Impact of Proximity to Agricultural Market on Farmer's Production-side Welfares

Hirpara Parth[1], Nagarajan Hari[2], Das Abhiman[3], Pandey Vivek[4]

Abstract

Indian agricultural markets exhibit a monopsonistic structure, largely a consequence of prevailing regulations. These regulations, especially market area specifications, foster regional disparities in coverage, disincentivize product differentiation, and render farmers price takers. Consequently, farmers are constrained from optimizing production choices and engaging in price or product differentiation due to restrictive policies. Using a spatial regression discontinuity design to endogenize market distance, we assess its effect on farmers' production-related welfare, focusing on input costs. Our results show that reducing unit distance to an in-state mandi decreases per-acre total household cost by 32.6 percentage points and machinery cost share by 2.1 percentage points. Farmers also become 49.8 percentage points more likely to grow more crops. Despite these cost reductions, profit per acre is insignificantly impacted, suggesting lower produce prices in locally saturated markets due to restrictive policies preventing access to distant, more competitive ones. Our findings highlight the critical need to deregulate Indian agricultural markets, enabling farmers to realize greater profits beyond mere cost reductions.

Keywords: Agriculture markets, Distance, Production costs, Market access, Farmers Welfare

JEL Classification: D24, D61, R11, Q12, Q13

[1] PhD fellow, Centre for Management in Agriculture, Indian Institute of Management, Ahmedabad.
[2] Professor, Centre for Management in Agriculture, Indian Institute of Management, Ahmedabad.
[3] Professor, Economics area, Indian Institute of Management, Ahmedabad.
[4] Professor, Economics area, Institute of Rural Management, Anand.

Introduction

The proximity of farms to agricultural markets plays a crucial role in shaping farmers' economic outcomes, impacting their welfare in terms of production cost, income, profit, and dietary diversity. In an agricultural market that is characterized by monopsonistic market structures, the power dynamics inherent in price setting reveal the complex relationship between market distance and farmers' profit-maximizing strategies (Chatterjee, 2023). Farmers are often restricted in their ability to adjust prices based on the distance to buyers and selection of markets to sell their produce, leading to inefficiencies and price distortions that directly affect their welfare. This is because the government sets an exogenous price of commodities (minimum support price - MSP), often enforced across regulated markets (Agricultural Produce Market Committees - APMCs) and restricts the operating area of such markets, implying that a farmer faces a monopsonistic buyer that restricts the per unit revenue, which remains constant in the farmer's market jurisdiction. Consequently, a farmer's profitability will depend on the per unit cost incurred in accessing the APMCs.

Consider farmer A, who is situated at a distance of d_{a1} and d_{a2} from markets m_1 and m_2, respectively. Similarly, farmer B is situated at a distance of d_{b1} and d_{b2}, respectively. The farmer's input choice is a function of two competing mechanisms. First, the government can constrain both the farmers from selling in m_1 by invoking constitutionally mandated market area restrictions (APMC Act). This happens if m_1 is situated in a jurisdiction that is different from that of A and B. Second, conditional on this constraint, both A and B face different cost functions in accessing the market m_2. This difference is attributed to the relative magnitude of d_{a2} and d_{b2}. If $d_{a2}>d_{b2}$, farmer A finds it costly to sell in m_2. Since the per unit price realization for the same crop grown by A and B is constant, it follows that farmer A is less likely to be profitable than farmer B due to cost differentials in accessing m_2.

Also, it can be said that the Indian agricultural markets are characterized by a monopsonistic structure (Chatterjee, 2023). This is due to the consequences of regularizing agricultural markets during the 1960s and 1970s through the Agricultural Produce Market Regulation (APMR) Acts of 1966. The APMC Act aims to regulate market activities and protect farmers' interests by establishing Agricultural Produce Market Committees (APMCs) for regulated markets (Mandis) in a specified area. Consequently, the restrictive nature of the APMC Act regarding market area specification has created regional variations in market coverage, reduced the incentives for farmers to engage in product differentiation, and rendered them price takers.

The restrictive nature of the APMC Act can be observed from the non-uniform placement of agricultural markets across states[5], creating regional variations in market coverage, reducing the incentives for farmers to engage in product differentiation, and rendering them price takers. This distributional disparity means that many farmers cannot easily access multiple markets to sell their produce, which is a fundamental feature in the existing system designed to manage an ever-increasing marketable surplus affecting their per unit price realization. Therefore, this paper assesses the impact of jurisdiction-based market area restrictions on farmer profitability. In particular, we estimate the cost functions of representative farmers A and B to confirm that farmer A indeed incurs higher per unit various input costs than farmer B, making farmer A less profitable. We will empirically endogenize the distance to agriculture markets in a regression discontinuity design using primary survey data collected from villages across the state border of Uttar Pradesh (UP) and Madhya Pradesh (MP).

[5] The uneven distribution of agriculture markets in India created considerable variations in the distance of agricultural farms from the designated markets. For example, many states, such as Meghalaya (with a total market area of 11214 sq km), Rajasthan (902 sq km), or Madhya Pradesh (833 sq km), experience severe shortages of regulated markets compared to states like Punjab (76-118 sq km) or Haryana (76 sq km). However, the national average area served by a regulated market is 457 sq km, and the recommended market area is 80 sq km (meaning the market should be within a 5 km radius of the farmer).

Various studies across different countries have shown that closer proximity to the markets in terms of distance or accessibility leads to reduced transportation costs, transaction costs, better market information, enhanced investment and adoption of conservation agriculture technologies, access to lower-cost agriculture inputs, increased probability of crop diversification, favorable produce prices, and enhanced farm returns. For example, in Zambia, Ethiopia, Nepal, India, Ghana, and Uganda by Abdulai (2016), Abebe *et al.* (2016), Ghimire and Huang (2016), Dey and Singh (2023a), Dey and Singh (2023b), Adam and Abdulai (2024), and Tesfaye and Tirivayi (2020) respectively. Studies in other regions support these findings. For example, closer proximity to markets enhances the adoption of climate-smart agricultural practices, access to better agriculture inputs in China (Liang *et al.*, 2021; Zheng & Ma, 2023), improves tomato seed varieties in Ghana (Shafiwu *et al.*, 2022), improving yields, income, and reduces poverty (Ma *et al.*, 2024). This influences farmers' risk perceptions and adoption of risk management strategies, as shown by Akhtar *et al.* (2018) in Pakistan, with those farther away from markets being more risk-averse. Further, the positive associative nature of market engagement, on-farm crop diversity, and dietary diversity was observed in Uganda, Vietnam, Yayu (southwestern Ethiopia), Northern Ghana, Nigeria, Tanzania, Ethiopia, Bangladesh, and India by Morrissey *et al.* (2024), Esaryk *et al.* (2021), Usman and Callo-Concha (2021), Addai *et al.* (2023), Awotide *et al.* (2016), Mmbando *et al.* (2015), Haile *et al.* (2022), and Meskel *et al.* (2020), Hoq *et al.* (2021), and John *et al.* (2021) respectively.

The results will provide a robust interpretation of endogenizing distance to the agriculture market on farmers' welfare, in particular, various costs involved in their production technology. The insight by Chatterjee (2023) shows that distance to agricultural markets affects market competition and, hence, farmers' price-realization mechanism. Given the fixed spatial distribution of agricultural markets in India, our study's results will help further segregate farmers' profit-maximization strategies that account for various production costs. Also, it will shed light on deregulating agricultural markets. We describe data collection in the next section, followed by empirical strategy, results, discussion, and conclusion.

Data Description

Survey design

We use household- and village-level "Rural Market Survey in Uttar Pradesh and Madhya Pradesh (2020)" primary data in our analysis to estimate the impact of distance to mandis (APMCs) from villages on farmers' production-side welfare. Farmers who reside in close proximity to the state border could manage to sell in the mandi of the other state. State APMC laws prevent farmers from selling them to other state mandis. Therefore, there is an issue of non-compliance near the state boundaries. During data collection, we improved the non-compliance by refining our study population. Specifically, we focused on villages situated 2 to 5 km from the shared state boundary of Uttar Pradesh (UP) and Madhya Pradesh (MP), while excluding villages within a 0 to 2 km range on either side of the border. Hence, the overall population of our study included villages within 2 to 5 km of the state border in the selected districts of both states.

In the survey, we randomly selected four districts in UP from the set of eleven districts that share their boundaries with thirteen districts of MP. The selection of these districts identifies the two districts in MP sharing their boundaries with the randomly selected districts from UP. To select the villages, we start with the village shapefile of India's UP and MP states. We have selected the continuum of blocks[6] from the Bhind and Rewa districts of MP, as Agra, Jalaun, Etawah, and Prayagraj districts of UP are their neighboring districts. We obtained the contiguous villages from the selected blocks that fall under the 5km area from the state

[6] Blocks are a district's planning and development units in addition to tehsils.

border of selected districts[7]. In these selected blocks, 1386 rural villages were obtained within 5 km of each state's border (Figure 1). We have used the administrative data to identify the mandi (APMC) locations. In MP, the mandi locations are obtained from the official APMC website. For UP, they are obtained from Google Maps using the address mentioned in the directory of wholesale agricultural produce assembly markets in India[8]. These locations were added to the QGIS, and projection was done using the "data analyst tool" to create the "nearest distance table" (linear distance)[9]. We have obtained the data for the nearest distance of all the villages to the mandis and the distance of each village from the state border.

Sample selection

We have sampled 45 villages in Madhya Pradesh and 52 villages from Uttar Pradesh in the selected districts, taking account of the appropriate probability weights that are equivalent to the inverse of the sample fraction.[10] From each village, we sampled 12 random households (1194 surveyed). Both the selected villages and households were assigned random unique identifiers. We defined the treatment and control villages by formulating a 'distance difference to the nearest mandi' matrix. We have taken the distance difference by considering the distance of a village to the nearest mandi of the respective state and the nearest mandi of another state. If the distance difference is negative, the nearest mandi is in the same state as the village. A pictorial representation of the same is presented in Figure 2. Villages with the nearest mandi in their respective state are considered in the treatment group, and if the nearest mandi is in another state, they are assigned to the control group. Therefore, our sample has

---

[7] We have used QGIS software to identify the contiguous villages in the selected blocks. QGIS is geographic information system software and open-sourced for all users.

[8] The government regularly updates the APMC mandi website for M.P., but not in the case of U.P. Therefore, we have traced the mandi address using G-Map for U.P. The address was obtained through the directory of wholesale agricultural produce assembly markets in India.

[9] We used the linear distance (displacement from mandi to village) to sample the villages. Using the official road data would be riskier as there are a lot of villages with no road data (villages might be using the 'Katcha' road) to reach the nearest mandi. Therefore, using the road distance data would be an imprecise way.

[10] We have used the finite population correction (FPC) for the sampling, as the sample represents a large proportion of the population. FPC is used to correct standard errors and improve the efficiency of estimators.

375 treated and 200 control households in MP and 228 treated and 391 control households in UP.

Variables collected in the survey

The survey has two modules: household-level and village-level. In the household-level module, we collected data on household composition, household income, consumption expenditure, financial literacy, land characteristics, crop insurance, crop output, cost of cultivation, awareness of crop prices, market channels, household assets, savings, and education. The member-level information was collected in some of these sections, such as household composition, household income, consumption expenditure, land characteristics, and education. We have constructed our household-level variables by aggregating the member-level information. The village-level module collected information on the location of amenities from the village, agriculture and livestock information, land tenure system, extension services, crop price, employment and business structure, and local governance. From this, we have used information on the agriculture costs to calculate the missing values at some places in the household agriculture input costs, such as machinery rental or labor costs.

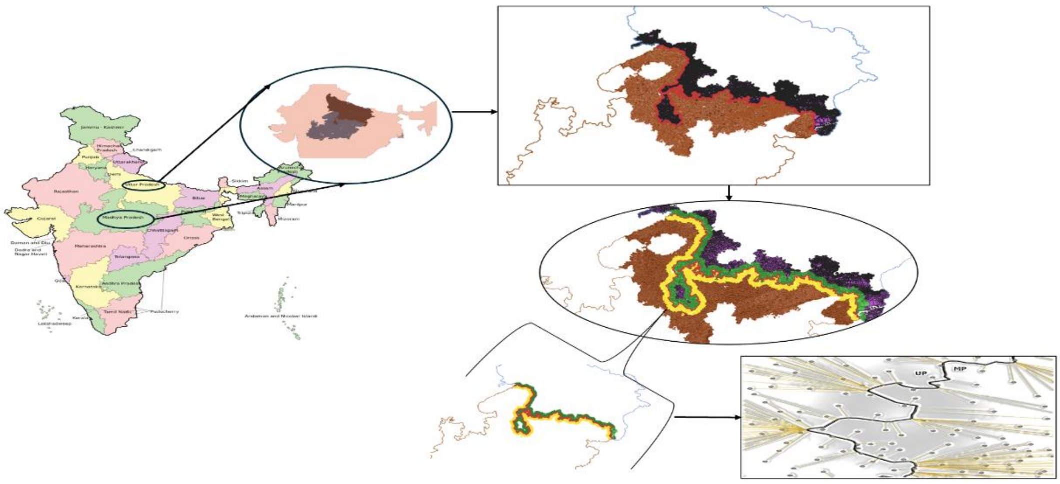

Figure 1: Village across borders of UP and MP within 5km range from border[11]

[11] **Figure 1:** The *blue, purple*, and *green* colors are for Uttar Pradesh, while the *brown* and *yellow* colors are for Madhya Pradesh

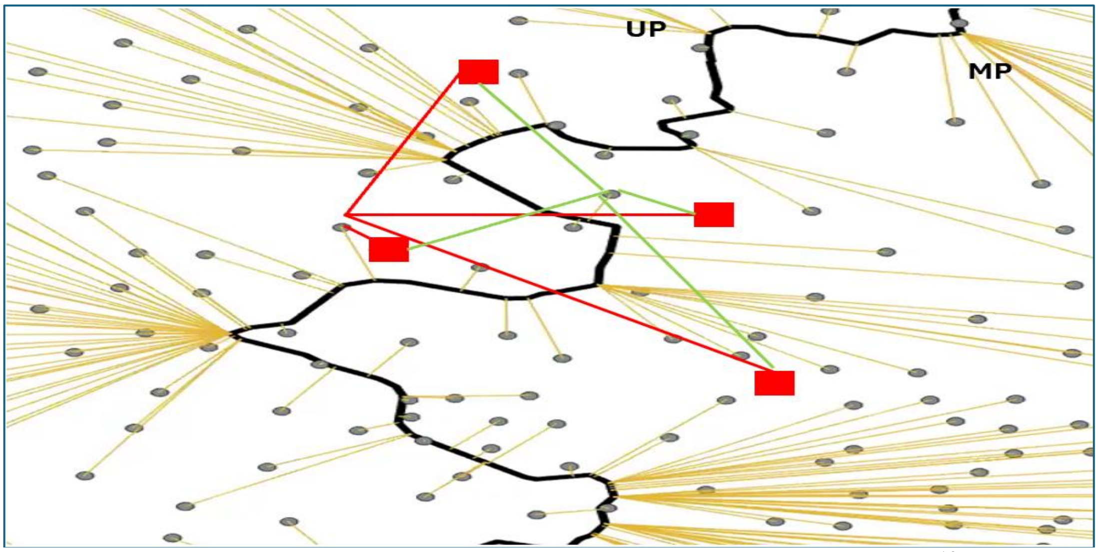

Figure 2: 'Distance difference to the nearest mandi' matrix[12]

Our primary outcome variables consist of household per acre total cost, share of machine cost in total cost, number of crops grown, household per acre profit, and dietary diversity index (Herfindahl index) with other input costs (seed, labor, chemicals, machine, and irrigation). We have formulated per-acre outcome variables using the land cultivated information. In the empirical model, we have used the age and education of male and female household members, dependency ratio, member illness, soil health card, and crop insurance information as our covariates. Also, we have used the 'distance difference to the nearest mandi' as our continuous running variable (CRV) to set up the threshold for the treatment and control villages in our empirical analysis. We have created an indicator variable, 'segment APMC,' representing the coverage of villages by one APMC. As multiple villages are taking the services of a single APMC within a district, this indicator variable is useful for controlling the spatial variation for mandi-level heterogeneity in our empirical estimation.

[12] **Figure 2:** This is just a representation. Each *gray* dot represents a village, with a *yellow* line indicating the distance from that village to the nearest point on the state border. The *red* rectangle represents APMCs, with *red* and *green* lines representing the distance between a village and all APMCs.

Summary statistics

Table 1 presents the summary statistics of the data used by comparing mean of the household characteristics and farmers' welfare (outcomes variables) of the treatment and control groups. Column 1 gives the number of total observations, while columns 2 and 4 give the number of observations in control and treatment groups, respectively. Columns 3 and 5 present a simple mean considering all variables for control and treatment groups, respectively. We highlight simple mean comparisons of the variables across groups. The sample's average age of males and females is between around 42 and 45 years in both the treatment and control groups. The average year of education is lower for females (5 to 6 years) than for males (8 to 9 years). Around 6 times a household member fell ill in the previous year (in 2019) across treatment and control groups, with a household dependency ratio of around 39. Only 1 to 2 percent of households in the sample avail crop insurance, while 6 to 7 percent of households have a soil health card. The average land holding size of households in the sample is 4 to 6 acres.

Note that almost all covariates are statistically insignificant, and outcome variables are statistically significant across treatment and control groups. However, we observed that the mean male members' age, frequency of illness of household members, and participation in household crop insurance schemes are statistically significant. So, we employed the optimal bandwidth with the regression discontinuity (RD) model with district- and state-fixed effects for a stringent check. The specification of the RD model and results are discussed in the following sections.

Empirical Strategy

In our empirical estimation model, we used the spatial regression discontinuity (RD) design to estimate the effect of endogenizing distance to market on production-side farmers' welfare outcomes. We will also describe and test the validity of RD-identifying assumptions. Using the discontinuity generated due to the differences in distance between the nearest mandi to the village within the state and across the border (another state), we employ a spatial regression discontinuity (RD) design to estimate the effect of distance to mandi on farmers' welfare outcomes. We used the 'distance difference' to the mandi as a continuous running variable (CRV) to set up the threshold for the treatment and control villages in the RD estimation model. Therefore, a unit increase in distance difference (in absolute terms, in kilometres) translates to an increase in the distance of the nearest mandi from the village since we have considered negative distance difference for the treatment group and vice versa.

We present results using a nonparametric approach to estimate the treatment effects using a one-dimensional forcing variable: the distance difference to the nearest mandi from the village and the transformed forcing variable (in polynomial form). Our primary outcome variables consist of household per acre total cost, share of machine cost in total cost, number of crops grown, household per acre profit, and dietary diversity index (Herfindahl index) with other input costs (seed, labor, chemicals, machine, and irrigation). We will estimate the various specifications of the equation below:

$$(1)\ y_{ij} = \alpha + \beta T_j + \gamma\left(dist\_near\_APMC_{j,own} - dist\_near\_APMC_{j,other}\right) + \delta T_j\left(dist\_near\_APMC_{j,own} - dist\_near\_APMC_{j,other}\right) + controls_{ij} + \Omega_{apmc} + \varepsilon_{ij}$$

Where y_{ij} is an outcome variable (household per acre total cost, share of machine cost in total cost, number of crop grown, dietary diversity index, household per care profit, and various input costs) of household i in village j; $dist\_near\_APMC_{j,own}$ is the distance of the village j to nearest APMC in own state; $dist\_near\_APMC_{j,other}$ is the distance of the village j to nearest APMC in other state; T_j is dummy variable that equals 1 if the nearest APMC to the village is within the same state, otherwise zero; $controls_{ij}$ are different covariates of household i in village j used in analysis, and Ω_{apmc} is mandi-level segment fixed effect that ensures that we are comparing households that are within the same mandi coverage. In addition, to satisfy the boundary positivity assumption described by Imbens and Zajonc (2011), we cluster standard errors at the village level.

First, we have used a nonparametric approach to estimate the treatment effects, which relaxes the functional form assumption in parametric regression using the *Epanechnikov* kernel function and optimal bandwidth selection as per Calonico *et al.* (2014) and Calonico *et al.* (2020). Second, we have used a more parametric approach using the polynomial in the forcing variable, which uses all observations on both sides of the cut-off. The optimal order of polynomials will be selected using Akaike's criterion. RD coefficient δ will give the causal effect of variation in the distance to the nearest mandi on farmers' welfare outcomes. We interpret the RD coefficient's (δ) sign and magnitude to show how a decrease in the unit distance to a mandi influences the farmers' welfare outcomes (leading to either an increase or a decrease). Furthermore, δ can be understood through the lens of local versus distant competition. Our measure is the comparative distance difference, calculated as the distance from the farm to its nearest mandi in its own state, relative to the distance from the farm to the nearest mandi in a neighboring state. A negative distance difference signifies that the mandi is within the farmer's own state, implying the farmer experiences greater local competition. In contrast, a positive distance difference suggests the farmer faces more distant competition.

Validity of RD assumptions

RD identification requires a fundamental assumption: the potential outcome functions $E[y_{ij}(1)|X]$ and $E[y_{ij}(0)|X]$ must be continuous at cut-off points, i.e., immediately surrounding the forcing variable at zero in the treatment boundary, where one and zero denote treatment assignment and non-assignment, respectively. In simple terms, the household covariates must transition smoothly across the cut-off. This assumption allows units in the control group to serve as a valid counterfactual for the units in the treatment group. To assess the validity of the RD assumption, we perform the McCrary (2008) test for breaks in the density of the forcing variable at the cut-off boundary.

Results

We begin by showing the RD assumption validity through the McCrary test and covariate balancing (Table 1). Then we describe the graphical analysis of the outcome variables, and the RD results estimated using APMC segment fixed effects and village clusters.

The McCrary (2008) test finds no evidence of endogenous assignment of sorting of villages near the cut-off. Figure A1 and Table A1 show McCrary's (2008) test results. We find a statistically insignificant jump in the distance difference to the mandis around the cut-off, suggesting that there is no manipulation in the allocation of the treatment units. In the context of this study, we perform this test even though it is not a concern because we have considered forcing variables concerning the mandis, and mandis are randomly assigned by the state regulations. Also, our sampling strategy eliminates the endogenous selection concerns by selecting villages between 2 to 5 km bands in each state. Despite this, in our main result estimates, we will consider the possibility of endogenous selection of farmers to a mandi, i.e., farmers selling across the nearest mandi.

Table 1, column 8, compares the mean of the covariates and outcome variables using polynomials in forcing variables in the RD design with district- and state-fixed effects. Column 7 presents the specific bandwidth for the covariates we have used in our estimation. For outcome variables, we have included polynomials in the forcing variable. Column 9 reports the cluster standard error of the mean difference between control and treatment villages. The RD estimates present in column 8 reveal that all the covariates are statistically insignificant for covariates, and from Figure 3 (RD plots), we can see that the covariates are smooth around the treatment. This suggests that all the covariates are balanced and are transmitted smoothly across the running variable. Similarly, statistical significance can be seen for the outcome variables for various outcome variables from simple mean comparison as well as from RD estimates.

Table 2 and A2 present the results from the RD design that estimates the causal impact of distance on farmers' welfare outcomes using equation (1). All specifications use APMC segment fixed effects, and standard errors are clustered by village to account for the potential spatial correlation of unobservables with APMC segments. In both tables, columns 1 and 5 present the number of observations, column 3 presents the RD coefficient estimated using the optimal bandwidth from column 2, and columns 4 and 7 present standard errors clustered at the village level. Column 6 presents the result of the RD design, which uses polynomials of the forcing variable and considers all observations without a specific bandwidth. All the specifications include covariates. The covariates used are mean age and education of male and female household members, frequency of falling ill, dependency ratio, soil health card status of a household, and household crop insurance participation. All the cost outcome variables are per acre and in the logarithmic form.

Table 2 shows that the total cost per acre and share of machinery in the total cost of the household are statistically lower in the treatment villages. At the same time, the number of crops grown is statistically higher in the treatment villages. There is insignificant evidence for households' per-acre profit and dietary diversity in the treatment villages. So, we can infer that if we reduce the distance to mandi by one unit (1 km), the per-acre household cost is statistically reduced by 32.6 percentage points, and the share of machinery cost in total household cost is reduced by 2.1 percentage points.

Table 1: Summary statistics of covariates and outcome variables

Variable	Mean comparison						RD estimates		
	Total count	Control		Treatment		Significance	Bandwidth	RD Coefficient	Standard error
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Covariates:									
Average age of female members	1194	591	42.26	603	42.72		7.792	-3.426	(2.245)
Average age of male members	1194	591	42.49	603	44.18	**	6.927	2.233	(3.394)
Average education of female members	1194	591	5.74	603	5.43		6.681	0.895	(0.981)
Average education of male members	1194	591	8.15	603	8.16		7.942	-0.203	(1.045)
Number of HH members fell ill (in last 12 months)	1194	591	6.12	603	5.76	**	6.925	-0.356	(0.579)
HH owns soil health card (=1)	1194	591	0.06	603	0.07		7.189	-0.003	(0.037)
HH avail any crop insurance scheme (=1)	1194	591	0.02	603	0.01	*	4.842	-0.019	(0.013)
Dependency ratio	1194	591	39.60	603	38.15		10.648	-0.921	(4.318)
Size of land holding	1172	584	4.64	588	5.86		18.022	2.839	(1.729)
Outcome variables:									
Per acre HH total cost	1172	584	11.07	588	10.87	***	-	-0.451***	(0.168)
Share of machine cost	1172	584	0.78	588	0.79	**	-	-0.011	(0.010)
Total number of crops grown (all seasons)	1194	591	2.86	603	2.97	*	-	0.613**	(0.267)
Dietary diversity index	1194	591	0.20	603	0.21	**	-	-0.002	(0.009)
Per acre HH profit	1172	584	12.92	588	12.93		-	-0.021	(0.023)
Per acre HH Rabi season cost	1172	584	9.28	588	9.22		-	-0.622**	(0.291)
Per acre HH Kharif season cost	1172	584	8.94	588	7.74	***	-	-0.424	(0.464)
Per acre HH total nursery cost	1172	584	4.03	588	2.76	***	-	-1.828*	(0.951)

Per acre seed cost	1172	584	8.04	588	7.76	***	-	-0.407***	(0.155)
Per acre chemical cost	1172	584	8.25	588	8.11	***	-	-0.409**	(0.183)
Per acre machine cost	1172	584	8.68	588	8.60		-	-0.478***	(0.177)
Per acre irrigation cost	1172	584	6.92	588	7.06		-	-1.168*	(0.641)
Per acre labor cost	1172	584	6.87	588	7.83	***	-	1.043	(1.200)
Per acre other costs	1172	584	6.45	588	6.35		-	-0.381	(0.293)

Note: Columns 1, 2, and 4 give the number of observations in the respective groups. Columns 3 and 5 give the mean of the respective variables. Column 6 gives the significance (*p<0.1 **p<0.05 ***p<0.01). Column 8 gives the RD coefficient as mentioned in the model (1) with cluster standard error at the village level in column 9 and optimal bandwidth in column 7. RD model used in columns 7, 8, and 9 does not use segment-fixed effects or covariates with district and state-fixed effects. It also contains polynomials in running variables with all observations. All costs considered in RD design are logarithmic values.

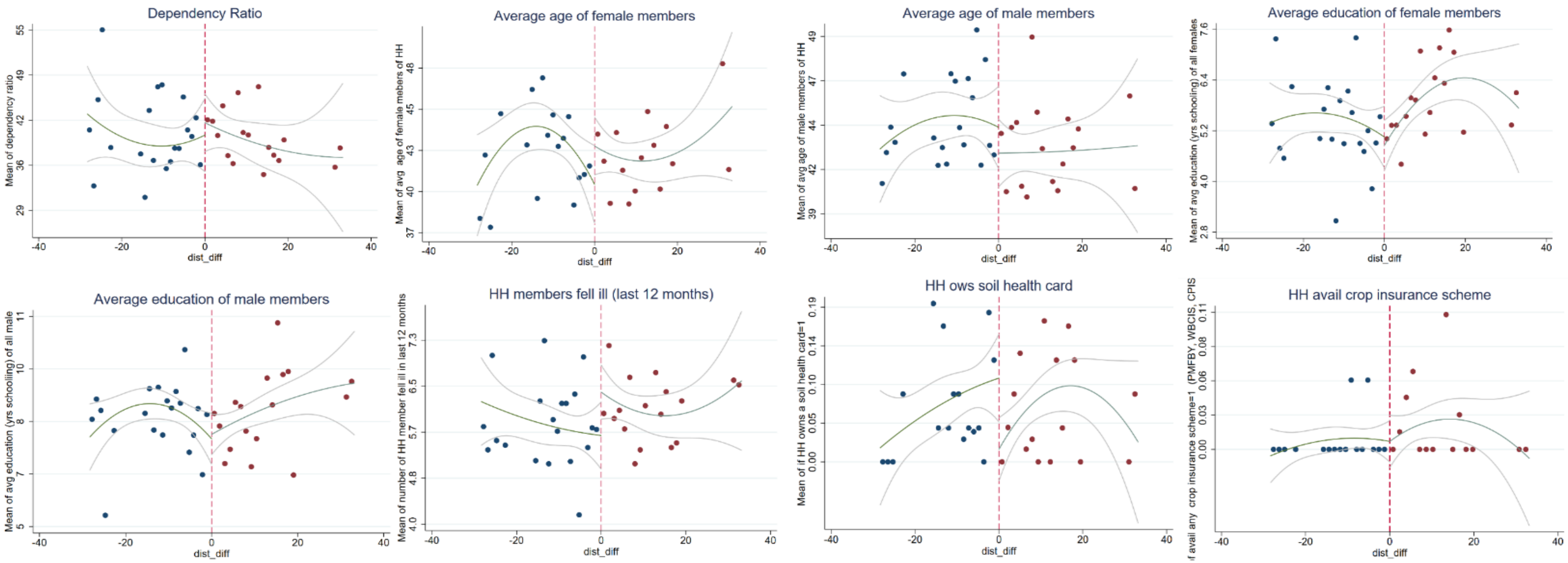

Figure 3: Binned average of various covariates (RD Plots)

While the household is statistically 49.8 percent more likely to grow more crops. Similarly, from Table A2, it is evident that per-acre Kharif season cost, seed cost, and machinery cost are statistically lower for the treatment group and insignificant for other input costs (nursery cost, labor cost, chemical input cost, and irrigation cost) for treatment villages. It means that if the distance to the mandis is reduced, then the costs of inputs, which are directly linked to the output markets, are also reduced. Reducing one unit distance to the mandis will lead to an 85.1 percent point, 71.5 percent point, and 51.2 percent point reduction in the Kharif season cost, per acre seed costs, and per acre machinery cost, respectively.

We have considered the logarithm of both the household per acre total cost and profit. It is evident from the RD plot (Figure 4) that household per-acre total cost and share of machinery cost in total cost are statistically lower in the treatment group, while the number of crops grown is statistically higher in the treatment group. There is an insignificant effect of distance on dietary diversity and household per-acre profit. This means that if the mandi is near the village and within the state, i.e., reducing the distance to mandi will significantly reduce household per-acre total cost and share of machinery cost in total cost while significantly increasing the number of crops grown by the household. Similarly, it is evident from Figure A2 that there is significantly lower household per acre rabi season cost, kharif season cost, seed cost, chemical input cost, machinery cost, irrigation cost, and other costs for the treatment group. We have used logarithmic values for all costs, and all costs are per-acre values. This shows that reducing the distance to the mandi is essential for reducing various costs associated with cultivation.

The results suggest that as the distance to mandis decreases, farmers are effectively facing greater local competition, which can be seen from a significant reduction in their total production cost and various input costs, and improved ability to grow more crops across seasons. In a local competition, farmers face less friction in market entry with

Table 2: Effect of distance on farmers' welfare outcomes

Variable	Optimal bandwidth				Polynomial in distance		
	Observations	Bandwidth	Coefficient	Standard error	Observations	Coefficient	Standard error
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Per acre HH total cost	480	7.167	-0.326*	(0.154)	1172	-0.371*	(0.181)
Share of machine cost	616	8.949	-0.021**	(0.009)	1172	-0.010	(0.010)
Total number of crops grown (all seasons)	492	7.211	0.498**	(0.202)	1172	0.481*	(0.264)
Dietary diversity index	480	7.186	-0.016	(0.010)	1172	-0.005	(0.010)
Per acre HH profit	440	6.826	-0.029	(0.018)	1172	-0.045	(0.030)

Note: Results use equation (1). Optimal bandwidths were chosen, as in Calonico, Cattaneo, and Titiunik (2014). Covariates were included in all the specifications. Covariates included were mean age and education of male and female household members, frequency of falling ill, dependency ratio, soil health card status of household, and household crop insurance participation. All specifications use APMC segment fixed effects and standard errors clustered at the village level. Polynomial order determined using Akaike's criterion. Significance level indicated are *p<0.1 **p<0.05 ***p<0.01.

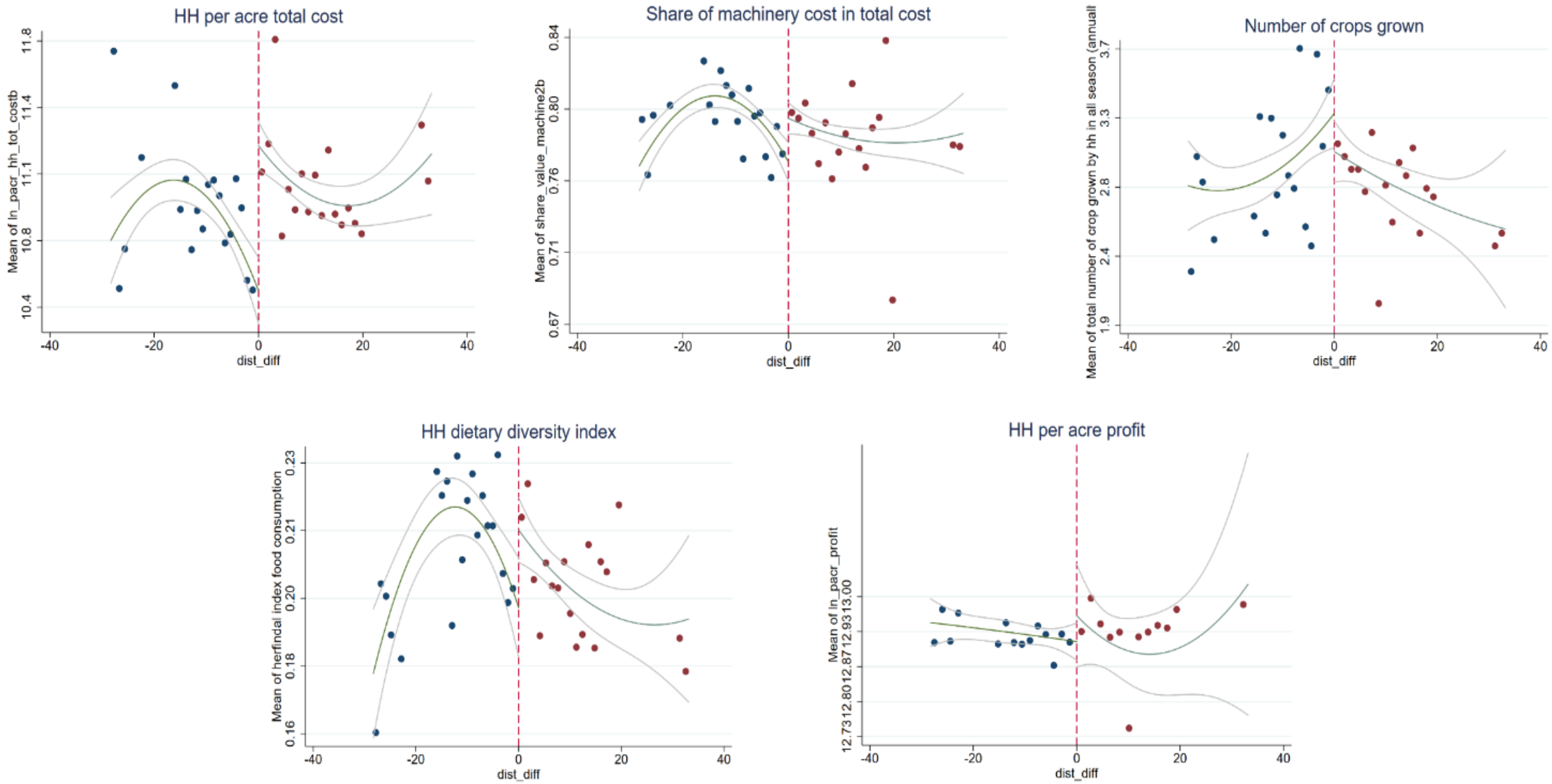

Figure 4: Binned average of various outcome variables (RD Plots)

lower transportation costs and improved efficiency in agricultural commodity transportation. Farmers can make more frequent market visits to local markets for the purchase of inputs, which provide them with quality inputs in a timely manner, and specifically for perishable commodities, they can visit with smaller produce quantities and gain favorable prices. Shorter distance to the output market also favors less spoilage and damage of agricultural produce, lowering the cost of losses incurred by the farmers.

Furthermore, increased local competition improves farmers' access to better price information, which enables them to make more informed decisions regarding both the sale of their produce and the purchase of inputs. This heightened competition among produce buyers and input sellers in local markets also tends to lower input prices and increase produce prices, ultimately leading to lower total costs for farmers. This effect is evident from the significant reduction we observed in total cost and specific input costs, particularly for seed and machinery.

The insignificant impact observed for chemical costs and total profit can be attributed to existing government policies in India. Specifically, the government largely determines prices and provides subsidies for major fertilizers, which could explain the insignificant effect of distance to mandi on farmers' chemical costs. Additionally, while increased local competition would ideally lead to increased profits, government policies such as market restrictions (e.g., the APMC Act) often limit farmers' ability to sell to other markets. This can lead to a saturation of the local market, preventing farmers from achieving efficient prices and thus dampening the potential for higher profits.

Interestingly, a significant reduction in the share of machinery cost in the total cost of production suggests that farmers are utilizing machinery more efficiently due to better access to local markets. Greater local competition typically leads to a higher concentration of machinery suppliers and service providers. This, in turn, can result in economies of scale for these suppliers, allowing them to reduce their product and service prices. This efficiency is then reflected in the lower machinery cost share for farmers. Additionally, a shorter distance to agricultural markets makes the maintenance and repair of farm machinery more efficient and increases access to machinery rental markets, thereby reducing farmers' dependency on outright machinery ownership.

The insignificant effect of distance to market on farm profits, despite reduced costs, suggests commensurately lower produce prices in local markets. This indicates that while the competitive local market effectively addresses the cost hurdle for farmers, its positive effects are not fully offset in terms of profits or dietary diversity. This implies that farmers might benefit from engaging with distant markets to encounter more competition and broader options for selling their produce. The availability of distant markets could also encourage farmers to diversify their agricultural products.

However, our observation that farmers are significantly growing more crops already suggests a tendency towards diversification. If farmers were to access distant markets, they might further diversify into higher-value crops or those that fetch greater profits. This could be achieved by leveraging local market information, efficiently utilizing local inputs, and exploiting the comparative advantages offered by local markets. By strategically combining these approaches, farmers could simultaneously benefit from local market efficiencies and implement profitable strategies to hedge against price fluctuations in distant markets.

Yet, farmers' ability to access these distant markets is not entirely within their control, as current agricultural market regulations often restrict them from entering such markets, even if they are comparatively nearer but located in other states. Therefore, government policy reforms are crucial to allow farmers to access distant markets. Some Indian states, such as

Gujarat, Maharashtra, and Karnataka, have already implemented such progressive policies. Since our study region is Uttar Pradesh and Madhya Pradesh, our findings particularly highlight that the inherently restrictive nature of existing market regulations needs to be reformed. This would enable farmers to capitalize on the benefits of distant markets, thereby realizing profits beyond the cost reduction advantages offered by competitive local markets. Such reforms would also facilitate product differentiation and information exchange on newer crops, ultimately making farming enterprises more profitable and providing farmers with superior economic opportunities.

Conclusion

In this paper, we rigorously examine the essential role of market proximity and its implications for farmers' production-side welfare, specifically through the lens of local versus distant market competition. Our findings consistently demonstrate that reduced distance to agricultural markets significantly enhances farmer welfare, primarily by fostering an environment of greater local competition.

For farmers whose nearest mandi is within their own state, signifying greater exposure to local competition, we observe substantial economic advantages. A one-unit (1 km) decrease in distance to the mandi is associated with a 32.6 percentage point reduction in total per-acre household costs and a 2.1 percentage point reduction in the share of machinery costs. This suggests improved production efficiency when farmers have easier access to nearby markets, likely due to reduced transaction costs, more efficient input procurement, and better access to machinery rental or maintenance services in a competitive local environment.

Furthermore, farmers operating under more local competition exhibit greater agricultural diversification, being statistically 49.8 percent more likely to cultivate different crops. This diversification is likely a strategic response to, or a benefit derived from, efficient local market access, potentially enabling better risk management and access to diverse market opportunities.

A deeper dive into input costs highlights key mechanisms: per-acre Kharif season, seed, and machinery costs are all statistically lower for farmers with closer market access. This indicates that direct market proximity leads to more cost-efficient procurement of crucial agricultural inputs, particularly those directly linked to output markets.

However, our research also uncovers a critical nuance: despite significant cost reductions, the insignificant impact on per-acre profit and dietary diversity suggests that these benefits are often offset by commensurately lower produce prices in saturated local markets. This highlights a constraint for farmers, even in a cost-efficient local market environment. While increased local competition improves price information and input efficiency, restrictive market regulations, such as the APMC Act, often limit farmers' ability to access more competitive distant markets. This can lead to local market saturation, preventing farmers from realizing the full profit potential of their reduced costs and diversified production.

In conclusion, our study provides compelling evidence that while local market access is instrumental in reducing production costs and encouraging diversification, current policy frameworks may inadvertently cap farmers' profitability by restricting access to distant markets. The experiences of states like Gujarat, Maharashtra, and Karnataka, which have initiated market reforms, offer a blueprint. Our findings using data from Uttar Pradesh and Madhya Pradesh underscore the urgent need for similar policy reforms. Allowing farmers to freely access distant markets, even across states, would enable them to capitalize on differential pricing, diversify into higher-value crops, and strategically hedge against price fluctuations. Such reforms are crucial not just for unlocking greater profits but for fostering product differentiation, facilitating information exchange on newer crops, and ultimately, building a more resilient and prosperous agricultural sector.

Reference(s)

Abdulai, A. N. (2016). Impact of conservation agriculture technology on household welfare in Zambia. *Agricultural Economics*, 47(6), 729-741.

Abebe, G. K., Bijman, J., & Royer, A. (2016). Are middlemen facilitators or barriers to improve smallholders' welfare in rural economies? Empirical evidence from Ethiopia. *Journal of Rural Studies*, 43, 203-213.

Adam, B., & Abdulai, A. (2024). Heterogeneous impact of crop diversification on farm net returns and risk exposure: Empirical evidence from Ghana. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*.

Addai, K. N., Temoso, O., & Ng'ombe, J. N. (2023). Heterogeneous effects of agricultural technology adoption on smallholder household welfare in Ghana. *Journal of Agricultural and Applied Economics*, 55(2), 283-303.

Akhtar, S., LI, G. C., Ullah, R., Nazir, A., Iqbal, M. A., Raza, M. H., ... & Faisal, M. (2018). Factors influencing hybrid maize farmers' risk attitudes and their perceptions in Punjab Province, Pakistan. *Journal of Integrative Agriculture*, 17(6), 1454-1462.

Awotide, B. A., Karimov, A. A., & Diagne, A. (2016). Agricultural technology adoption, commercialization and smallholder rice farmers' welfare in rural Nigeria. *Agricultural and Food Economics*, 4, 1-24.

Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, *23*(2), 192-210.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, *14*(4), 909-946.

Chatterjee, S. (2023). Market power and spatial competition in rural India. *The Quarterly Journal of Economics*, 138(3), 1649-1711.

Dey, S., & Singh, P. K. (2023a). Role of market participation on smallholder vegetable farmers' wellbeing: Evidence from matching approach in Eastern India. *Agribusiness*, 39(4), 1217-1237.

Dey, S., & Singh, P. K. (2023b). Market participation, market impact and marketing efficiency: an integrated market research on smallholder paddy farmers from Eastern India. *Journal of Agribusiness in Developing and Emerging Economies*.

Esaryk, E. E., Reynolds, S. A., Fernald, L. C., & Jones, A. D. (2021). Crop diversity is associated with higher child diet diversity in Ethiopia, particularly among low-income households, but not in Vietnam. *Public Health Nutrition*, 24(17), 5857-5868.

Ghimire, R., & Huang, W. C. (2016). Adoption pattern and welfare impact of agricultural technology: empirical evidence from rice farmers in Nepal. *Journal of South Asian Development*, 11(1), 113-137.

Haile, K., Gebre, E., & Workye, A. (2022). Determinants of market participation among smallholder farmers in Southwest Ethiopia: double-hurdle model approach. *Agriculture & Food Security*, 11(1), 18.

Hoq, M. S., Uddin, M. T., Raha, S. K., & Hossain, M. I. (2021). Welfare impact of market participation: the case of rice farmers from wetland ecosystem in Bangladesh. *Environmental Challenges*, 5, 100292.

Imbens, G., & Zajonc, T. (2011). Regression discontinuity design with multiple forcing variables. *Report, Harvard University.[972]*, 26.

John, A. T., Makkar, S., Swaminathan, S., Minocha, S., Webb, P., Kurpad, A. V., & Thomas, T. (2021). Factors influencing household pulse consumption in India: A multilevel model analysis. *Global food security*, 29, 100534.

Liang, Z., Zhang, L., Li, W., Zhang, J., & Frewer, L. J. (2021). Adoption of combinations of adaptive and mitigatory climate-smart agricultural practices and its impacts on rice yield and income: Empirical evidence from Hubei, China. *Climate Risk Management*, 32, 100314.

Ma, W., Sonobe, T., & Gong, B. (2024). Linking Farmers to Markets: Barriers, Solutions, and Policy Options. *Economic Analysis and Policy*.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698-714.

Meskel, T. W., Ketema, M., Haji, J., & Zemedu, L. (2020). Welfare impact of moringa market participation in Southern Ethiopia. *Sustainable Agriculture Research*, 9(3), 98-113.

Mmbando, F. E., Wale, E. Z., & Baiyegunhi, L. J. (2015). Welfare impacts of smallholder farmers' participation in maize and pigeonpea markets in Tanzania. *Food Security*, 7, 1211-1224.

Morrissey, K., Reynolds, T., Tobin, D., & Isbell, C. (2024). Market engagement, crop diversity, dietary diversity, and food security: evidence from small-scale agricultural households in Uganda. *Food Security*, 16(1), 133-147.

Shafiwu, A. B., Donkoh, S. A., & Abdulai, A. M. (2022). The welfare impact of improved seed variety adoption in Ghana. *Journal of Agriculture and Food Research*, 10, 100347.

Tesfaye, W., & Tirivayi, N. (2020). Crop diversity, household welfare and consumption smoothing under risk: Evidence from rural Uganda. *World Development*, 125, 104686.

Usman, M. A., & Callo-Concha, D. (2021). Does market access improve dietary diversity and food security? Evidence from Southwestern Ethiopian smallholder coffee producers. *Agricultural and Food Economics*, 9(1), 18.

Zheng, H., & Ma, W. (2023). Impact of agricultural commercialization on dietary diversity and vulnerability to poverty: Insights from Chinese rural households. *Economic Analysis and Policy*, 80, 558-569.

Appendix 1 (RD Assumption test)

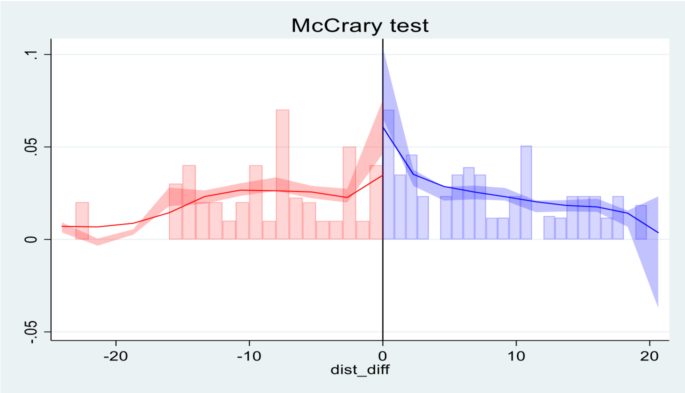

Figure A1: McCrary test

Table A1: RD Manipulation test (McCrary test)

Method	T	p>|T|
Robust	1.53	0.126

Note: We have used an unrestricted model with Epanechnikov kernel considering cut-off at zero

Appendix 2: Binned average of various input cost variables (Figure A2)

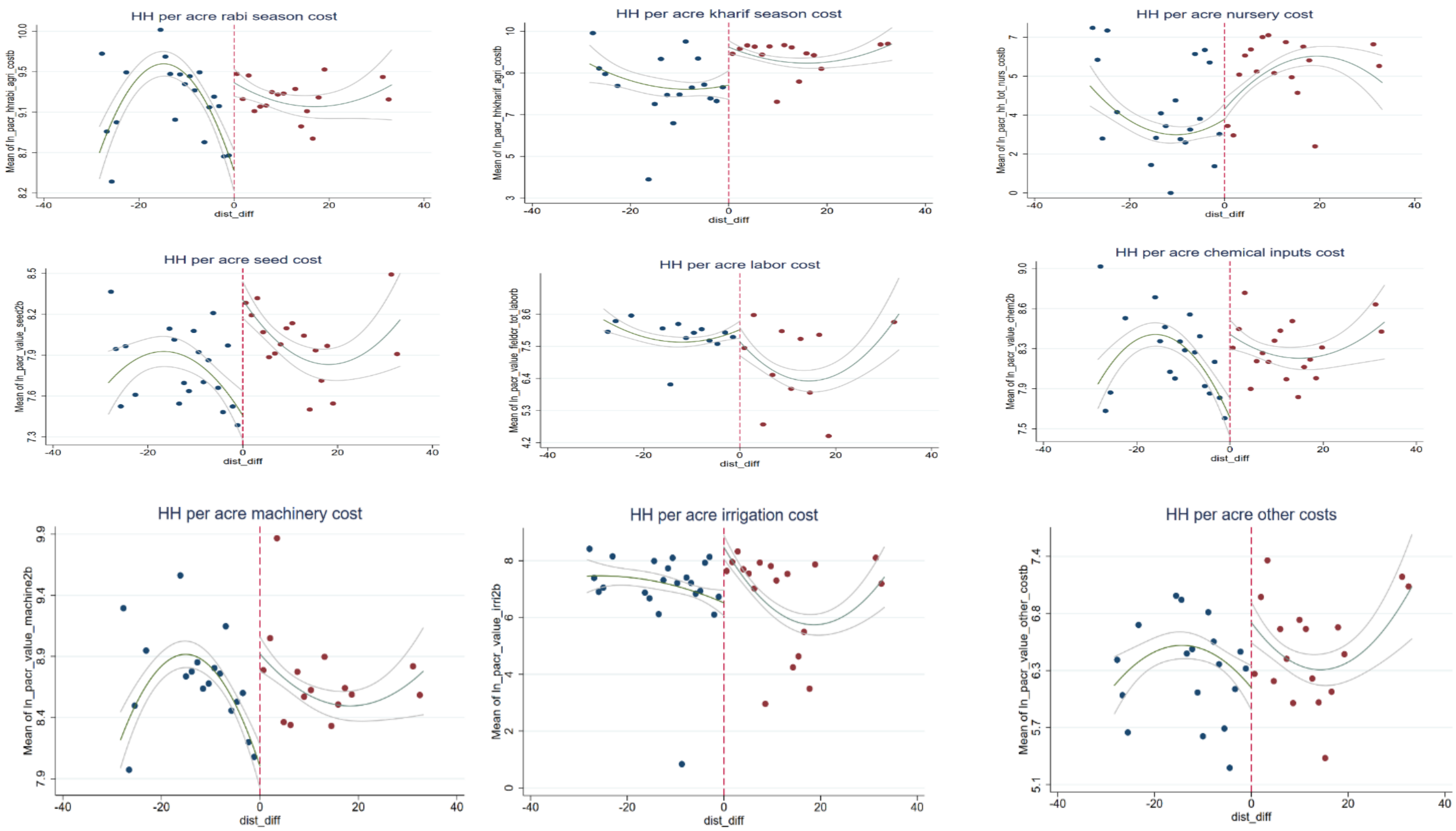

Table A2: Effect of distance on various input costs (Appendix 3)

Variable	Optimal bandwidth				Polynomial in distance		
	Observations	Bandwidth	Coefficient	Standard error	Observations	Coefficient	Standard error
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Per acre HH Rabi season cost	557	7.778	-0.600	(0.430)	1172	-0.587*	(0.320)
Per acre HH Kharif season cost	640	9.427	-0.851**	(0.344)	1172	-0.498	(0.323)
Per acre HH Zaid season cost	616	9.074	0.064	(0.096)	1172	0.015	(0.072)
Per acre HH total nursery cost	468	6.915	-1.541	(1.167)	1172	-0.985	(0.705)
Per acre seed cost	429	6.764	-0.715**	(0.282)	1172	-0.523**	(0.202)
Per acre chemical cost	492	7.223	-0.414	(0.282)	1172	-0.325*	(0.184)
Per acre machine cost	557	7.891	-0.512***	(0.096)	1172	-0.414**	(0.181)
Per acre irrigation cost	414	6.105	0.110	(0.592)	1172	-0.712	(0.829)
Per acre labor cost	557	7.824	1.113	(1.619)	1172	0.773	(1.360)
Per acre other costs	378	5.588	0.481	(0.312)	1172	-0.429	(0.390)

Note: Results use equation (1). Optimal bandwidths were chosen, as in Calonico, Cattaneo, and Titiunik (2014). Covariates were included in all the specifications. Covariates included were mean age and education of male and female household members, frequency of falling ill, dependency ratio, soil health card status of household, and household crop insurance participation. All specifications use APMC segment fixed effects and standard errors clustered at the village level. Polynomial order determined using Akaike's criterion. Significance level indicated are *p<0.1 **p<0.05 ***p<0.01.