	-0.003
	(0.000)	(0.000)	(0.000)	(0.002)
Age	-0.000	-0.000	-0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Education	0.001	0.001	-0.001	-0.003
	(0.001)	(0.001)	(0.003)	(0.004)
Member of Coop or SHG	0.015	0.017*	0.025*	0.010
	(0.010)	(0.010)	(0.014)	(0.021)
Landholdings	-0.010***	-0.010***	-0.009***	-0.010***
	(0.002)	(0.002)	(0.003)	(0.004)
Household Size	-0.008***	-0.008***	-0.008*	-0.010*
	(0.002)	(0.002)	(0.005)	(0.005)
Value of Farm Equipment	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Off Farm Work	0.023**	0.022**	0.052**	0.062**
	(0.010)	(0.010)	(0.022)	(0.026)
Constant	0.409***	0.414***	0.443***	0.496***
	(0.029)	(0.031)	(0.054)	(0.083)
Observations	613	613	613	613
R-squared	0.080	0.076	0.098	-
F-statistic (Instrumental Variable)	-	12.26	-	6.61

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4. Estimation Results for Output and Participation in the Supermarket Value Chain at the Extensive Margin

Variables	(1) OLS	(2) 2SLS	(3) FE	(4) 2SLS-FE
Dependent Variable: Herfindahl-Hirschman Index for Output				
Supermarket Channel Participant	-0.054***	-0.052**	-0.126***	-0.173***
	(0.012)	(0.022)	(0.021)	(0.044)
Age	-0.001	-0.001	0.001	0.000
	(0.001)	(0.001)	(0.002)	(0.002)
Education	0.003	0.003	0.003	0.001
	(0.002)	(0.002)	(0.004)	(0.004)
Member of Coop or SHG	0.087***	0.087***	0.112***	0.099***
	(0.013)	(0.013)	(0.020)	(0.024)
Landholdings	-0.008**	-0.008**	-0.007	-0.008*
	(0.003)	(0.003)	(0.005)	(0.005)
Household Size	-0.008**	-0.008**	-0.014**	-0.014**
	(0.003)	(0.003)	(0.006)	(0.006)
Value of Farm Equipment	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Off Farm Work	-0.000	-0.000	0.025	0.024
	(0.013)	(0.013)	(0.027)	(0.027)
Constant	0.510***	0.509***	0.468***	0.527***
	(0.036)	(0.039)	(0.072)	(0.091)
Observations	613	613	613	613
R-squared	0.130	0.130	0.383	-
F-statistic (Instrumental Variable)	-	22.80	-	13.69

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5. Estimation Results for Output and Participation in the Supermarket Value Chain at the Intensive Margin

Variables	(1) OLS	(2) 2SLS	(3) FE	(4) 2SLS-FE
Dependent Variable: Herfindahl-Hirschman Index for Output				
Share of Marketed Output Sold to Supermarket	-0.001***	-0.001**	-0.002***	-0.005***
	(0.000)	(0.000)	(0.001)	(0.002)
Age	-0.001	-0.001	0.002	0.002
	(0.001)	(0.001)	(0.002)	(0.002)
Education	0.003	0.003	0.006	0.003
	(0.002)	(0.002)	(0.004)	(0.004)
Member of Coop or SHG	0.092***	0.095***	0.130***	0.112***
	(0.013)	(0.013)	(0.020)	(0.025)
Landholdings	-0.007**	-0.007**	-0.003	-0.005
	(0.003)	(0.003)	(0.005)	(0.005)
Household Size	-0.009**	-0.009**	-0.014**	-0.015**
	(0.003)	(0.003)	(0.006)	(0.006)
Value of Farm Equipment	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Off Farm Work	0.003	0.001	0.039	0.051
	(0.013)	(0.013)	(0.030)	(0.032)
Constant	0.486***	0.494***	0.363***	0.427***
	(0.035)	(0.037)	(0.076)	(0.091)
Observations	613	613	613	613
R-squared	0.113	0.107	0.334	-
F-statistic (Instrumental Variable)	-	12.26	-	6.61

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1