Agricultural Mechanization Services, Adverse Selection and By-Stage Productivity of Small Farms

Evidence from Wheat Production in Northern China

Yu Sheng, Jiping Ding, Hangyu Zhang

Crawford School of Public Policy, ANU
College of Economics Management, Northwest AF University
China Center for Agricultural Policy, Peking University

14 May 2024
Washington D.C.

Introduction

Rapid growth of mechanization services has revolutionized farming practice of small-holders in developing countries (Zhang et al. 2017; Sheng et al. 2019; Diao et al. 2020).

- Promote the substitution of capital for labor through saving sunk costs related to capital investment.
- Standardize farming practices and facilitated advanced technology embodied in machinery and equipment.

However, many challenges preventing small farms from adopting mechanical services to improve agricultural productivity.

- High transaction costs (Daum et al. 2021; Foster and Rosenzweig,2022).
- Limited supply and Moral hazards (Caunedo et al. 2022; Foster and Rosenzweig,1994).

Questions remain on how "adverse selection" matter for mechanization service adoption and its effect farms' performance?

Introduction

This paper examines mechanization services in China and its impact on farm productivity, with the aim to:

- Investigate the adoption of mechanization services at different stages and assess their productivity effects, both at each stage and overall.
- Examine how "Principle Agent" problem may cause farms to opt for less and low-quality use of mechanization services in the plant protection stage.

We estimate a multiple-stage production function by using a unique farm-suvey data for wheat production in Northern China.

- Cover a balanced panel of 145 wheat farms for the 2013-2020 period.
- Compile the by-stage input-output data: ploughing/sowing, plant protection and harvesting.
- Distinguish between "mechanization services" and "self-owned machinery".

A theoretical model is also developed to explain the lack use of mechanisation services.

Background

China is the largest wheat producer in the world.

- In 2022, China's wheat sowing area and production are 23.57 million hectares and 137.7 million tons, representing 10.59% and 20% of the global total.
- Winter wheat in Northern China (in particular, Henan and Shandong) has accounted for around a half of total wheat production.

Wheat production in Northern China is dominated by small farms, and comprises three stages (About 240 days):

- Ploughing/sowing (10 days), plant protection (3 months) and harvesting (10 days)
- The ploughing/sowing and harvesting stages need to be on-time and are power intensive
- The plant protection stage (i.e. fertilizing, watering, weed and pest controls) is complex and take a longer time

Background: Wheat Production Stages

Stage 1 (ploughing/sowing):
Time: Within 10 days.
Tasks: Simple

ploughing

sowing

Stage 2 (plant protection) :
Time: 90 days.
Tasks:Complex.

fertilizing

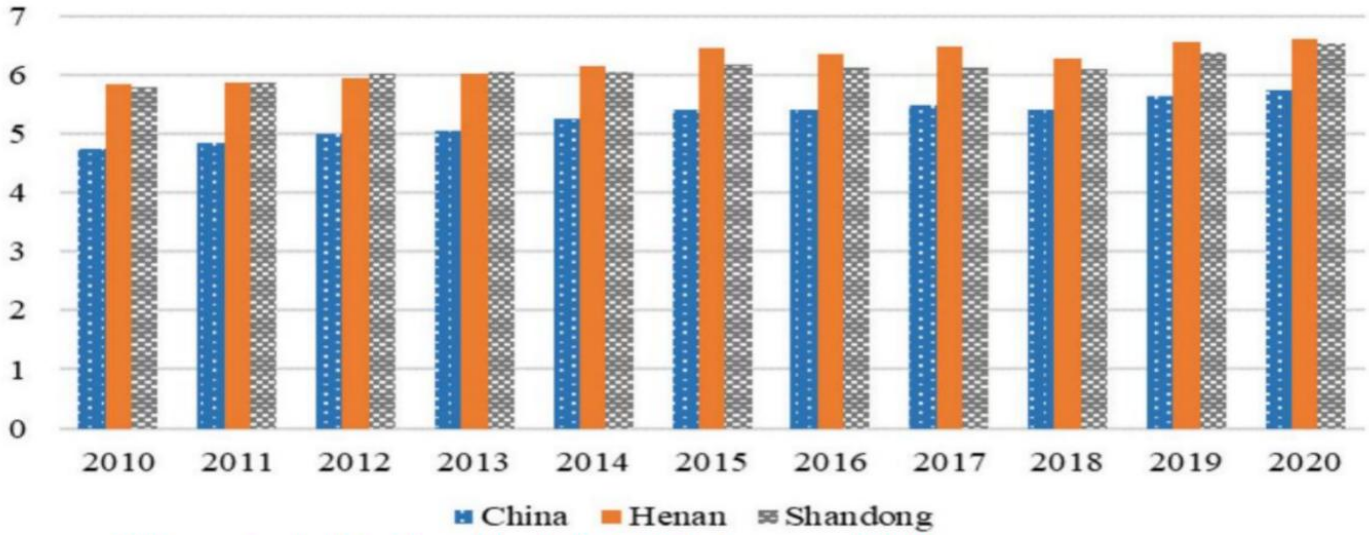
irrigation

plant protection

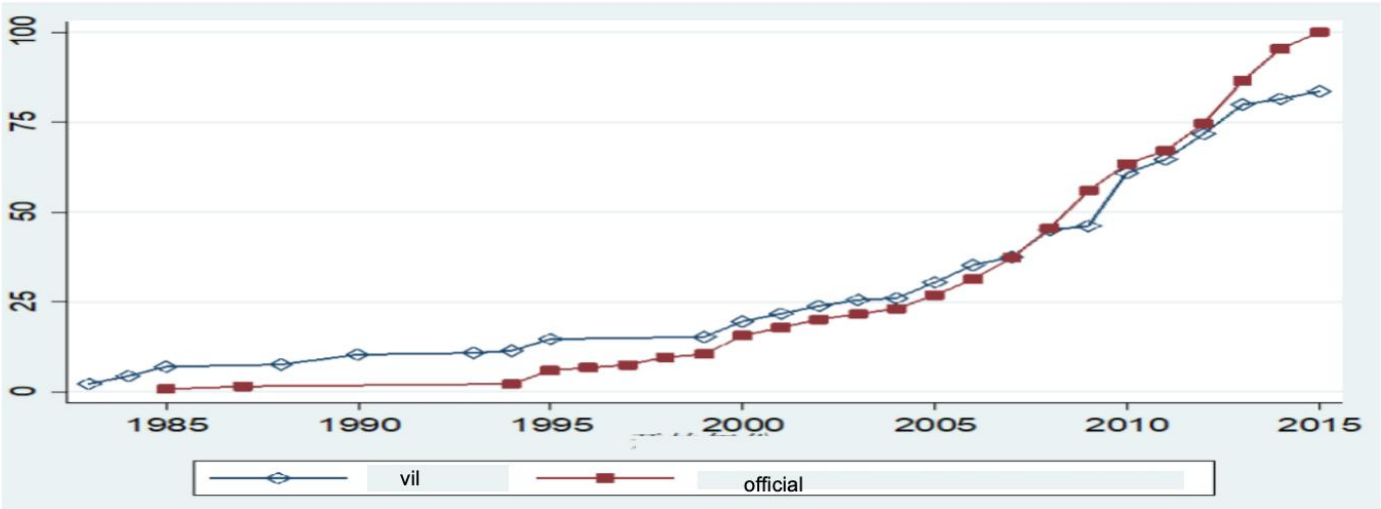
Stage 3 (harvesting):
Time: Within 10 days.
Task: Simple.

harvesting

Background: Yield Growth and Custom Service Diffusion

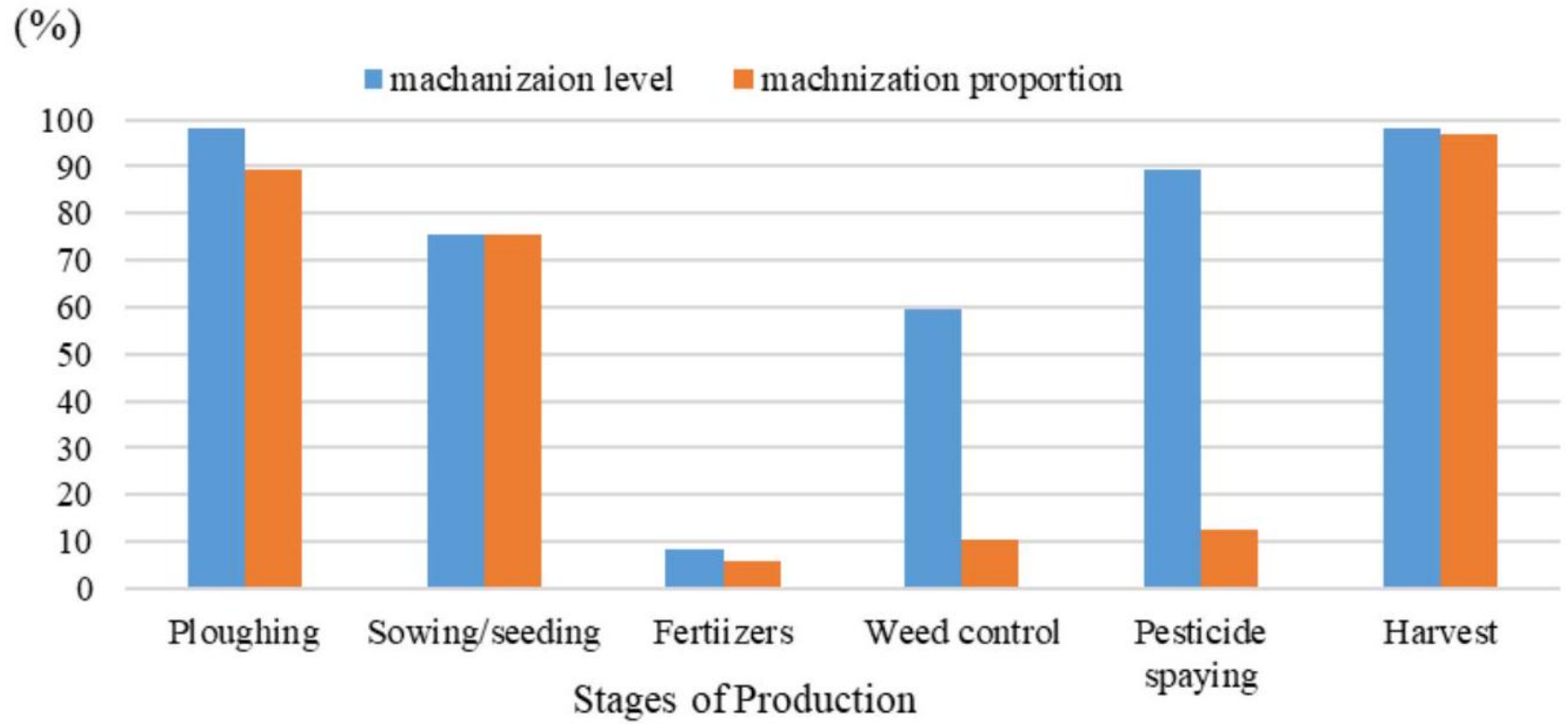


Wheat yield (ton/ha) [1.3% Annually]



Increased accessibility mechanization service by region (%)

Background: By-Stage Capital Usage And mechanization Service



Source: The CCAP farm survey data for 8 provinces in 2022.

Model and Method

A farm will maximize total profits by choosing capital and labor for each stages, given the predetermined output of the previous stage, such that:

$$\max \pi = pQ - \sum_{s=1}^{N} rK_s - \sum_{s=1}^{N} wL_s \tag{1}$$

$$\text{s. t. } \sum_{s=1}^{N} K_s < \bar{K}, \sum_{s=1}^{N} L_s < \bar{L}$$

- K_s : Capital in stage s , L_s : Labor in stage s
- $q_s = F(K_s, L_s)$
- $\bar{K}$ is the total capital. $\bar{L}$ is the total labor.
- r: rental price. w: labor wage. (Same across different stages.)

Model and Method

We assume that $F(K_s, L_s)$ is concave in (K, L) and take the first order condition, thus equilibrium factor prices equal to the marginal products of inputs

$$\begin{cases} w = \frac{\partial F\left(K_s^i, L_s^i L_{-1}\right)}{\partial L_s^i} \\ r = \frac{\partial F\left(K_s^i L_L^j; q_{-1}\right)}{\partial K_s^i} \end{cases} \quad \forall i \in \{1, 2, \ldots, N\}, \forall s \in \{1, 2, \ldots, n\} \tag{2}$$

Assume that $L_1^i = L_2^i = L_3^i = L^i$ (labor is fixed in each stage), we can normalized capital intensity (or capital-labor ratio) for each stage as:

$$k_s^i = \frac{K_s^i}{L_s^i} = \frac{K_s^i}{\bar{L}^i} \tag{3}$$

Denote A_s is an effiecicy paramter, the by-stage production function can be simplified as:

$$F(K_s, L_s; q_{-1}) = A_s \, \mathrm{h}\left(k_s^i; q_{-1}\right) \tag{4}$$

Model and Method

By incorporating the two parameter α and T into the by-stage production function, we have:

- $F\left(K_s, \overline{L_s}; q_{-1}\right) = A_s(T,a)\mathrm{h}\left(k_s^i; q_{-1}\right)$, where $\mathrm{h}'(.) > 0, \mathrm{h}''(.) < 0$

We assume that efficiency of production

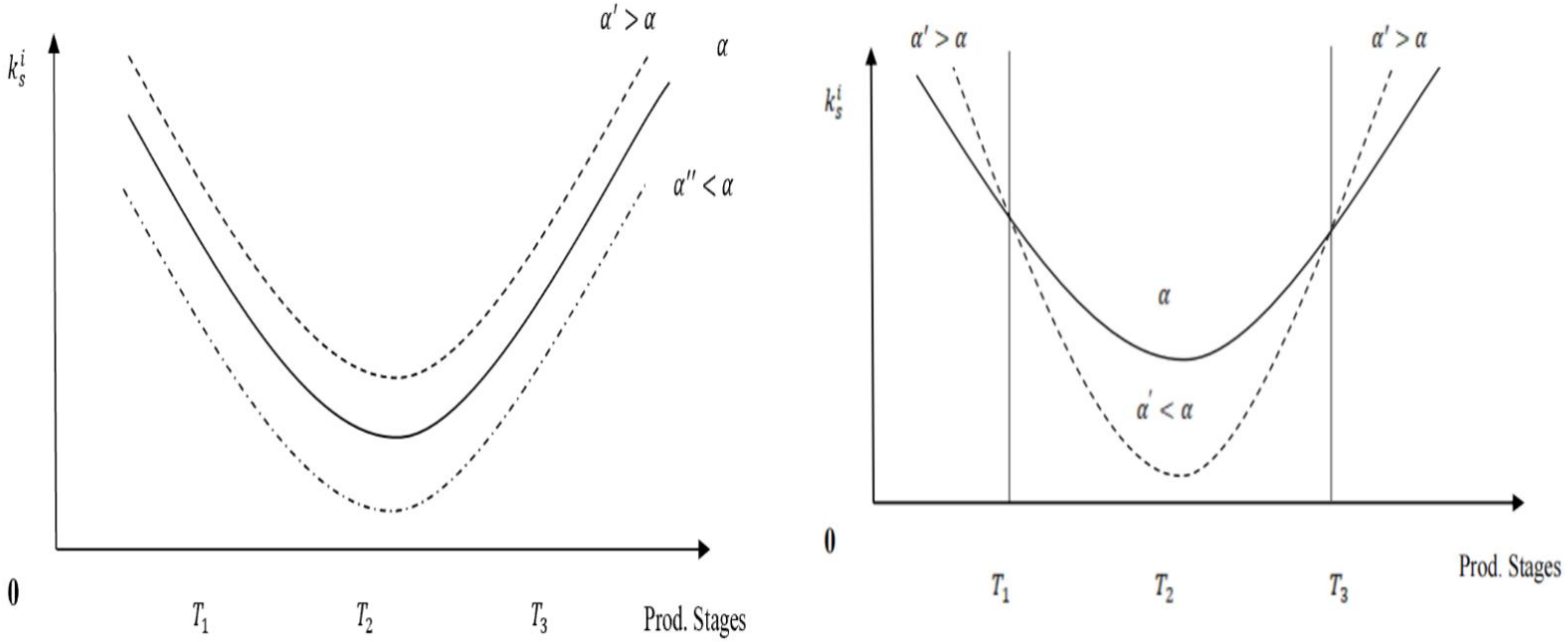
- Decreases with the length of time T_s such that $\frac{\partial A_s}{\partial T_s} < 0$
- Increases with capital deepening $\frac{\partial A_s}{\partial a} > 0$.

We assume that productivity gains brought by specialization in mechanization level will decrease with T (i.e. $\frac{\partial^2 A_s}{\partial a \partial T_s} < 0$), such that:

- $k_s^i = h'^{-1}\left(\frac{r}{A'(T_m,a)}\right) \rightarrow \frac{\partial k_s^i}{\partial a} > 0$ and $\frac{\partial k_s^i}{\partial T_s} < 0$

Model and Method

Self-owned machinery vs. mechanization service by production stages



(a) self-owned machine α (b) mechanization custom service α'

$\alpha' < \alpha$, tasks are complex and time is longer in plant protection stage

Model and Method

We use the output of the previous stage as the input of the next stage.

- The empirical specification of the three-stage production function for wheat production can be written as:

$$
\begin{aligned}
Y_{i1t} &= e^{w_{1it}+\varepsilon_{i1t}} L_{i1t}^{\alpha_1} K_{i1t}^{\beta_1} N_{i1t}^{\tau_1} \\
Y_{i2t} &= e^{w_{2it}+\varepsilon_{2it}} L_{i2t}^{\alpha_2} K_{i2t}^{\beta_2} N_{i2t}^{\tau_2} R_{i2t}^{\gamma_2} \\
Y_{i2t} &= e^{w_{3it}+\varepsilon_{3it}} L_{i3t}^{\alpha_3} K_{i3t}^{\beta_3} N_{i3t}^{\tau_3} R_{i3t}^{\gamma_3}
\end{aligned}
$$

  where $Y_{1it} = R_{2it}, Y_{2it} = R_{3it}$.

- Taking the log form:

$$
\begin{aligned}
y_{i1t} =& \alpha_1 l_{i1t} + \beta_1 k_{i1t} + \tau_{i1t} n_{i1t} + w_{i1t} + \varepsilon_{i1t} \\
y_{i2t} =& \alpha_2 l_{i2t} + \beta_2 k_{i2t} + \gamma_2 r_{i2t} + \tau_{i2t} n_{i2t} + w_{i2t} + \varepsilon_{i2t} \\
y_{i3t} =& \alpha_3 l_{i3t} + \beta_3 k_{i3t} + \gamma_3 r_{i3t} + \tau_{i3t} n_{i3t} + w_{i3t} + \varepsilon_{i3t}
\end{aligned}
$$

  where $y_{i1t} = r_{i2t}, y_{i2t} = r_{i3t}$.

Model and Method:Econometric Issues

We use the control function approach to restore unobserved TFP.

- Olley and Parks (1995), Levin and Petrin (2015) and Achberk et al. (2019) etc.

The demand functions for intermediate inputs of the three production stages $\phi_{st}(\cdot), s = 1, 2, 3.$ are thus written as follows:

$$
\begin{aligned}
e_{i3t} &= \phi_{3t}\left(k_{i3t}, l_{i3t}, n_{i3t}, w_{i3t}\right) \\
e_{i2t} &= \phi_{2t}\left(k_{i2t}, l_{i2t}, w_{i2t}, k_{i3t} l_{i3t}, w_{i3t}, n_{it}\right) \\
e_{i1t} &= \phi_{1t}\left(k_{i1t}, l_{i1t}, w_{i1t}, k_{i2t}, l_{2t}, w_{i3t}, k_{i3t}, l_{i3t}, w_{i2t}, n_{it}\right)
\end{aligned}
$$

Inverse the above equations provide the control functions (or the inverted demand function) for the three production stages:

$$
\begin{aligned}
w_{i3t} &= \phi_{3t}^{-1}\left(k_{i3t}, l_{3jt}, e_{3jt}\right) \\
w_{i2t} &= \phi_{2t}^{-1}\left(k_{i2t}, l_{i2t}, e_{i2t}, k_{i3t}, l_{i3t}, \hat{w}_{i3t}\right) \\
w_{i1t} &= \phi_{1t}^{-1}\left(k_{i1t}, l_{1jt}, e_{1jt}, k_{2jt}, l_{2jt}, \hat{w}_{i3t}, k_{3jt}, l_{3jt}, \hat{w}_{i2t}\right)
\end{aligned}
$$

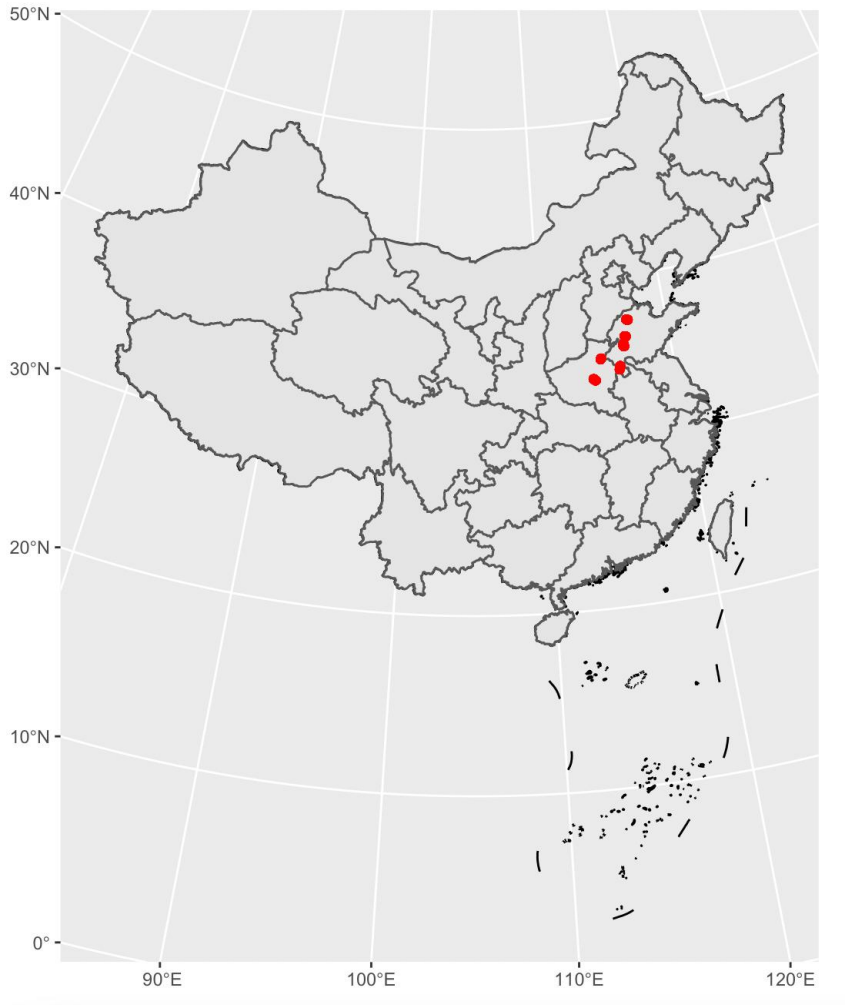
A set of sensitivity analyses are carried out.

Model and Method

$$w_{its} = \beta_0 + \beta_1 CS_{\text{ratio},its} + \beta_2 \left(\frac{K}{L}\right)_{its} + \gamma Z_{its} + u_{is} + v_t + e_{its} \quad (5)$$

- w_{its} is the by-stage productivity estimates
- $\left(\frac{K}{L}\right)_{its}$ refers to the capital-labor ratio in each stage of production
- $CS_{\text{ratio},its}$ refers to the proportion of mechanization services in capital input.
- u_{is} denotes the farm and stage fixed effects
- v_t denotes technology progress common to all wheat farms over time.
- e_{its} is the residual values.

Data Source: Sample maps and methods



We conduct three waves of farm survey.
Henan (120 households)

- County: Fengqiu, Yucheng, and Linying. 6 townships and 12 villages (2 in each town).

Shandong (120 households)

- County: Linyi, Wenshang, and Feicheng. 6 townships and 12 villages (2 in each town)

Outliers were excluded due to incomplete data, outliers, and other statistical reasons.

Data Source: Definition of Major Variables

Wheat production is categorized into three stages

- Ploughing and seedling
- Plant protection: fertilizing, weeding and pest control, irrigation
- Harvesting: collecting, drying and packaging

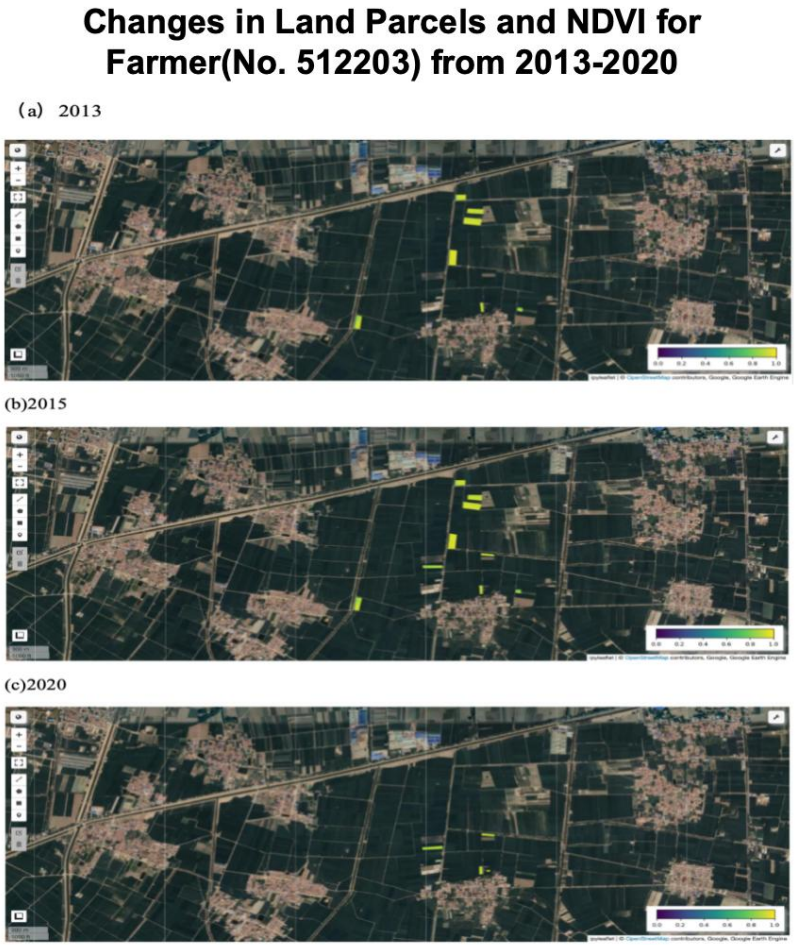
Input and Output

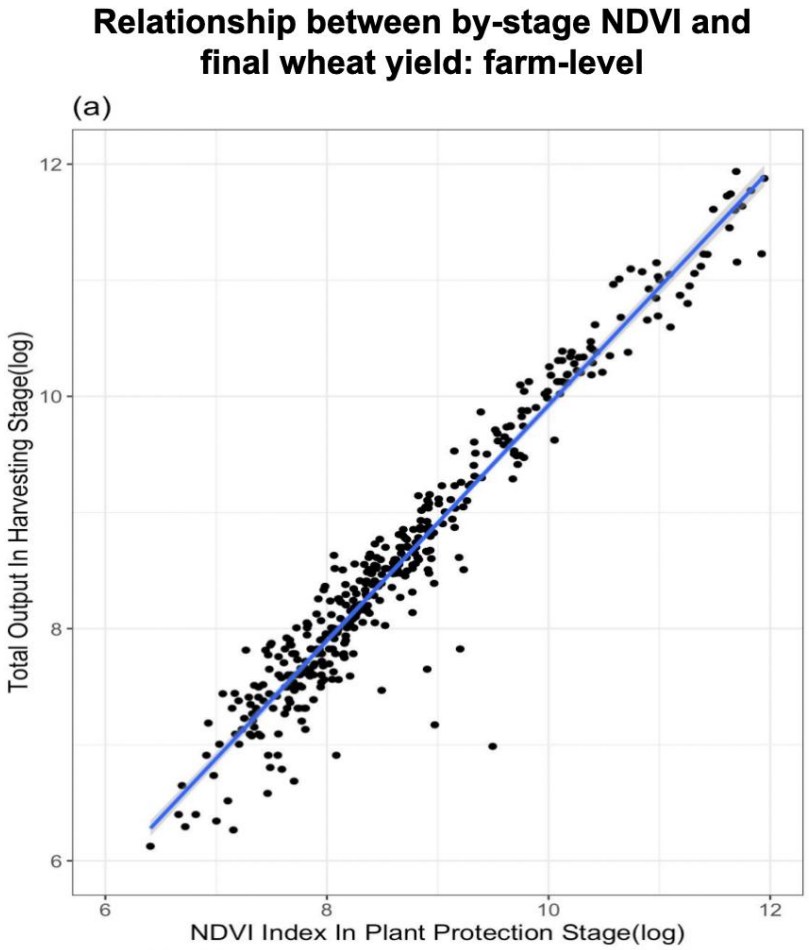
- Agricultural output: wheat harvested (ton), NDVI index.
- Agricultural input (cost in real term): machinery use (tractor, harvester...), labor input and intermediate material (seed, fertilizer...).

Mechanization service

- Relative ratio of mechanization service costs in total capital costs (in real term)

Data Source: By-stage NDVI Calculation





Note: The image on the left shows the changes in plot boundaries over three years for household ID 512203, as well as the maximum NDVI values for the plant protection stages on the plot. The image on the right displays scatter plots and linear regression lines for second-stage NDVI versus third-stage output.

Empirical Results

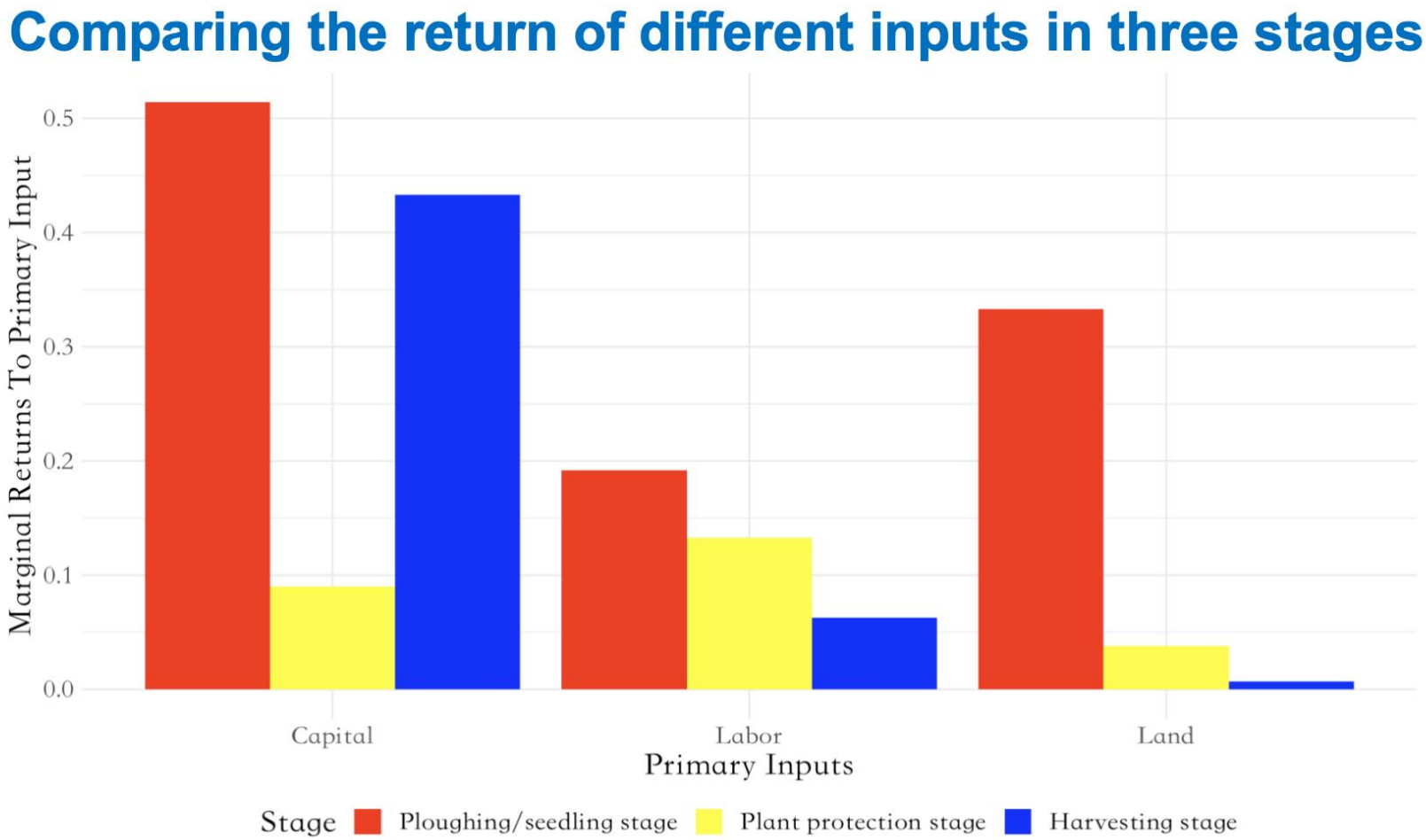
1. Multi-stage production function estimation
   - Comparing the returns to various inputs in three stages
   - Comparing the distribution of by-stage farm TFP estimates
   - Comparing the distribution of by-stage aggregation and farm-level
2. Linking the capital-labor ratio and mechanization service proportion to by-stage productivity
3. Robustness checks

Empirical Results: Multi-stage production function estimation

	Plowing/Seedling			Plant Protection			Harvesting		
	OLS	OLSW	GMM	OLS	OLSW	GMM	OLS	OLSW	GMM
K	0.609***	0.556***	0.514***	0.012***	0.018***	0.090***	0.458***	0.485***	0.433***
	(0.047)	(0.011)	(0.076)	(0.004)	(0.001)	(0.001)	(0.026)	(0.006)	(0.031)
L	0.034	0.019**	0.192***	0.185***	0.183***	0.133***	0.047***	0.061***	0.063***
	(0.034)	(0.008)	(0.017)	(0.015)	(0.003)	(0.037)	(0.015)	(0.004)	(0.000)
M	-	-	-	0.654***	0.641***	0.741***	0.462***	0.431***	0.450***
	-	-	-	(0.021)	(0.005)	(0.115)	(0.031)	(0.007)	(0.000)
A	0.299***	0.309***	0.333***	0.144***	0.135***	0.038	0.039**	0.045***	0.007
	(0.038)	(0.009)	(0.088)	(0.019)	(0.004)	(0.039)	(0.020)	(0.005)	(0.009)
County-FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Return to scale	0.90	0.88	1.03	1.00	0.98	1.00	1.01	1.02	0.95
Observations	435	435	435	435	435	435	435	435	435

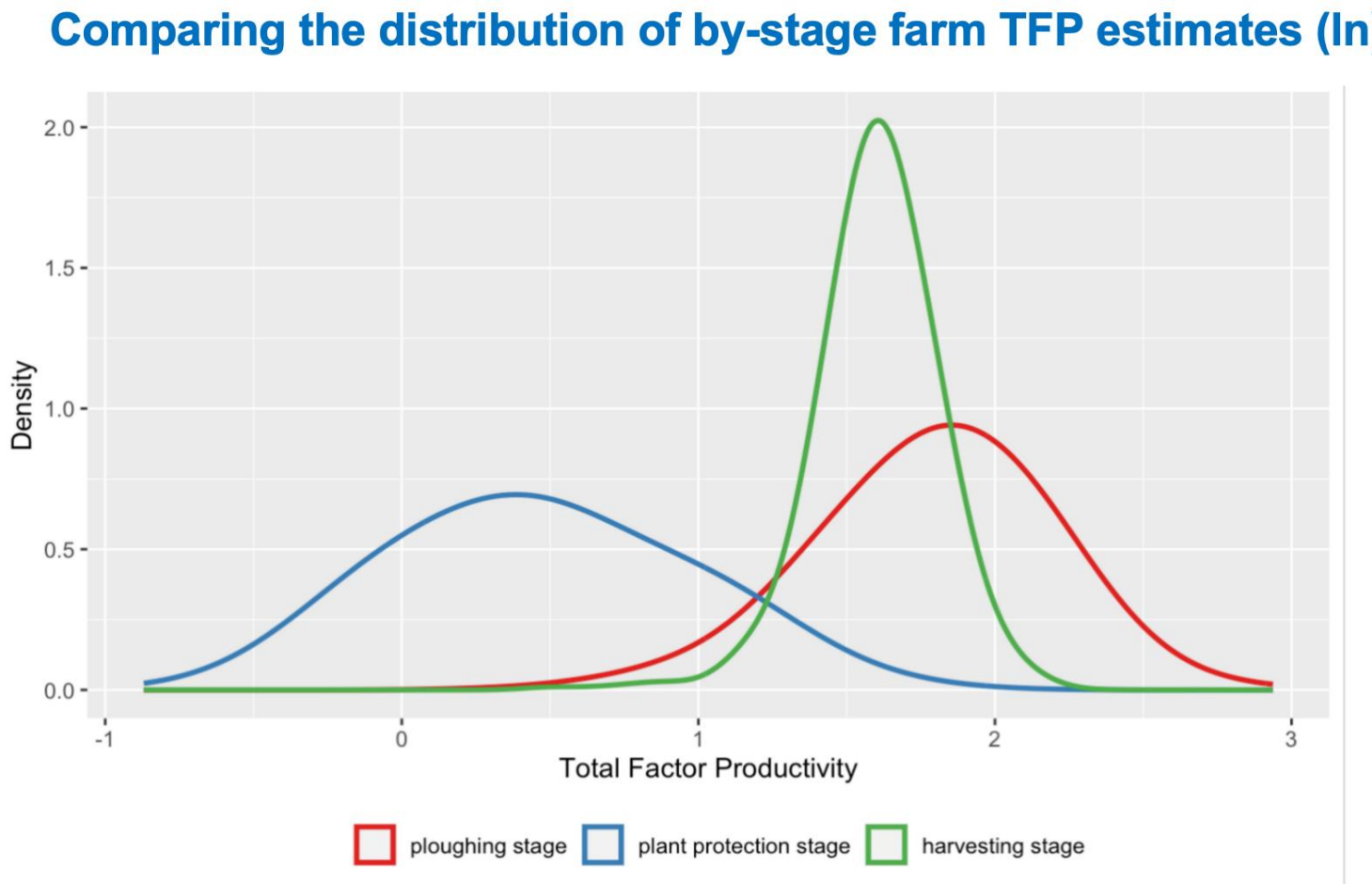
Note: In the first column, K, L, M and A represent capital, labor, intermediate input and land input used for wheat farm production in Northern China. OLS and OLSW refer to the farm fixed effect model with and without adjustment for farm weights, and GMM refers to the GMM model. The numbers in parentheses below the estimated coefficients are standard errors. The symbols ,, and denote significance at the 1%, 5%, and 10% levels.

Empirical Results: Comparing the returns to various inputs in three stages



It remains a question why returns to capital and labor are so different across different production stages?

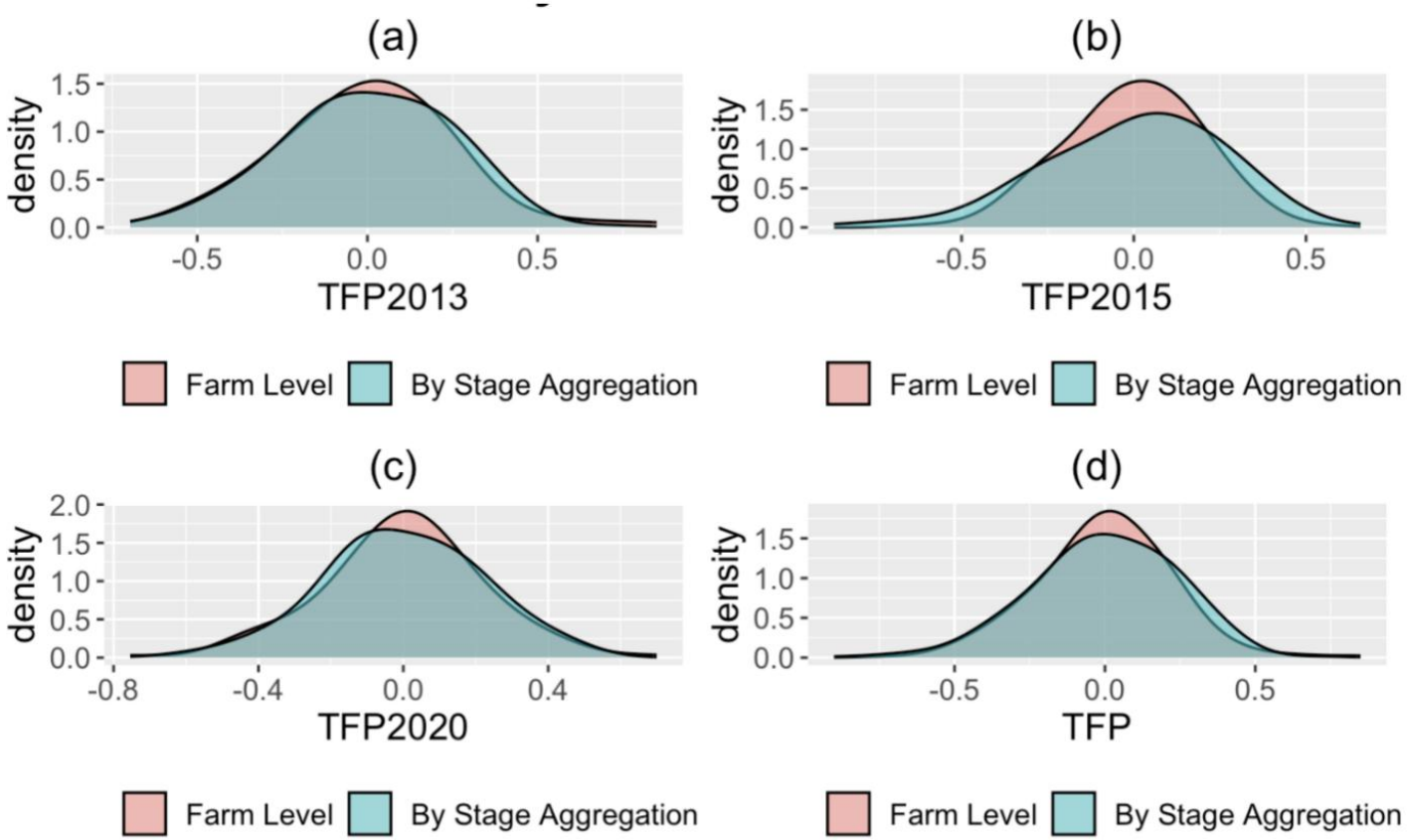
Empirical Results: Comparing the distribution of by-stage farm TFP estimates



Source:Authors make estimation by using the by-stage production function.
Note: Red line is the ploughing and sowing stage. Blue line is the plant protection stage. Green Line is the harvesting stage.

Empirical Results: Comparing the distribution of by-stage aggregation and farm-level

Comparing the distribution of aggregated by-stage TFP estimates with farm-level TFP estimates by year.



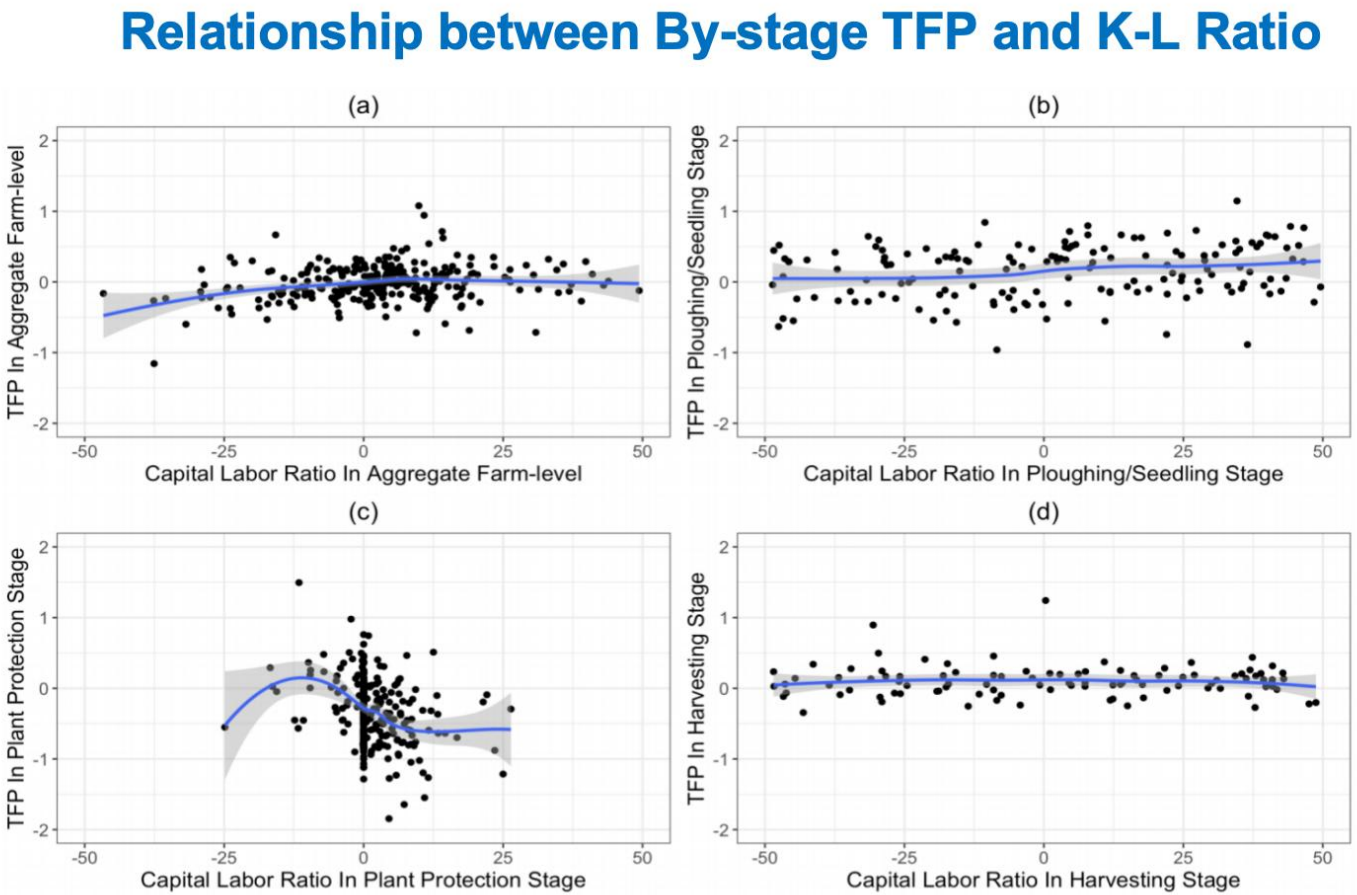
Source:Authors make estimation by using the by-stage production function.

Empirical Results

1. Multi-stage production function estimation
   - Comparing the returns to various inputs in three stages
   - Comparing the distribution of by-stage farm TFP estimates
   - Comparing the distribution of by-stage aggregation and farm-level

2. Linking the capital-labor ratio and mechanization service proportion to by-stage productivity
3. Robustness checks

Empirical Results: Linking the capital-labor ratio and custom service ratio with by-stage productivity



Note: x(capital labor ratio) and y(TFP) are both subtracted the values from period t-1 at the household level. Panel (a) represents the farm-level TFP against the capital-labor ratio. Panel (c)-(d) represent the ploughing/seedling stage, the plant protection stage, and the harvesting stage respectively.

Empirical Results: Linking the capital-labor ratio and custom service ratio with by-stage productivity

	(1) Ploughing/Seedling	(2) Plant Protection	(3) Harvesting	(4) Overall
K-L ratio	0.001**	-0.009***	0.000	-0.001
	(0.001)	(0.003)	(0.000)	(0.001)
Mechanization proportion	0.004**	-0.002**	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Farm-level fixed effects	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y
Number of Observations	435	435	435	435
R-squared	0.744	0.708	0.452	0.545

Note: The dependent and independent variables in Column (4) are computed at the aggregate level, while in Columns (1)-(3), are computed with respect to the ploughing/seedling stage, the plant protection stage, and the harvesting stage, respectively.

Empirical Results

1. Multi-stage production function estimation
   - Comparing the returns to various inputs in three stages
   - Comparing the distribution of by-stage farm TFP estimates
   - Comparing the distribution of by-stage aggregation and farm-level
2. Linking the capital-labor ratio and mechanization service proportion to by-stage productivity
3. Robustness checks

Empirical Results: Robustness Checks

Re-do the estimation by adjusting for the difference in labor wages by production stages.

- Although there are wage differences for labor across different stages, it does not affect our estimation.

We re-estimate the by-stage NDVI by using the quadratic function form (rather than a linear function form).

- We use a nonlinear relationship between wheat yield and the NDVI index to estimate the by-stage output index.

The GMM estimation of the multiple-stage production function could be sensitive to the choice of initial parameters (Brandt et al., 2022).

- We expand the range for the initial value of input elasticity.

Empirical Results: Wage Differences Across Different Stages

Table 3: Median Wages for Various Agricultural Production Stages (2013-2020)

Activity	Average (yuan/hour)	2013 (yuan/hour)	2015 (yuan/hour)	2020 (yuan/hour)
Plowing	15.39	13.85	15.39	28.13
Sowing	16.14	18.64	11.50	12.98
Fertilizing	13.49	10.00	14.78	20.92
Weeding	12.48	10.07	19.76	12.48
Pesticide Application	13.83	10.00	26.67	19.90
Harvesting	12.12	10.84	8.83	48.83

Note: All wages are calculated at the 2013 constant price. The percentage of households having employment labor in plowing, sowing, fertilizing, weeding, pesticide application, and harvesting was 5.8%, 6.6%, 9.9%, 11.61%, 12.2%, and 6.5%

Empirical Results: Robustness Checks

	RC1			RC2			RC3		
	S1	**S2**	**S3**	**S1**	**S2**	**S3**	**S1**	**S2**	**S3**
K	0.454***	0.102***	0.395***	0.531***	0.094***	0.403***	0.524***	0.155***	0.323**
	(0.022)	(0.009)	(0.023)	(0.030)	(0.004)	(0.040)	(0.148)	(0.059)	(0.144)
L	0.262***	0.117***	0.105***	0.255***	0.169***	0.085***	0.210***	0.229*	0.110
	(0.012)	(0.014)	(0.000)	(0.008)	(0.032)	(0.000)	(0.030)	(0.134)	(0.098)
M	-	0.690***	0.450***	-	0.683***	0.450***	-	0.526**	0.471***
	-	(0.055)	(0.000)	-	(0.069)	(0.000)	-	(0.201)	(0.098)
A	0.318***	0.102*	0.000	0.245***	0.083**	0.012	0.251	0.116	0.058
	(0.022)	(0.055)	(0.000)	(0.021)	(0.034)	(0.011)	(0.158)	(0.165)	(0.163)
County-FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Return to scale	1.03	1.03	0.95	1.03	1.01	0.95	0.98	1.02	0.96
Observations	435	435	435	435	435	435	435	435	435

Note: RC1 re-do the estimation by adjusting for the difference in labor wages by production stages. RC2 re-estimate the by-stage NDVI by using the quadratic function form.RC3 expand the range for the initial value of input. S1 is ploughing/sowing stage. S2 is the plant protection stage. S3 is the harvesting stage.

Conclusions

Mechanization services could negatively affect farm productivity in the plant protection stage.

- hampers the further increase of capital intensity
- causes a decline in overall farm productivity

The existence of adverse selection by farmers could be a potential cause.

- impede the substitution of mechanization services
- generate negative spill-over effects downstream production stages

These findings provide insights in regulating the market for mechanization services.

- An optimal choice of by-stage mechanization service may not help to enhance productivity.

Questions and Comments