



# Efficiency or equity? The effect of an exogenous agricultural commercial shock: Evidence from Manchuria in the 1930s

Tianyu Huang<sup>a,b,\*</sup>, Nan Li<sup>c</sup>

<sup>a</sup> School of Economics, Guangxi University, 100 Daxuedong Road, Nanning, Guangxi Province 530004, China

<sup>b</sup> China Frontier Economic Research Institute, Guangxi University, 100 Daxuedong Road, Nanning, Guangxi Province 530004, China

<sup>c</sup> School of Economics, Fudan University, 600 Guoquan Road, Shanghai 200433, China

## ARTICLE INFO

### JEL classification:

N55  
O13  
Q12  
Q13

### Keywords:

Cash crops  
Agricultural productivity  
Welfare  
Inequality  
Factor markets

## ABSTRACT

The adoption of high-value cash crops (HVCs) is considered an efficient way to improve farmers' productivity and welfare, although the systematic empirical evidence is limited. By exploiting an exogenous commercial shock for soybeans and microlevel rural survey data from Manchuria of China in the 1930s, where factor markets were relatively well functioning, this study investigates the influence of HVC cultivation on farmers' agricultural performance, welfare, and inequality both theoretically and empirically. We find that (i) larger farms devote a larger proportion of land to HVCs and that (ii) farmers who cultivate a larger proportion of HVCs have higher agricultural productivity, profit, net income, and expenditure. Furthermore, the adoption of HVCs lowers profit, income, and expenditure inequality by enhancing income from wages relative to land rents, which benefits households with less land.

## 1. Introduction

The improvement of productivity, welfare, and equity in rural areas are major issues for economists and policy makers. As the rural economy develops, production systems evolve out of autarky and become more integrated into commercial markets. Therefore, with the development of commercialization, transformation from traditional grain crops to high-value cash crops (HVCs; such as fruit, cotton, coffee beans, and medicinal herbs) is considered an effective way to improve the productivity and welfare of smallholder farmers (Birthal, Roy, & Negi, 2015; Rivera-Padilla, 2020; Weinberger & Lumpkin, 2007).<sup>1</sup> However, in practice, many farmers, especially smallholders, still grow mainly low-value traditional grain crops. Accordingly, this study focuses on two urgent questions. First, what are the determinants of farmers' crop choices? Second, if only some farmers respond to the market opportunities of HVCs, then what impact will this have on agricultural productivity, especially on farmers' welfare and inequality?

The existing literature have investigated several determinants of farmers' crop choice, such as production ability, subsistence, risk coping, trade cost, and institutions.<sup>2</sup> A recent branch of studies integrates such choice into the context of frictional factor markets

\* Corresponding author at: School of Economics, Guangxi University, 100 Daxuedong Road, Nanning, Guangxi Province 530004, China.

E-mail addresses: [tyhuang@gxu.edu.cn](mailto:tyhuang@gxu.edu.cn) (T. Huang), [nanli@fudan.edu.cn](mailto:nanli@fudan.edu.cn) (N. Li).

<sup>1</sup> Since there is no strict definition of cash crops, this study defines those crops that are used mainly for sale rather than for self-consumption and whose unit price is higher than that of main food crops (sorghum and millet) as cash crops, including beans such as soybeans, wheat, perilla-seed, rice, peanuts, etc. in Manchuria in the 1930s. See footnote 24 for more information.

<sup>2</sup> See Fafchamps (1992); Kebede (2022); Montero (2022); Rivera-Padilla (2020).

(especially land) since they are prevalent in the developing world.<sup>3</sup> Empirically, although the results of existing studies tend to hold a positive attitude toward HVC adoption, there are still some limitations to the few systematic investigations for causality.<sup>4</sup> First, as a type of endogenous choice of farmers, crop choice is affected by a variety of observable and unobservable factors, so the estimated results are inevitably threatened by the potential endogenous bias caused by omitting variables, especially for the nationwide sample.<sup>5</sup> Second, since crop choice is endogenous, different farmers react differently to the same market opportunity. Especially in relatively competitive factor markets, the existing research on whether the introduction of HVCs aggravates welfare inequality among farmers is scarce.<sup>6</sup>

To address the above limitations, this paper exploits the booming international soybean trade in the early 20th century as an exogenous shock of commercialization. We then empirically investigate the causal effects of HVCs (mainly soybeans and wheat) adoption on farmers' performance such as agricultural profits, factor productivity and welfare indicators like income, expenditure and the corresponding inequality, with the household-level survey data collected by the Survey Department of South Manchuria Railways Co. in Manchuria (Northeast China) in the 1930s.

To derive a reliable causal effect of HVC cultivation on productivity, welfare, and its inequality in the context of free factor markets, a major challenge is to find the place, where, (1) is suitable for growing HVCs, (2) the shock of commercial opportunity is relatively exogenous, (3) has high factor mobility. Manchuria, where was forced to opened to international markets in the late 19th century in the early 20th century, provides us such an appropriate case. Since there was suitable for soybean cultivation and the exogenous positive demand shock from Europe was very strong, soybeans became the most important commodity for export in that period. In addition, at that time, factors, especially land, can be redistributed without any institutional restrictions in the market by means of sale, mortgage, tenancy, etc.<sup>7</sup>

To deal with the endogeneity of crop choice for HVCs, we rely on the variation in soil suitability for crops across regions to construct exogenous variables. Specifically, we use estimates of potential crop yields across surveyed villages from the Food and Agriculture Organization (FAO)'s Global Agro-Ecological Zones (GAEZ) database. These yields are calculated by incorporating local soil and weather characteristics into a model that predicts the maximum attainable yields for each crop in a given area. Because of its exogeneity, this potential yield has been used as an instrumental variable (IV) in some important studies (e.g. [Bustos, Caprettini, & Ponticelli, 2016](#)).<sup>8</sup> To further measure the competitive advantage of cash crops over traditional grain crops, we calculate the ratio of potential yields between the major cash (almost soybeans) and grain crops (almost sorghum or millet) in each surveyed village as the IV of crop choice.

To guide the empirical work, we first build a static general equilibrium model in which there is no friction in the labor and land markets within a village. The model predicts that farmers will grow a higher proportion of HVCs and earn higher profits when they have higher production ability, face a higher price ratio of soybeans to food crops, and live in localities where are more suitable for growing soybeans. Furthermore, a higher proportion of HVCs reduces the income gap between landlords and farmers within a village by enhancing income from wages relative to land rentals, which benefits households with less land endowment under the assumption that HVCs are more labor intensive than traditional grain crops. The consequences mean that both efficiency and equity can be derived.

The empirical results are consistent with the model's predictions. Our baseline ordinary least squares (OLS) estimates and two-stage least squares regression with IV (2SLS-IV) find that the proportion of land grown HVCs for households is positively correlated with agricultural performance indicators such as per capita farm profit, labor and land productivity (calculated by value added) and welfare indicators such as per capita expenditure on living and per capita net cash income. The results hold in several robustness checks by exploiting a cohort difference-in-differences (DID) estimate, subsamples of different regions, operating farm sizes and soybean suitability at the village level, and dropping the top and bottom 5% of observations.

Furthermore, we examine the corresponding effect on inequality from two dimensions. First, the heterogeneity analysis using quantile regression shows that the positive effect of HVC adoption on agricultural profits is more significant for farmers in the lowest profit quantile, while the effect on welfare indicators such as expenditure and income shows little difference across different quantile groups. Second, the analyses at the village level reveal that the higher the proportion of land grown HVCs is, the lower the Gini indices within the village, indicating that cultivating more labor-intensive HVCs in free labor and land markets can reduce the degree of inequality (and, at the very least, not exacerbate it).

The behavior of farmers in production decisions is a traditional issue in the microeconomics literature. Generally, the crop, input, technique choices of a utility-maximizing farmer are determined by her preferences on consumption and risk ([Fafchamps, 1992](#);

<sup>3</sup> See [Adamopoulos and Restuccia \(2020\)](#) for a review.

<sup>4</sup> See [Birthal et al. \(2015\)](#); [Kung and Li \(2011\)](#); [Ogutu and Qaim \(2019\)](#); [Zhang et al. \(2021\)](#).

<sup>5</sup> [Birthal, Hazrana, and Negi \(2020\)](#) point out that due to the strong heterogeneity of natural and social environment in various regions, only controlling for regional fixed effects at a macro level (such as the city level) is likely to fail to completely eliminate the problem of omitting variables at a smaller level (such as at the county and village levels).

<sup>6</sup> Through quantile regression, [Ogutu and Qaim \(2019\)](#) find that the group with higher per capita income is more positively affected by commercialization, thus raising concerns about inequality. However, they did not conduct further investigation.

<sup>7</sup> [Brandt \(1987\)](#) and [Benjamin and Brandt \(1997\)](#) argue that factor markets in North and Northeast China in the 1920s and 1930s were not perfect but were free and efficient.

<sup>8</sup> Existing empirical studies of this topic use mainly the crop choices of other farmers in the same village or the number households' transportation tools as IV ([Birthal et al., 2015](#); [Ogutu & Qaim, 2019](#); [Zhang et al., 2021](#)). It is a concern for their validities since they may correlate to households' wealth or other omitted variables.

Kebede, 2022), constraints on capital (Blanco & Raurich, 2022), transportation cost (Rivera-Padilla, 2020), land ownership (Montero, 2022), climate shocks (Makate, Angelsen, Holden, & Westengen, 2022) and so on. Combining with the classic literature explaining cross-country agricultural productivity disparities, some recent empirical studies with structural models have begun to incorporate farmers' occupational choices (agricultural and nonagricultural) into structural models to further discuss the topic about misallocation (e.g. Adamopoulos, Brandt, Leight, & Restuccia, 2022; Lagakos & Waugh, 2013). The latest study extends the discussion to farmers' choices in agriculture (cash vs. grain crops) (Adamopoulos & Restuccia, 2020; Rivera-Padilla, 2020). For example, Adamopoulos and Restuccia (2020) find that restrictive land policies for large farms are the main reason for the low adoption rate of HVCs.

In terms of the impacts of the adoption of HVCs on farmers' performances, plentiful early empirical studies with qualitative and basic statistical analyses have revealed the positive relationships (e.g. Weinberger & Lumpkin, 2007). Recent systematic empirical studies with rural household survey data have found that cultivating HVCs for commercial purposes can improve income, profits, and reduce poverty in the developing world (Birthal et al., 2015; Masanjala, 2006; Ogutu & Qaim, 2019; Zhang, Mishra, & Hirsch, 2021). However, specializing in HVCs also means that smallholders, who are vulnerable, need to bear greater risks. Thus, diversifications in crops and income sources seem to be an effective way to balance profits and risks (Jin, Min, Huang, & Waibel, 2021; Michler & Josephson, 2017; Tesfaye & Tirivayi, 2020).

Moreover, the adoption of HVCs and its implication for smallholders' welfare are also very important and controversial topics in modern Chinese history. Traditional perspectives hold a negative attitude on cash crops since the markets were controlled by big imperialistic countries (e.g. Brandt, 1985). Huang (1985) considers that although cotton is more profitable, smallholder farmers can not cope with risk when they specialize in cotton planting. On the contrary, some other influential works approve that commercialization and market-oriented production stimulate growths in agriculture and income in pre-war China (Brandt, 1989; Myers, 1970; Rawski, 1989). Thus, systematic investigations are still needed.

Overall, this paper first contributes to the empirical literature on the impact of HVC adoption on farmers' productivity and welfare in both historical and contemporary dimensions. We incorporate this issue into a free-factor-markets context, which is also a crucial determinant of efficiency and equity (Otsuka, 2007). Compared with the existing literature, this study not only uses various methods, such as IV and DID to deal with endogeneity but also investigates the impact of HVC adoption on the whole rural economy from the perspective of farmers' factor allocation and inequality.

Second, this paper is also related to the literature on farmers' choices and agricultural productivity. While existing literature focusing only on whether farmers grow HVCs, this study further investigates farmers' proportion choices for different crops, and unlike Rivera-Padilla (2020), we also introduce the land tenancy market into the analysis.<sup>9</sup>

Finally, this paper has implications for contemporary rural land policy. Recently, a series of studies have documented the productivity losses caused by the factor misallocation resulting from land restrictive policies (e.g., Adamopoulos et al., 2022; Adamopoulos & Restuccia, 2014, 2020; Chari, Liu, Wang, & Wang, 2020; Chen, Restuccia, & Santaaulàlia-Llopis, 2022),<sup>10</sup> but there has been rarely any further discussion of the effect of such policies on inequality.<sup>11</sup> This paper is not a direct assessment of these restrictive policies but rather uses the historical context of free labor and land markets to examine their influence on inequality. Our results show that, at least in a specific context, land transfer liberalization can not only improve the efficiency of factor allocation but also reduce (and, at the very least, not exacerbate) inequality, which provides historical experience that supports the market-oriented reform of the land market in contemporary China.<sup>12</sup>

The remainder of this paper is organized as follows. Section 2 introduces the historical background of agricultural production and soybean trade in Manchuria. In Section 3, we model farmers' factor allocation and crop choice and analyze their impacts. Section 4 describes the data and variables used in this study. The corresponding empirical methods and results are given in Section 5. In Section 6, the influence of HVC cultivation on the inequality indices related to households' welfare is examined. Finally, Section 7 concludes.

## 2. Historical background

### 2.1. Development and agricultural management in modern Manchuria

As a newly opened region, Manchuria was different from other parts of China in terms of its agricultural production and management. Although records of early development date back to the Warring States period,<sup>13</sup> migration to Manchuria began around the

<sup>9</sup> In the model of the existing literature, farmers choose the most profitable occupation among three mutually exclusive options: nonagricultural sector, HCVs, and cereal crops (Adamopoulos & Restuccia, 2020; Rivera-Padilla, 2020). However, in reality, farmers often grow both cash and grain crops (e.g., Adamopoulos and Restuccia (2020), pp. 12, footnote 8).

<sup>10</sup> These restrictions often include caps on farm size, bans on land transfers and higher taxes on large farmers; see Online Appendix D of Adamopoulos and Restuccia (2014) for a detailed summary.

<sup>11</sup> An exception is Chen et al. (2022), who calibrated a structural model and carried out counterfactual inference to obtain the conclusion that the marketization of land factors can reduce inequality.

<sup>12</sup> Contemporarily, China is pursuing institutional reform for rural land marketization. A series of policies, such as the separation of ownership, contracting and management rights for rural land, has been implemented. Here, the historical experience refers to the gains from free factor markets only in terms of productivity, but not equity, since income gap between landlords and tenants is not an issue in contemporary China. The results may have implications of both efficiency and equity for other developing countries where tenancy system is still prevalent.

<sup>13</sup> During the Warring States period (475–221 BCE.), Manchuria belonged to the State of Yan.

1860s—when the Qing government removed the restrictions imposed on the Han Chinese against freely settlement there to mitigate its financial and frontier crises (Kong, 1986).<sup>14</sup> The Qing government opened up state-owned, banner (qidi) and waste land from south to north Manchuria sequentially for migrants to farm. The population of Manchuria grew from 3 million in 1850 to 40 million in 1940, while migrants from the North China Plain accounted for two-thirds of that growth (Eckstein, Chao, & Chang, 1974). Although modern industry grew rapidly until 1930, agriculture remained important, accounting for 36% of GDP (Ma, 2008), and >90% of the incomes of farmers came from crop production (Kung & Li, 2011).

In terms of land management mode, Manchuria shows the characteristics of agricultural production and management with high level of land concentration and developed land and labor markets. On the one hand, this situation was the result of the annexation of land by some large warlords, merchants and landlords due to the social turmoil in the late Qing Dynasty and the early Republic of China. On the other hand, this situation was also the result of the immigration timing in Manchuria; the first immigrating farmers often became landlords, and as a result, later immigrants often had to either sell their labor force as farm laborers or rent in land as tenants to make a living. Therefore, the above factors have led to a high concentration of land in Manchuria, and the Gini coefficient of land is between 0.7 and 0.9. Moreover, tenancy has become a very important method of production. According to the survey data of Manchuria Railway Corporation in Manchuria, 56.2% of family farms were involved in land rental, while up to 63.2% of family farms larger than 10 ha involved in land rent-in. In addition, approximately 50% of farm households are involved in rental labor in the labor market.<sup>15</sup>

## 2.2. Soybean trade in Manchuria in modern times

China has a long history of planting cash crops, such as tee, cotton, silk and so on. Until the late Qing dynasty, 15% of arable land had been devoted to cash crops (Shi, 2017). In the republican era, the share of cash crops in total agricultural production rose from 17% to 21% between the mid-1910s and mid-1930s (Perkins, 1970). Although China is the earliest soybean-planting country in the world, soybean planting in the northeast region has developed with the influx of immigrants and agricultural development in modern times. On the one hand, due to the characteristics of soybean itself, its roots have a high nitrogen content of rhizobia, which is convenient for agricultural development to change “raw land”, which is not easy to cultivate, into “ripe land”, which is easy to cultivate. On the other hand, the Vienna World Exposition in the early 20th century made Europe realize the advantages of soybeans in oil yield and that it could replace peanut as a chemical raw material. Therefore, the soybeans’ cultivation area expanded, which made Manchuria the largest soybean growing and production area in China and even the world at that time. In the early 20th century, soybean became the most commercialized crop in Manchuria. Such soybean trade played a very important role in foreign trade in the northeast region. Especially after WWI, the international market demand for soybeans further expanded, prompting soybeans to replace tea and silk and become the largest export commodity of modern China’s foreign trade (Kung & Li, 2011). At this time, the soybean export volume of Manchuria had reached approximately 60–70% of China’s total soybean export volume (Perkins, 1970). Moreover, soybean exports accounted for approximately 50–60% of the total exports of Manchuria, which shows the importance of soybean trade for economic development in Manchuria.

The rise of soybean trade in Manchuria in modern times has experienced several important stages in its development. In the late Qing Dynasty, the government used to strictly control soybean trade. Japan began to purchase a large number of soybeans and bean cakes in Manchuria after the Sino-Japanese War (1895), which promoted soybean production and trade in this region. However, the rapid development of Manchuria’s soybean trade occurred after the Russo-Japanese War. Around 1908, Japanese and Russian merchants began to ship Northeast Chinese soybeans to Europe and the United States. Since then, Northeast Chinese soybean trade has developed rapidly (Lei, 1981). During this period, soybean exports were also negatively affected by the impact of WWI on European demand. With the end of WWI, soybean exports from Manchuria enjoyed a golden age for most of the 1920s until the Great Depression era (Kung & Li, 2011). Fig. 1 depicts the time trend of soybean export volume and price in Manchuria in modern times. As mentioned in the above text, the solid-square price curve and the hollow-circle export volume curve both indicate that international soybean trade in Manchuria experienced a rising trend overall after 1895 and peaked in the 1920s, although there were fluctuations during this period.

## 3. Theoretical model

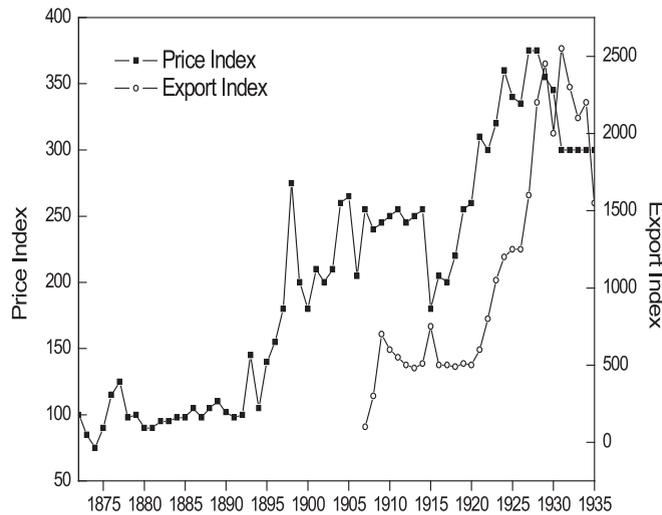
To better clarify the effects of HVC adoption on agricultural performance and welfare theoretically and derive hypotheses for empirical test, this section constructs a small open economy general equilibrium model by referring to the framework of Adamopoulos and Restuccia (2020). In this village economy, farmers are heterogeneous in their production ability and can choose to grow one or two types of crops: cash and grain crops. The factors needed for agricultural production (land and labor) can be transferred freely within markets but not across villages.

### 3.1. Setup

In each village economy, suppose that there is one landlord family and  $N_h$  farmer families. There are a total of  $L$  arable land and  $N$

<sup>14</sup> The crisis of the frontier in the Northeast deepened after the Second Opium War (circa 1860) because of the forced opening of Yingkou as a treaty port and the invasion of Russia, which occupied a territory of one million km<sup>2</sup> in North Manchuria.

<sup>15</sup> The introduction and description of this survey data are described in Section 4 of this paper.



**Fig. 1.** Chinese soybean price and export indices, 1872–1935.  
Source: citing from Fig. 2, Kung and Li (2011).

labor force in the village. All cultivated land is owned by the landlord family, while the labor force is evenly distributed across farmer families, with each labor force being able to provide 1 unit of labor inelastically. The landlord family does not produce but rents out the land, and production is carried out by family agricultural production. Farmer families are free to rent-in land and allocate each of their own labor (either to produce at home or to rent out as farm laborers) to grow crop  $s$ ,  $s \in \{g, c\}$  with  $g$  and  $c$  referring to grain and cash crops, respectively. According to Lucas (1978), the production function of farmer  $i$  regarding crop  $s$  is in the Cobb–Douglas form of diminishing returns to scale as follows:

$$Y_{is} = (A_s \kappa_i)^{1-\sigma} (N_{is}^{\alpha_s} L_{is}^{1-\alpha_s})^\sigma \tag{1}$$

where  $Y_{is}$  denotes the output and  $A_s$  denotes the productivity determined by the natural conditions of the village.  $N_{is}$  and  $L_{is}$  denote the labor and land input by farmers in crop cultivation, respectively.  $\kappa_i$  is the production ability (including operation, management, etc.) of farmer  $i$  for crop  $s$ , which is assumed to be determined by the capacity of the head of each farming household. The distribution of  $\kappa_i$  in one village is given exogenously, and such distribution is the same in each village. In addition,  $\alpha_s \in (0, 1)$  is a coefficient related to the output elasticity of the labor force, assuming that  $\alpha_c > \alpha_g$ , which means that cash crops are more labor intensive compared to grain crops.<sup>16</sup>  $\sigma \in (0, 1)$  is a span-of-control parameter.

Villagers can sell the two types of crops in markets after harvest. Since the village is a small open economy,  $P_s$  is determined by the external market.<sup>17</sup> Grain crops are used as the numeraire, so  $P_g = 1$ , and the ratio of the two prices is defined as  $P \equiv P_c/P_g$ ; thus, we can obtain the profit function of farmers growing a certain crop as follows:

$$\pi_{is} = P_{is} Y_{is} - w N_{is} - r L_{is} - C_s \tag{2}$$

where  $C_s$  is the fixed cost required to produce crop  $s$ , assuming that the fixed cost of growing cash crops is higher, that is, that  $C_c > C_g$ .  $w$  and  $r$  are wage and land rent in equilibrium, respectively, which are determined by the supply and demand in factor markets within village. The labor and land factor markets are clear when the market is in equilibrium, that is, when  $\sum_i \sum_s N_{is} = N$ ,  $\sum_i \sum_s L_{is} = L$ .

### 3.2. Model analysis

First, we analyze the optimal decision-making behavior of farmers' demand for production factors. According to Eq. (2), the profit function is maximized by choosing optimal factor inputs for each crop. Then, the maximized profit and optimal input demand can be calculated as follows:

<sup>16</sup> Since the labor input of farmers for each crop is not included in our data, we cannot directly estimate labor-elasticity parameter  $\alpha_s$ , but we can learn from other studies. For example, Chari et al. (2020) estimate the output elasticity of various crops with respect to different inputs by using nationwide survey data of farmers in contemporary China (see Table A.7 in their Appendix); among them, the output elasticity with respect to labor of soybeans (the most important cash crop in Manchuria) was 0.226, while corn and other grain crops are 0.148 and 0.198, respectively, and thus, our assumption should be reasonable.

<sup>17</sup> Another setting is to assume that grain crops are autarky in the village and that  $P_g$  could be endogenously determined by further assuming farmers' utility function. Since this approach does not change the main conclusions of the model,  $P_g$  is set here as exogenous given for simplicity.

$$N_{is}^* = A_s(\sigma P_s)^{\frac{1}{1-\sigma}} \left(\frac{1-\alpha_s}{r}\right)^{\frac{\sigma(1-\alpha_s)}{1-\sigma}} \left(\frac{\alpha_s}{w}\right)^{\frac{1-\sigma(1-\alpha_s)}{1-\sigma}} \kappa_i \tag{3}$$

$$L_{is}^* = A_s(\sigma P_s)^{\frac{1}{1-\sigma}} \left(\frac{1-\alpha_s}{r}\right)^{\frac{1-\sigma\alpha_s}{1-\sigma}} \left(\frac{\alpha_s}{w}\right)^{\frac{\sigma\alpha_s}{1-\sigma}} \kappa_i \tag{4}$$

$$\pi_{is}^* = A_s(1-\sigma)P_s^{\frac{1}{1-\sigma}}\sigma^{\frac{\sigma}{1-\sigma}} \left(\frac{1-\alpha_s}{r}\right)^{\frac{\sigma(1-\alpha_s)}{1-\sigma}} \left(\frac{\alpha_s}{w}\right)^{\frac{\sigma\alpha_s}{1-\sigma}} \kappa_i - C_s \tag{5}$$

3.2.1. Analysis of production ability

It can be seen from Eqs. (3)–(5) that since other parameters are the same, for farmers in the same village, the demand for labor and land inputs and the maximum profit that can be obtained by a farmer are proportional to its production ability  $\kappa_i$ , regardless of what kind of crop is grown, that is,  $\partial N_{is}^*/\partial \kappa_i > 0$ . It is clear that in frictionless factor markets, farmers with higher production ability operate larger farms, employ more labor, and obtain higher profits. At this time, the total output is also the largest because there is no misallocation of production factors.

Next, we analyze farmers’ choice on crops. Since there is a fixed cost for growing both grain and cash crops, the head of a household does not grow a crop until the profit is positive. Similar to the conclusion of Adamopoulos and Restuccia (2020), since profit is proportional to ability and  $C_c > C_g$ , there are two critical values of production ability,  $\underline{\kappa}_i$  and  $\bar{\kappa}_i$ , where  $\pi_{ig}^*(\underline{\kappa}_i) = 0$  and  $\pi_{ic}^*(\bar{\kappa}_i) = 0$ . The meaning here is that when a farmer’s ability is lower than  $\underline{\kappa}_i$ , he or she chooses to be a farm laborer for another family’s farm since the profits of the two crops are negative. According to a similar analysis, when  $\underline{\kappa}_i \leq \kappa_i \leq \bar{\kappa}_i$ , farmers choose to grow grain crops. When  $\kappa_i > \bar{\kappa}_i$ , both types of crops are grown. Combining the results of the above analysis, we can obtain Result 1 as follows:

**Result 1:** In a village, when a farmer’s production ability,  $\kappa_i$ , is higher, (1) he or she is more likely to grow cash crops; (2) the greater the demand for labor and land is; (3) higher the profit can be obtained.

3.2.2. Analysis of suitability for cash crops

In addition to farmers’ production ability, natural environmental factors (coefficient  $A_s$ ), such as the soil and climate in each village, also play a crucial role in crop production. Since soybean was the most important cash crop in the “Manchuria” at that time, whether the local area was suitable for growing soybeans had an impact on farmers’ crop choice and profits. Although the change in  $\pi_{is}^*$  cannot be obtained directly from Eq. (5) because  $w$  and  $r$  change simultaneously when  $A_s$  changes, it can still be obtained simply from Eqs. (1) and (2) according to the Envelope Theorem as follows:

$$\partial \pi_{ic}^*/\partial A_c = \partial \pi_{ic}/\partial A_c = (1-\sigma)A_c^{-\sigma} P \kappa_i^{1-\sigma} (N_{ic}^{\alpha_c} L_{ic}^{1-\alpha_c})^{\sigma} > 0 \tag{6}$$

We know that  $\partial^2 \pi_{ic}^*/\partial A_c^2 < 0$  after calculating the second derivative and derive  $\partial(\partial \pi_{ic}^*/\partial A_c)/\partial \pi_{ic}^* < 0$ , which means that the improvement of  $A_c$  can increase profits, and the effect of which is more significant for farmers with low profits.

For farmers’ crop choice, because the increase in productivity  $A_c$  of cash crops causes an increase in profit  $\pi_{ic}^*$ , it can be seen from the above crop type selection analysis that critical value  $\bar{\kappa}_i$  decreases at this time, that means more farmers will grow cash crops. For farmers who have already grown both types of crops, the proportion of land devoted to cash crops  $\mu$  is as follows:

$$\mu = AP^{\frac{1}{1-\sigma}} \left(\frac{w}{r}\right)^{\frac{\sigma(\alpha_g-\alpha_c)}{1-\sigma}} (1-\alpha_c)^{\frac{1-\sigma\alpha_c}{1-\sigma}} (1-\alpha_g)^{\frac{1-\sigma\alpha_g}{1-\sigma}} \tag{7}$$

where  $A \equiv A_c/A_g$ , and  $\partial \mu/\partial A > 0$  can be proven (see the Appendix for the proof). Combining the above conclusions, we can derive the following result:

**Result 2:** When the relative productivity of cash crops to grain crops  $A$  is higher, (1) the proportion of land devoted to cash crops in the village increases, and (2) the profits of growing cash crops in the village increases, and this increasing effect becomes more significant for farmers with low profits.

In addition to its impact on household income, the impact of the cultivation of cash crops on income inequality within and between villages has also become an attractive issue for economists and policy makers. From the above analysis, it can be seen that farmers with higher production ability use more inputs and gain more profits through the free factor markets. Does this situation increase the degree of inequality? Farmers’ income depends mainly on their factor endowments in free markets, and we thus analyze the impact of productivity  $A_c$  on endogenous factor prices  $w$  and  $r$  within the village. Note that for simplicity, we look only at the case where all farmers have the ability to grow cash crops.<sup>18</sup> When the market is in equilibrium, we can derive  $\sum_i N_{ig}^* + \sum_i N_{ic}^* = N$  and  $\sum_i L_{ig}^* + \sum_i L_{ic}^* = L$  according to the conditions of labor and land market clearing. After substituting the expressions of  $N_{is}^*$  and  $L_{is}^*$  ( $s = g, c$ ) in Eqs. (3) and (4) into the above equations, we can derive

<sup>18</sup> This conclusion can be extended to the situation in which when  $A_c$  changes, the types of crops grown do not change for all farmers. As this work includes only a suggestive analysis, we do not further discuss the complex situation in which some farmers’ change the types of crops they grow theoretically.

$$\frac{\left(\frac{W}{r}\right)^{\frac{\sigma(1-\alpha_g)}{1-\sigma}} Z_1 + A \left(\frac{W}{r}\right)^{\frac{\sigma(1-\alpha_c)}{1-\sigma}} Z_2}{\left(\frac{W}{r}\right)^{\frac{1-\sigma\alpha_c}{1-\sigma}} Z_3 + A \left(\frac{W}{r}\right)^{\frac{1-\sigma\alpha_c}{1-\sigma}} Z_4} = \frac{N}{L} \quad (8)$$

where  $Z_i$  ( $i=1, 2, 3, 4$ ) is a positive constant related only to  $P$ ,  $\alpha_g$ ,  $\alpha_c$  and  $\sigma$ .  $\partial(w/r)/\partial A > 0$  can be proven (see the Appendix for the proof). Intuitively, since the wage share of cash crops is higher ( $\alpha_c > \alpha_g$ ), the ratio of wage-to-land-rent increases as the proportion of cash crops increases. In the model setting of this study, the landlord does not produce but only leases the land, so an increase in  $w/r$  means that the income share of land endowment declines, which leads to a reduction in the income gap between farmers and landlords. We summarize as the following result:

**Result 3:** When productivity  $A_c$  of cash crops is higher, the ratio of wages to land rent in villages increases, which reduces the income gap between farmers and landlords.

### 3.3. Hypotheses

To transform the results of the above model into hypotheses that can be empirically tested, one must note the following two points. First, due to data availability, we cannot estimate farmers' total factor productivity (TFP) for various crops, such as was done by Chari et al. (2020), but we can use farm size as a proxy for  $\kappa_i$  while controlling for village characteristics since farmers' input demands are proportional to their production ability. Second, although the influence of the wage-to-land-rent ratio on inequality indicators (such as the Gini index) in the village depends on the distribution of the initial factor endowment, we focus mainly on the income difference between landlords and farmers in the historical background of this paper, which is directly related to the wage-to-land-rent ratio. Based on the above results, we can obtain the following hypotheses:

**Hypothesis 1.** A farmer operating a larger farm area has a higher proportion of cash crops grown and earns more profit.

**Hypothesis 2.** Farmers in a village that is more suitable for growing soybeans has a higher proportion of cash crops grown and greater agricultural profits.

**Hypothesis 3.** The positive effect of the cultivation of cash crops on profits is more significant for farmers with low profit.

**Hypothesis 4.** The ratio of wage to land rent is greater in villages where the proportion of land devoted to cash crops is larger, which reduces the inequality index of farmers' welfare.

In the theoretical analysis of this study, the exogenous variables are farmers' ability  $\kappa_i$ , the relative productivity  $A$  and relative price  $P$  between cash and grain crops, and the endogenous variables are crop choice, input demands and factor prices.<sup>19</sup> The specific empirical test methods and results are detailed in the next sections of this paper.

## 4. Data

### 4.1. Introduction

In this study, data from a unique household-level rural survey conducted in the 1930s are used to examine determinants of crop choice and its impacts on productivity, welfare and inequality. The survey was conducted by the Investigation Department of the South Manchurian Railways Co. under the auspices of the Provisional Industrial Investigation Bureau, which belonged to the Ministry of Enterprises of the State Council of "Manchukuo" in the mid-1930s. The target of this survey was to provide information for the Manchurian Agricultural Development Five-Year Plan, which aimed to increase agricultural output. The survey was implemented in two waves. The first wave was conducted in 19 villages chosen from 16 counties in North Manchuria in late February 1935, while the second took place 1 year later in 26 villages chosen from 25 counties that were mainly located in Middle and South Manchuria. Altogether, the two surveys covered information on family members, migration histories, properties, production, consumption, loans and so on for 1911 rural households in 45 villages located in 41 counties (Fig. 2).

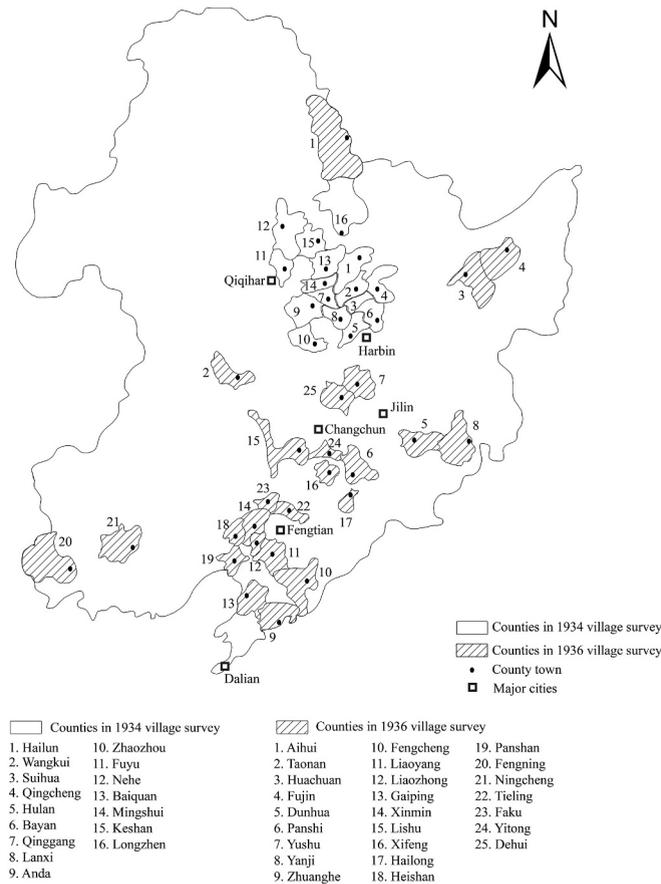
Although the surveyed villages were not selected randomly, their wide spatial dispersion makes the surveys geographically representative.<sup>20</sup> Additionally, considering its regularity and comprehensiveness, the quality of this survey is among the best in the Republican era. Thus, these survey data have been widely used by scholars (e.g., Myers, 1976; Benjamin & Brandt, 1997; Kung & Li, 2011; Li & Li, 2017).

### 4.2. Descriptive statistics

According to the data of this rural survey, we first demonstrate the planting structure of the "Manchukuo" in the 1930s in Appendix Fig. A.1. The proportion of land grown soybeans is approximately 20% in both North and South Manchuria, and the

<sup>19</sup> As for the influence of relative price  $P$  on related variables, it can be inferred from the position of  $P$  in the model that the impacts of  $P$  and  $A$  should be similar. We carry out a relevant empirical investigation Section 5.2.2 (B).

<sup>20</sup> Considering for the safety of the investigators, the surveyed villages were mainly located near county towns.



**Fig. 2.** Location of villages in the 1934–1936 Manchurian village survey. (Source: Location of observation is from Manchuria Village Surveys in the 1930s)

proportions of corn and millet are also not regionally different. The difference between North and South Manchuria is reflected mainly in the fact that North Manchuria grows more wheat, while South Manchuria grows more sorghum. Then, we show the geographical distribution of HVCs and productivity. In Appendix Fig. A.2(a), which shows the proportion of HVCs grown in each village, the North Manchuria area has a higher proportion and is the main region of soybean and wheat production in Manchuria during the study period. The South Manchuria area, similar to the large cities of Mukden and Dalian, is dominated by the cultivation of food crops. In addition, the relationship between the HVC adoption and productivity is depicted in Appendix Fig. A.2(b). In terms of the output per unit of labor force, its regional distribution exhibits a similar pattern with HVCs, revealing a possible positive correlation between the two variables.

The statistical information of the variables used in the empirical work of this study is shown in Table 1. The empirical investigations are conducted mainly from two observation levels: the household and village levels. At the micro level (Panel A of Table 1), we investigate the determinants of crop choice, and the impact of HVC adoption on farmers' per capita farm profits, per capita expenditure on living, per capita net cash income and other related welfare indicators.<sup>21</sup> At the same time, some household characteristic variables (including the amount of land and labor force owned and used by households, household population, age of household head, socioeconomic status of household, and the proportion of nonagricultural labor force) and village-level characteristic variables (including the ratio of cash to grain crops price, per capita cultivated land, distance to the county seat, distance to big cities, average land quality, etc.) are controlled for.<sup>22</sup> At the village level (Panel B of Table 1), we investigate the influence of HVC cultivation on inequality

<sup>21</sup> The sample used in this study includes a total of 1911 rural households. Since we investigate mainly the influence of crop choices on relevant welfare indicators, we delete the samples of pure landlords and pure hired laborers that do not engage in farming and obtain 1198 observations. Among them, the sample size is reduced to 1170 because the data of household living expenditure is missing for some households. For the net cash income, this information is collected only in the second survey in 1936, so there are only 836 observations.

<sup>22</sup> According to the classification of Myers (1976), we divide the socioeconomic status of farmers into four types: landlord = 4, owner cultivator = 3, tenant = 2, and laborer = 1. In addition, since the data do not contain information on farmers' land quality, but there is the quality information of each plot of rented land (top, medium land quality = 1, and inferior land quality = 0), we take the average quality of all rented land in the village as the control variable.

**Table 1**  
Descriptive statistics for the variables used in the regression analysis.

Variables	Description	Obs.	Mean	Sd	Max	Min
Panel A: household level						
HVCs prop.	Proportion of HVCs (%)	1198	0.21	0.234	1	0
HVCs grow	Whether HVCs are grown (yes = 1)	1198	0.59	0.49	1	0
Profit	Per capita farm profit (Yuan)	1198	20.48	39.299	780.64	-108.60
Expenditure	Per capita expenditure on living (Yuan)	1170	10.99	11.804	125.97	0.02
Income	Per capita net cash income (Yuan)	836	2.73	49.752	1394.95	-82.46
Land_ope.	Land area sown (ha)	1198	10.94	22.328	296	0.1
Labor	Labor used	1198	3.35	3.100	29	0
Land_own	Land owned (ha)	1198	10.48	33.956	610.21	0
Family	Family size	1198	7.87	5.886	68	1
Age	Household head age	1198	49.62	11.956	61	18
Status	Socioeconomic Status (1–4)	1198	2.21	0.729	4	1
Land_ren	Land rented-in (ha)	1198	5.13	11.93	115	0
Off-farm	Ratio of off-farm labor to population (%)	1198	0.03	0.101	1	0
Panel B: village level						
HVCs prop.	Proportion of HVCs (%)	45	0.40	0.158	0.78	0.06
Gini_pro.	Gini index of per capita farm profit	45	0.62	0.140	0.881	0.36
Gini_exp.	Gini index of per capita expenditure	45	0.37	0.082	0.67	0.252
Gini_inc.	Gini index of per capita income	26	0.62	0.109	0.915	0.427
Gini_land	Gini index of per capita land owned	45	0.76	0.110	0.965	0.449
Land_rent	Ratio of land for rental (%)	45	0.48	0.250	0.959	0.009
Gini_lab.	Gini index of per capita labor owned	45	0.32	0.061	0.495	0.222
Hired_day	Days being hired as laborer	45	0.19	0.132	0.718	0.011
Off-farm	Ratio of off-farm labor to population (%)	45	0.08	0.095	0.403	0
Yield_ratio	Cash to grain crops potential yield ratio	45	0.58	0.146	1	0.28
Price_ratio	Cash to grain crops price ratio	45	1.46	0.413	3.21	1
Wage_land	Wage to land ratio	45	5.53	6.133	31.62	0.180
To_county*	Distance to county seat (Huali)	45	30.36	26.300	167	8
To_city*	Distance to city (Huali)	45	107.7	73.557	414	15
Land_ara*	Per capita arable land (ha)	45	1.15	0.769	3.24	0.15
Size*	Village Size	45	42.60	18.100	91	9
Land_qua*	Land quality on average	45	0.70	0.209	1	0
Age_vil*	Village age	45	106	89.606	300	5
Railway*	Railway passing by (yes = 1)	45	0.62	0.490	1	0

Note: (1) The calculation for Gini indices of per capita expenditure on living and net cash income include all (non- negative value) households in village. (2) Labor used equals family members plus hired laborers. (3) Socioeconomic status: landlord = 4 , owner-cultivator = 3 , tenant = 2 , laborer = 1. (4) Yield ratio is the ratio of the potential yield between the most grown cash and food crops in each village, as well as the price ratio. (5) One Huali = 0.311 mile. (6) Yuan is the unit of local currency in Manchuria. (7) The village-level variables with \* are included in all regressions with village controls.

indicators such as the Gini index of the above welfare indicators.

## 5. Empirical results

### 5.1. Determinants of crop choice

According to the above analysis, as an endogenous variable, farmers' crop choices are affected by many external factors. Therefore, before investigating the influence of HVC adoption on other indicators, we first conduct a systematic empirical investigation of the determinants of farmers' crop choice to reveal their correlations. The specific empirical equation is as follows:

$$y_{vi} = \alpha + \beta_1 land\_ope_{vi} + \beta_2 yield\_ratio_v + \beta_3 price\_ratio_v + \gamma' z_{vi} + \varepsilon_{vi} \quad (9)$$

where  $y_{vi}$  denotes the crop choice of farmer  $i$  living in village  $v$ , including the proportion of land grown HVCs, whether growing HVCs and the count of crops.  $land\_ope_{vi}$  is the area of farmers' operating land, and  $yield\_ratio_v$  and  $price\_ratio_v$  are the ratios of potential yield and price between the cash crop with the largest planted area and the grain crop in the village, respectively.<sup>23</sup>  $z_{vi}$  is a vector of a series of household- and village-level control variables, as specifically described in Table 1.  $\alpha$ ,  $\beta_i$ , and  $\gamma'$  are the corresponding coefficients to be estimated, and  $\varepsilon_i$  is the random error term.

The corresponding empirical results are reported in Table 2. In column (1), after controlling for individual characteristics and village fixed effects, farmers with larger operating land areas are found to grow a greater proportion of HVCs. Numerically, a 1%

<sup>23</sup> For the details of the potential yield ratio between cash and food crops, please refer to Section 5.2.2 (A) of this paper.

increase in the area of operating land leads to an 8.4% increase in the proportion of land grown HVCs. After adding the potential yield and price ratios and replacing village fixed effects with village-level control variables and provincial fixed effects in columns (2), the results are still robust, which is consistent with [Hypothesis 1](#).<sup>24</sup> In addition, the results of columns (2) also show that when the relative productivity of HVCs is higher, the proportion of HVCs is also higher, thus confirming the first part of [Hypothesis 2](#). However, the coefficient of the price ratio is significantly negative, which seems to be opposite to expectations. The possible reason for this is that crop prices are not exogenous. In regions where more HVCs are grown, prices of food crops rise since they are produced less, thus exhibiting a negative correlation.

An interesting issue is that how do farmers manage the risk induced by cultivating HVCs. Thanks to the richness of the survey data, we uncover two facts regarding the income-risk tradeoff. First, devoting at least a portion of land to food crops to reduce exposure to weather or price shocks. In fact, only 10 of 1198 households devoted to cash crops in entirety. Second, the varied use of share contracts as a risk-sharing device. In column (3) of [Table 2](#), tenant subsample regression is performed using a tenant subsample. The results indicate that a farmer who signs a share contract grows more HVCs, but this effect weakens as the frequency of exceptional disasters increases.<sup>25</sup>

In columns (4) to (5) of [Table 2](#), the dependent variables are changed to whether farmers grow HVCs and the count of crop types grown, respectively. As a result, both the coefficients of operating land area are still significantly positive. The result in column (5) suggests that the diversity of farmers' crop choice may be the consequence of the optimization behaviors by competent farmers rather than the forced risk-coping behaviors by small farmers.<sup>26</sup> In addition, since the data in the survey do not include information on the labor allocation for each crop, we cannot estimate the TFP of farmers for each crop. Instead, we exploit a small subsample of farmers who grow only grain crops for investigation.<sup>27</sup> After using the Cobb–Douglas production function to estimate the output elasticity with respect to land and labor, we find that the elasticities are 0.795 and 0.085, respectively, and that the sum of the two is <1, which is consistent with the assumption that the production function in this model is diminishing returns to scale. After calculating TFP through the residual term, we put it into the regression as an explanatory variable, the results of which are reported in column (6) of [Table 2](#). The coefficient of TFP is significantly positive, indicating that farmers with higher ability operate larger farm areas, which provides some evidence for Result 1.

## 5.2. Effects of HVC adoption on agricultural performance and welfare: Micro level

### 5.2.1. Baseline regression

To systematically and empirically investigate the impact of HVC adoption on the indicators of farmers' agricultural performance and welfare, we exploit OLS model for estimation, the specific equation of which is set as follows:

$$y_{vi} = \alpha + \beta_1 HVCs\ prop_{vi} + \gamma' z_{vi} + \varepsilon_{vi} \quad (10)$$

where  $y_{vi}$  denotes agricultural performance indicators such as agricultural profit and factor productivity of household  $i$  living in village  $v$ , as well as welfare indicators such as per capita household net cash income and expenditure on living. The key explanatory variable  $HVCs\ prop_{vi}$  is the proportion of land grown HVCs by farmer  $i$ .<sup>28</sup>  $z_{vi}$  is a vector of a series of control variables at the household and village levels.  $\alpha$ ,  $\beta_1$ , and  $\gamma'$  are the coefficients to be estimated, and  $\varepsilon_{vi}$  is the random error term. Since prices vary from region to region and have the most important impact on villagers' welfare, we divide the per capita farm profit, per capita expenditure on living and per capita net cash income by local food prices in the regression; that is, we convert the nominal value into the real value. In addition, for the value added of output, we replace the local price with the average price of all villages and then multiply it by the corresponding crop yield to deal with the effect of the regional heterogeneity of prices.

We first investigate the agricultural performance in the baseline regression, and the regression results are reported in [Table 3](#). In

<sup>24</sup> Notice that we can not exclude the effect of economic scale when using farm size as a proxy of ability to predict crop choice, since the limitation of the data in estimating TFP. See the interpretation of the result of Column (6) in [Table 2](#) for more discussion.

<sup>25</sup> This result may be explained as that, share contracts were dominating in the locations where exceptional disaster is frequent, so contract choice did not affect crop choice significantly. The data of historical disasters in each village are from *Yearly Charts of Dryness/Wetness in China for the Last 500-Year Period*.

<sup>26</sup> In the development economics literature, there are two opposing views as to the reasons for farmers' planting diversity (or the diversity of income sources): one is the initiative of capable farmers (unconstrained by funds and other constraints), and the other is the behavior of small farmers forced to cope with risks. For details, please refer to [Martin and Lorenzen \(2016\)](#).

<sup>27</sup> Since maize and sorghum have similar unit area yields, farmers with the planting proportion of the sum of these two crops >80% are used as a subsample. In this subsample, the information of both land and labor inputs are available for TFP estimate.

<sup>28</sup> Among crops, wheat is another important cash crop in North Manchuria, and its volume for sale accounts for 50.7% of total production. Rice is only grown on a large scale in two villages in Yanji County. In addition, the most important crops (and their unit prices) are soybean (4.02), wheat (5.62), sorghum (3.12), millet (3.30) and corn (2.92) in dan/yuan. Information on the highest proportion of cash and food crops planted in each village can be found in Appendix [Table A.1](#).

**Table 2**  
Determinants of crop choice.

Dependent variable	HVCs prop.			HVCs grow	Count of crops	Land_ope. (ln)
	(1)	(2)	(3)			
Land_ope (ln)	0.083*** (0.009)	0.065*** (0.012)	0.073*** (0.028)	0.205*** (0.025)	1.677*** (0.138)	
Yield_ratio		0.403*** (0.064)	0.530*** (0.114)	0.646*** (0.145)	-2.461*** (0.603)	
Price_ratio		-0.045** (0.020)	-0.035 (0.028)	-0.033 (0.037)	0.094 (0.185)	
Contract_type (share = 1)			0.201** (0.078)			
Disaster			0.131 (0.250)			
Contract_type *Disaster			-0.700** (0.320)			
TFP						0.031*** (0.008)
Hou. Cons.	Y	Y	Y	Y	Y	Y
Village FE	Y	N	N	N	N	Y
Vil. Cons.	N	Y	Y	Y	Y	N
Province FE	N	Y	Y	Y	Y	N
Obs.	1198	1198	630	1198	1198	252
R-squared	0.555	0.432	0.460	0.429	0.667	0.706

Note: (1) The models are OLS. (2) Robust standard errors in parentheses. (3) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

columns (1) to (3) the coefficients of the proportion of HVCs are significantly positive after gradually controlling for household characteristics and village fixed effects, which verifies [Hypothesis 2](#).<sup>29</sup> Since the survey does not contain information about the quality of farmers' cultivated land, we therefore use a subsample of tenants in which the quality of the renting land is available in tenancy contracts to alleviate this concern. The results are reported in column (4) of [Table 3](#), and the previous conclusions remain robust after adding land quality. In column (5), we change village fixed effects to a combination of village control variables and provincial fixed effects, and the result does not change. In addition, we use the unit output values as dependent variables in columns (6) and (7), the results of which show that when the proportion of HVCs increases by 10%, the per capita and per hectare output value increase by 8.4% and 7.1%, respectively, revealing that the adoption of HVCs can effectively improve agricultural productivity. The coefficients of cultivated area are found to be both significant and nonsignificant, indicating that there is no negative relationship between productivity and farm size.

Next, to rule out that the empirical results are driven only by some specific variables and to further investigate the welfare index, we replace some key explanatory variables and dependent variables in Eq. (10) and then conduct similar analysis. The results are reported in [Table 4](#).

We use whether to grow HVCs as independent variable in columns (1) to (4) of [Table 4](#). The results in column (1) show that the coefficient is not significant at the 10% level but that the corresponding coefficient is significant at the 1% level when the analysis is limited to tenant subsamples containing land quality information in column (2). The results in columns (3) to (4) show that the factor productivity of farmers is significantly improved by growing HVCs. In columns (5) to (6), we replace the dependent variables with the per capita expenditure on living and per capita net cash income to investigate the impact of HVC adoption on welfare indicators.<sup>30</sup> The results show that when controlling for village-level variables and provincial fixed effects, the former is positive and significant, while the latter is positive and nonsignificant. When strictly controlling for village fixed effects, neither is significant (the results are not reported). This lack of significance may be due to the following reasons. First, in addition to the income from the sale of agricultural products, the variables of expenditure and income also include the land rent income of the landlord and the wage income of some family members employed by other farms, which are not included in the income of their own farm operations. Second, the estimated results (including those in [Table 3](#)) may be affected by endogeneity, which is further investigated in the next subsection.

### 5.2.2. Robustness tests

Although the results of baseline regression reveal the positive effect of the adoption of HVCs on agricultural performance and are consistent with the predictions of the theoretical model, the estimation results may be biased due to a series of problems such as endogeneity. Therefore, this subsection investigates the robustness of the results in terms of the below aspects.

<sup>29</sup> Per capita farm profit is calculated as (total value of farm output of the household - total cost)/(household population \* unit price of food). Total cost includes the wages of employees (the sum of annual, monthly and daily wages) plus land rent. When taking the logarithm of per capita agricultural profit, since some of the values are negative, we first shift all the values as a whole to ensure the minimum equal one before taking the logarithm, so as to not lose the observations.

<sup>30</sup> These two variables are divided by the unit price of food crop and translated into logarithms, just like the per capita farm profit.

**Table 3**  
The effects of HVC adoption on agricultural performance.

Dependent variable	Profit (ln)					Output per labor (ln)	Output per ha (ln)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HVCs prop	0.160*** (0.027)	0.167*** (0.026)	0.200*** (0.033)	0.259*** (0.044)	0.215*** (0.035)	0.845*** (0.166)	0.710*** (0.148)
Land_ope (ln)	0.144*** (0.010)	0.224*** (0.024)	0.230*** (0.026)	0.273*** (0.056)	0.230*** (0.026)	1.204*** (0.080)	-0.031 (0.066)
Land quality				0.018 (0.017)			
Hou. Cons.	N	Y	Y	Y	Y	Y	Y
Vil. Cons.	N	N	N	N	Y	N	N
Village FE	N	N	Y	Y	N	Y	Y
Province FE	N	N	N	N	Y	N	N
Obs	1198	1198	1198	630	1198	1198	1198
R-squared	0.317	0.404	0.548	0.526	0.455	0.675	0.545

Note: (1) The models are OLS. (2) Robust standard errors in parentheses. (3) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

**Table 4**  
The effects of HVC adoption on agricultural performance and welfare.

Dependent variable	Profit (ln)		Output per labor (ln)	Output per ha (ln)	Expenditure (ln)	Income (ln)
	(1)	(2)	(3)	(4)	(5)	(6)
HVCs grow (yes = 1)	0.015 (0.014)	0.053*** (0.019)	0.231*** (0.061)	0.115** (0.056)		
HVCs prop					0.242*** (0.081)	0.061 (0.051)
Land quality		0.018 (0.017)				
Hou. cons	Y	Y	Y	Y	N	N
Village FE	Y	Y	Y	Y	N	N
Vil. cons	N	N	N	N	Y	Y
Province FE	N	N	N	N	Y	Y
Obs.	1198	630	1198	1198	1170	836
R-squared	0.536	0.509	0.668	0.533	0.461	0.084

Note: (1) The models are OLS. (2) Robust standard errors in parentheses. (3) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

5.2.2.1. *IV regression.* According to the above analysis, the previous regression results are threatened by the endogeneity problem because some important variables, such as farmers’ production ability, are omitted in our regression, and there is likely to be a two-way causal relationship between crop choice and agricultural performance. To solve these issues, we exploit the potential yield ratio of cash to food crops in each village as the IV of the proportion of land grown HVCs. The information of the IV are derived from the GAEZ database of the FAO of the United Nations (FAO GAEZ database, version 4). The database calculates the maximum yield that is theoretically possible for each crop in each grid based on local natural conditions such as soil and climate. These data have also been used as IV in some important studies due to their good exogeneity (e.g., Bustos et al., 2016). According to geographical coordinates, we obtain the maximum potential yield of the cash and food crops with the largest planted area (Appendix Table A.1) in the grid of each surveyed village and then calculate their ratio as the IV, so the variable exactly corresponds to productivity ratio A in Section 3. We conduct regression analysis using 2SLS, and the corresponding model is set as follows:

$$\text{First stage : } HVCs \widehat{prop}_{vi} = \alpha_1 + \lambda \cdot yield\_ratio_v + \gamma' z_{vi} + \mu_{vi} \tag{11}$$

$$\text{Second stage : } y_{vi} = \alpha_2 + \beta \cdot HVCs \widehat{prop}_{vi} + \gamma' z_{vi} + \varepsilon_{vi} \tag{12}$$

Eq. (11) is the first-stage regression, the dependent variable  $HVC \widehat{prop}_{vi}$  is the farmers’ proportion of land devoted to HVCs, and the IV  $yield\_ratio_v$  is the potential yield ratio. Eq. (12) is the second-stage regression, in which the dependent variable is the related variable of agricultural performance and welfare, and the explanatory variable is the estimated value of  $HVC \widehat{prop}_{vi}$  in Eq. (11). The results of 2SLS with IVs are reported in Table 5.

As shown in panel A of Table 5, in the first-stage regression, the coefficients of the potential yield ratio are significantly positive, indicating that the proportion of HVCs grown is higher as a consequence of the higher relative productivity of soybeans to food crop, and an F value >10 indicates that the IV passes the test of weak IV. Panel B of Table 6 reports the regression results of the second stage, showing that the cultivation of HVCs has a significant positive effect on the per capita farm profit of households, the per capita expenditure for living and the per capita net income. The results remain consistent when we change the explanatory variable from the proportion of HVCs to whether to grow HVCs (Table A.2 in Appendix). It is worth noting that for the welfare indicators, the coefficients of the HVCs proportion become more significant in 2SLS estimates with IV, revealing that the baseline OLS estimation results are likely

to exhibit endogenous bias.

**5.2.2.2. Cohort DID regression.** In the previous text, we use 2SLS regression with relative potential yield as the IV to deal with the endogeneity problem. However, as discussed in Section 3, the price ratio  $P$  between cash and food crops is also a potential mechanism similar to productivity ratio  $A$ . Although the price is included in the data, the results of Subsection 5.1 show that food price is determined by the crops' proportion in the village, so it cannot be directly treated as an exogenous variable. The survey data are cross-sectional data, although, we can still use the fact that the international soybean price changed over time combined with the timing when farmers migrate to the village to do some investigations by using the DID model.

Fig. 1 shows that the soybean price and export indices in Manchuria rose mainly after the Russo-Japanese War in 1908 and peaked in the 1920s after the negative demand impact of WWI. Therefore, farmers who migrated to the regions that produced mainly soybeans during this peak period should have benefited more than those who arrived at other times. Referring to Kung and Li (2011), the DID estimation equation is set as follows:

$$y_{vit} = \alpha + \sum_{period2}^{period6} \delta_t mig_{it} + \beta_0 vil_{vi} + \sum_{period2}^{period6} \beta_t (mig_{it} * vil_{vi}) + \gamma' z_{vi} + \varepsilon_{vit} \quad (13)$$

where  $y_{vit}$  is the agricultural profit of farmer  $i$  who immigrated to village  $v$  during period  $t$  and  $mig_{it}$  and  $vil_{vi}$  are dummy variables for the migration period and whether the region produces mainly soybeans, respectively.<sup>31</sup> The corresponding regression results are reported in Table 6, and coefficient  $\beta_t$  is the result of DID estimation of interest.

The results in columns (1) to (3) of Table 6 show that when the per capita farm profit, per capita profit of labor and profit of unit area are, respectively used as the dependent variables, only the interaction terms between immigrants in the 1920s and major soybean-producing region are significantly positive, which indicates that the profit difference between farmers who emigrate to regions that mainly producing soybeans and those who emigrate to other regions is larger in the period of high soybean prices. Although the dependent variables are only cross-sectional data in the 1930s, farmers may have accumulated more wealth by growing soybeans before, thus enabling them to grow HVCs without financial constraints (the fixed cost item in Eq. (5)) during the survey period, thus increasing their profits. In addition, although the coefficients of interaction terms during WWI (1914–1918) are all negative and significant only in column (2), they still reflect the negative impact of WWI on soybean trade. The results of DID estimation here are similar to those of Kung and Li (2011), both of which confirm the role of soybean trade in improving the welfare of farmers in Manchuria.

**5.2.2.3. Subsample regression.** To exclude that the results of baseline regressions are driven by some specific subsamples (such as regions) and to further investigate the heterogeneity, we divide the samples into subsamples on the basis of North and South Manchuria, the median farm size and whether the region is a major soybean-producing region, and conduct regression analysis. The results are reported in Table 7. As shown in Table 7, two main pieces of information can be obtained. First, the coefficients of the proportion of HVCs are all positively significant at the 1% level, indicating that the improving effect of HVC adoption on agricultural performance is general in each subsample. Second, the results show that the effect is more significant in those subsamples where HVC cultivation is more common (North Manchuria region, farm sizes larger than the median, and in the major soybean production regions). The above results further confirm the robustness of the baseline regression results.

**5.2.2.4. Other discussions.** The above robustness checks basically support the positive effect of HVC cultivation on agricultural performance and farmers' welfare found in the baseline regression, but there are still some concerns that may lead to bias in the results that are worth discussing. The first issue is the influence of outliers. To eliminate the interference of some relatively large or small outliers in the regression results, we delete the top and bottom 5% observations of per capita farm profit, per capita expenditure on living and per capita net cash income and then use the remaining observations for the same regressions. The results in Appendix Table A.3 show that our findings are not affected.

The second problem is the self-selection of immigrants. In all the discussions above, the unbiasedness of the estimators depends on the assumption of random farmer immigration at the village level, that is, it was not the case where incoming migrant farmers with higher production ability migrated to soybean-producing regions. Fortunately, this assumption is reasonable after considering the history of immigrant in Manchuria. As described in Section 2, the Northeast gradually opened up in the order from south to north during the late Qing Dynasty and the Republic of China, so the destinations of the first generation of immigrants were strictly restricted by authorities. For later immigrants, their destinations were often where they had relatives or friends and, thus, not mainly determined by economic motivation.<sup>32</sup>

Finally, the third concern is the rotation of crops. After growing soybeans in the first year, growing sorghum in the second year with

<sup>31</sup> The division of immigration period is shown in Table 6. Among the 45 villages, the average soybean planting area is 21.7%, and villages with a soybean planting proportion larger than this average are defined as major soybean producing regions in this study, with reference to the practice of Kung and Li (2011).

<sup>32</sup> According to the survey data used in this paper, more than one-third of farmers migrated to a certain place due to the recommendation of local relatives or friends. Regarding the choice of destination, Kung and Li (2011) discussed it in detail in the last paragraph of page 577 and the first paragraph of page 578.

**Table 5**  
The effects of HVCs' proportion on agricultural performance and welfare with IV.

Panel A: First stage			
Dependent variable	HVCs prop		
	(1)	(2)	(3)
Yield_ratio	0.403*** (0.064)	0.409*** (0.064)	0.313*** (0.074)
Controls	Y	Y	Y
Obs	1198	1170	836
F-value	39.517	40.598	18.113
R-squared	0.432	0.436	0.545

Panel A: Second stage			
Dependent variable	Profit (ln)	Expenditure (ln)	Income (ln)
	(1)	(2)	(3)
HVCs prop	0.376** (0.202)	4.515*** (0.805)	0.467** (0.223)
controls	Y	Y	Y
Obs	1198	1170	836

Note: (1) The models are 2SLS. (2) The controls include household and village control variables and province dummy. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

**Table 6**  
The effects of HVC adoption on profit with DID.

Dependent variable	Profit (ln)	Profit per labor (ln)	Profit per ha (ln)
	(1)	(2)	(3)
Soybean producing region*immigrant period			
1860–1894	0.072 (0.044)	0.098** (0.043)	0.076** (0.033)
1895–1907	0.185** (0.091)	0.085 (0.071)	0.091** (0.044)
1908–1915	0.098 (0.067)	0.120 (0.089)	0.127*** (0.043)
1916–1920	−0.071 (0.060)	−0.167* (0.101)	−0.022 (0.087)
Post 1921	0.099** (0.050)	0.113*** (0.043)	0.125*** (0.030)
Controls	Y	Y	Y
Obs	1198	1198	1198
R-squared	0.472	0.518	0.224

Note: (1) The models are OLS. (2) The controls include household and village control variables and province dummy. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

**Table 7**  
The effects of HVCs' proportion in subsamples.

Dependent variable	North	South	Farm Size (>median)	Farm Size (<median)	Soybean region (No)	Soybean region (Yes)
	(1)	(2)	(3)	(4)	(5)	(6)
Profit (ln)						
HVCs prop	0.290*** (0.054)	0.088** (0.035)	0.086*** (0.023)	0.380*** (0.076)	0.104** (0.042)	0.206*** (0.044)
Hou. cons	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y
Obs	438	760	600	598	737	461
R-squared	0.677	0.370	0.508	0.587	0.452	0.674

Note: (1) The models are OLS. (2) Robust standard errors in parentheses. (3) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

nitrogen fixed by soybeans can achieve a higher yield. Therefore, farmers often choose the mode of rotation between soybeans and food crops, which leads to the unstable crop choice observed in the survey year. In fact, our results are not heavily influenced by the crop rotation problem. Although only one year's harvest is included in the survey data, information on the proportion of each crop over

two or three years is recorded. In columns (3) and (6) of [Table A.1](#) in the Appendix, we show the change in the proportion of land devoted to the major cash and food crops in each village in the surveyed year and the year adjacent, respectively, the results of which show that on average, they change only 4.1% and 3.3%, respectively. The reasons for this can refer to the introduction of local crop rotation in the survey report of Tieling County. First, as soybeans are the main cash crop in the local area, some farmers continue to grow them until the yield drops significantly and then begin to implement crop rotation. Second, farmers usually rotate crops within different plots instead of growing just one crop each year, so the proportion of each crop may not change much.

## 6. Extension: Impacts of HVC adoption on inequality

The previous section systematically and empirically examines the influence of HVC adoption on farmers' agricultural performance and welfare, in the context of free factor markets. The results show that larger farms derive more benefits because they tend to grow larger proportion of HVCs, raising concerns about worsening wealth inequality among farmers operating different farm sizes in villages. In fact, this is the main reason for many developing countries currently implementing various types of restrictive land policies. In this section, we use quantile regression and the data of 45 villages to conduct empirical investigations on inequality at the village level. Compared with many contemporary agricultural surveys, although the historical survey used in this study is not representative enough—it does not randomly select some farmers in the village and rather investigates all farmers—so we can better examine the influence of HVC adoption on index of inequality in the village.

### 6.1. Quantile regression

Since the heterogeneity of the effects of HVCs growing on agricultural profit, income and expenditure in different groups is also a very noteworthy issue, the quantile regression method is used in this section for relevant investigation. Different from [Section 5.2.2 \(C\)](#), which examines the heterogeneity of explanatory variables through subsample OLS regression, in this section, quantile regression is exploited to examine the heterogeneity of the dependent variables. The specific model is set as follows:

$$y_{vi} = \alpha + \beta_{\tau} HVCs\ prop_{vi} + \gamma' z_{vi} + \varepsilon_{vi} \quad (14)$$

where  $\tau$  is a quantile between 0 and 1. After the regression of Eq. (14) for a series of  $\tau$  values, we are interested in the difference in the estimated coefficients  $\beta_{\tau}$  obtained in the heterogeneity analysis.

The regression results are reported in Appendix [Tables A.4–A.6](#), and the corresponding estimation coefficients are shown in [Fig. 3](#). [Fig. 3\(a\)](#) shows that the coefficient seems to have a gradually decreasing trend, although it is not obvious, indicating that the influence of the proportion of HVCs grown on per capita farm profit seems to be greater for farmers with smaller per capita profit compared to those with larger per capita profit, which provides some evidence in support of [Hypothesis 3](#). There is almost no heterogeneity in the impact of the proportion of HVCs grown on per capita expenditure on living and per capita net cash income, as shown in [Figs. 3\(b\) and 3\(c\)](#), and the coefficients for the latter are basically not significant. Overall, the quantile regression results show that the cultivation of HVCs does not lead to larger gains for wealth farmers, thereby alleviating our concerns about worsening inequality.

### 6.2. Investigation of inequality at the village level

To provide more examination on inequality at different levels, we use the Gini and Theil indices as the dependent variables in the regression analysis at the village level. The corresponding equation of the OLS model is set as follows:

$$y_v = \alpha + \beta \bullet HVCs\ prop_v + \gamma' z_v + \varepsilon_v \quad (15)$$

where  $y_v$  is the Gini or Theil index of per capita farm profit, per capita expenditure on living and per capita net cash income of village  $v$ . It should be noted that we use all households in the village to calculate the Gini or Theil index of per capita expenditure on living and per capita net cash income to measure inequality within villages more comprehensively and accurately.<sup>33</sup> The key explanatory variable  $HVCs\ prop_v$  is the proportion of land grown HVCs in the village.  $z_v$  is a vector of village control variables, and  $\varepsilon_v$  is a random error term.

The results are reported in [Table 8](#). Columns (1) to (3) show that there is a negative relationship between the proportion of land grown HVCs and the Gini index of per capita farm profit, per capita expenditure on living and per capita net cash income when no village-level control variables are controlled for. However, such effect on the Gini index of per capita farm profit is not significant at the 10% level after controlling for village-level control variables in columns (4) to (6), while the effects on the latter two Gini indices change little. The above results indicate that those villages with higher proportion in HVC cultivation have less inequality in some welfare indicators. To further rule out the endogenous estimation bias, similar to the micro-level investigation, we still exploit the potential yield ratio of cash to food crops as the IV of the proportion of HVCs and use 2SLS model. The results of which are given in Appendix [Table A.7](#) which show that the coefficients in column (1) are positive but not significant, while the results in the last two columns are basically the same as those in [Table 8](#). In fact, according to the results of the microlevel investigation, it is reasonable to say that the adoption of HVCs worsens the inequality of agricultural profits because farmers with higher proportion of HVCs grown are

<sup>33</sup> This means that the calculation includes not only the samples of farmers who operating land (own or rent-in), which are used in the previous microlevel investigation, but also those samples of pure landowners and labors who do not operate farms.

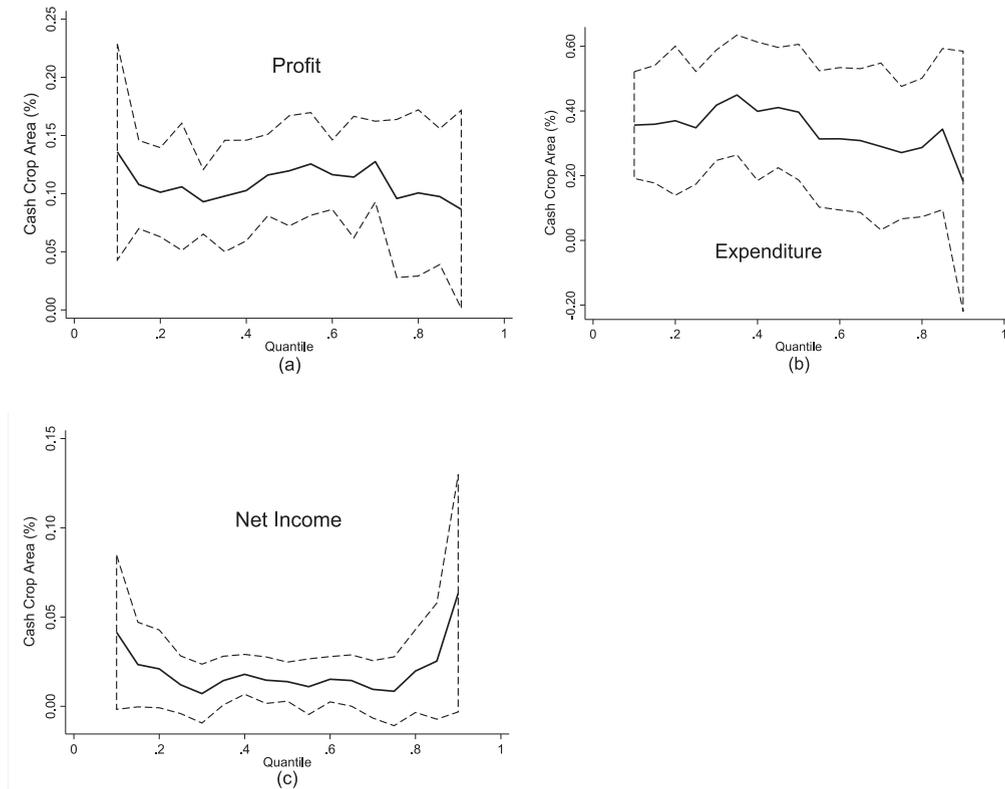


Fig. 3. Coefficients of the quantile regressions for the effects of HVCs' proportion on (a) profit, (b) expenditure and (c) income.

more profitable.

Finally, to test [Hypothesis 4](#) in [Section 3](#), we add the ratio of village wage to land rent as a mediator to the regression model in [Table 8](#), the results of which are reported in [Table 9](#).<sup>34</sup> Column (1) of [Table 9](#) show that the ratio of wage to land rent is greater in villages where the proportion of HVCs is higher. In columns (2) to (4) of [Table 9](#), we add the ratio of wage to land rent as an explanatory variable into the regression, the results of which show that compared with the results in columns (4) to (6) of [Table 8](#), the negative impact of the proportion of HVCs on the Gini indices of per capita farm profits and per capita net cash income is weakened, while the impact on the Gini coefficient of per capita expenditure on living is small. The most obvious effect is the result in column (4) of [Table 9](#). The coefficient of the proportion of HVCs grown (which is significantly negative) changes from negatively significant (column (6) of [Table 8](#)) to insignificant after controlling for the wage to land-rent ratio (its own coefficient is significantly negative), indicating that the cultivation of HVCs reduces the Gini index through the mechanism of raising the wage to land-rent ratio.

It is a concern that the variable “per capita expenditure on living” may not be a good proxy for welfare, since HVCs growers had to purchase grain from the market whereas food growers did not. To address this issue, we subtract the expenditure on food and calculate the corresponding Gini index. The results in column (5) show that the coefficient of the proportion of HVCs grown is again significantly negative and becomes less significant after controlling for the wage-to-land-rent ratio in column (6). These results reveal that the cultivation of HVCs in villages narrows the gap in endowment income between farmer households and landowners by increasing the wage-to-land-rent ratio, thus reducing the inequality of welfare indicators such as expenditure, which provides some empirical evidence for [Hypothesis 4](#). Overall, our conclusions lack some robustness due to the limitation on sample size, but the present evidence reveal that, the cultivation of HVCs has at least not worsened inequality, thus providing some support for market-oriented reform policies that facilitate the mobility of agricultural production factors.<sup>35</sup>

## 7. Conclusions

A number of developing countries have always had a large gap in agricultural labor productivity compared to developed countries,

<sup>34</sup> Salary data are derived from the average salary of laborers in the survey year working in the village. The land rent data are the average annual rent of the land of median quality in the village.

<sup>35</sup> The conclusion of this paper is similar to that of [Benjamin and Brandt \(1997\)](#) in terms of the influence of the free factor markets on farmers' income inequality at that time, whereas we further investigate the effect of HVC adoption.

**Table 8**  
The effects of HVCs' proportion on inequality.

Dependent variable	Gini_pro	Gini_exp	Gini_inc	Gini_pro	Gini_exp	Gini_inc
	(1)	(2)	(3)	(4)	(5)	(6)
HVCs prop	-0.288** (0.124)	-0.248** (0.101)	-0.209* (0.111)	-0.142 (0.128)	-0.212** (0.085)	-0.270* (0.128)
Gini_land				0.350 (0.285)	0.224 (0.188)	-0.128 (0.604)
Land_rent				0.008 (0.100)	-0.087 (0.077)	0.054 (0.183)
Gini_lab				0.163 (0.283)	0.215 (0.163)	-0.503 (0.551)
Hired_day				0.393** (0.168)	-0.180 (0.137)	-0.053 (0.319)
Price_ratio				-0.007 (0.038)	0.017 (0.039)	-0.000 (0.116)
Other controls	N	N	N	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y
Obs	45	45	26	45	45	26
R-squared	0.323	0.299	0.411	0.755	0.545	0.683

Note: (1) The models are OLS. (2) The other controls include village control variables in Table 1. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

**Table 9**  
The mechanism between HVCs' proportion and inequality.

Dependent variable	Ratio of wage to land rent (ln)	Gini_pro	Gini_exp	Gini_inc	Gini_exp (exclude grain)	Gini_exp (exclude grain)
	(1)	(2)	(3)	(4)	(5)	(6)
HVCs prop	1.706* (1.001)	-0.121 (0.135)	-0.227** (0.092)	0.065 (0.212)	-0.224** (0.097)	-0.213* (0.106)
Ratio of wage to land rent (ln)		-0.012 (0.025)	0.009 (0.020)	-0.099* (0.046)		-0.010 (0.024)
Vil. cons	Y	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y
Obs	45	45	45	26	45	45
R-squared	0.628	0.757	0.548	0.781	0.622	0.624

Note: (1) The models are OLS. (2) The controls include village control variables in Table 1. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

so how to improve agricultural productivity is crucial for addressing the development issues, such as poverty reduction, of these countries. Taking advantage of trade opportunities in domestic and international markets, the cultivation of HVCs is considered an effective way to increase farmers' productivity and welfare. In addition, although the liberalization of factor markets is considered a way to improve efficiency, it is also controversial due to concerns about inequality. However, systematic empirical examinations of the impact of HVC cultivation on agricultural performance and welfare, especially inequality, is still limited in the existing literature. To complement this issue, this study first constructs a theoretical model of farmers' factor allocation and crop choice under free factor markets and then exploits household-level agricultural survey data in Manchuria in the 1930s, combined with the special historical background of the free factor markets and the prosperity of international soybean trade at that time, to conduct a systematic empirical investigation of the corresponding impacts.

The results show that at the micro level, first, farmers with higher production ability operate larger farms and are more likely to grow HVCs. Second, the adoption of HVCs can significantly improve performance indicators, such as the profit and single-factor productivity of farmers, and welfare indicators, such as expenditure and net cash income. These results remain robust after implementing a series of robustness tests including IV regression. For example, numerically, the results of using IV to address endogeneity show that when the proportion of land devoted to HVCs increases by 10%, the per capita farm profit, per capita expenditure on living and per capita net cash income increase by 4.3%, 45.3% and 4.9%, respectively. Third, the quantile regression shows that the effect of HVC adoption on agricultural profits is more significant for small farmers with lower profits. In addition, at the village level, the increase in the proportion of HVCs cultivated reduces the Gini indices of welfare indicators such as expenditure and net cash income by raising the ratio of wage to land rent within villages.

This paper provides not only systematic theoretical analysis and empirical evidence for the impact of HVC cultivation on agricultural performance, welfare and inequality, which is a very important topic both historically and contemporary, but also supports historical experience for contemporary land reforms. It is well known that due mainly to concerns about inequality, many developing countries have implemented a variety of land and taxation policies that restrict land transfers or that are detrimental to large farmers. However, the results of this study show that at least for the cultivation of labor-intensive HVCs, the market-oriented transfer of land can

not only achieve efficient allocation but also increase the proportion of labor income in total output, thus helping narrow the income gap between smallholder farmers and landlords.

**Data availability**

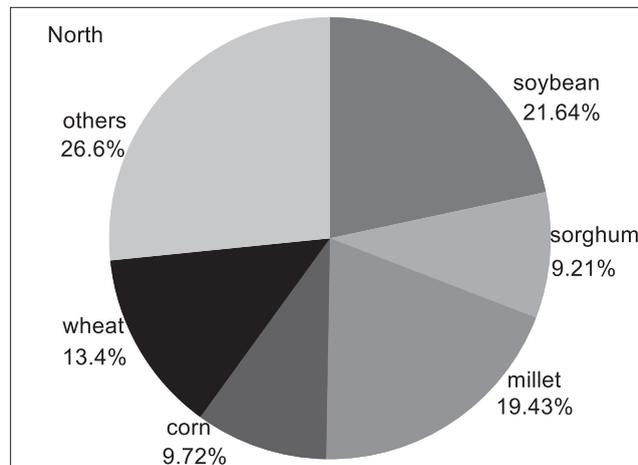
Data will be made available on request.

**Acknowledgements**

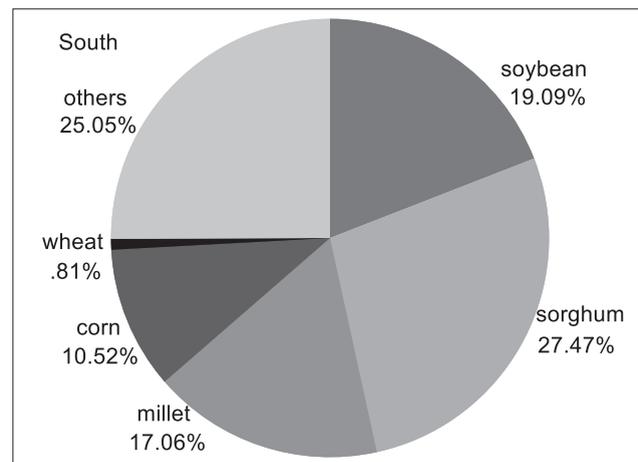
We are grateful to the two anonymous referees and the editor for valuable and kindly comments, and the financial support of the National Ethnic Affairs Commission of the People’s Republic of China (2021-GMB-011). All remaining errors are our own.

**Appendix A**

*A.1. Part A: Figures and tables*



(a)



(b)

**Fig. A.1.** Crops grown in North and South Manchuria.  
Source: the 1934–1936 Manchurian rural survey.

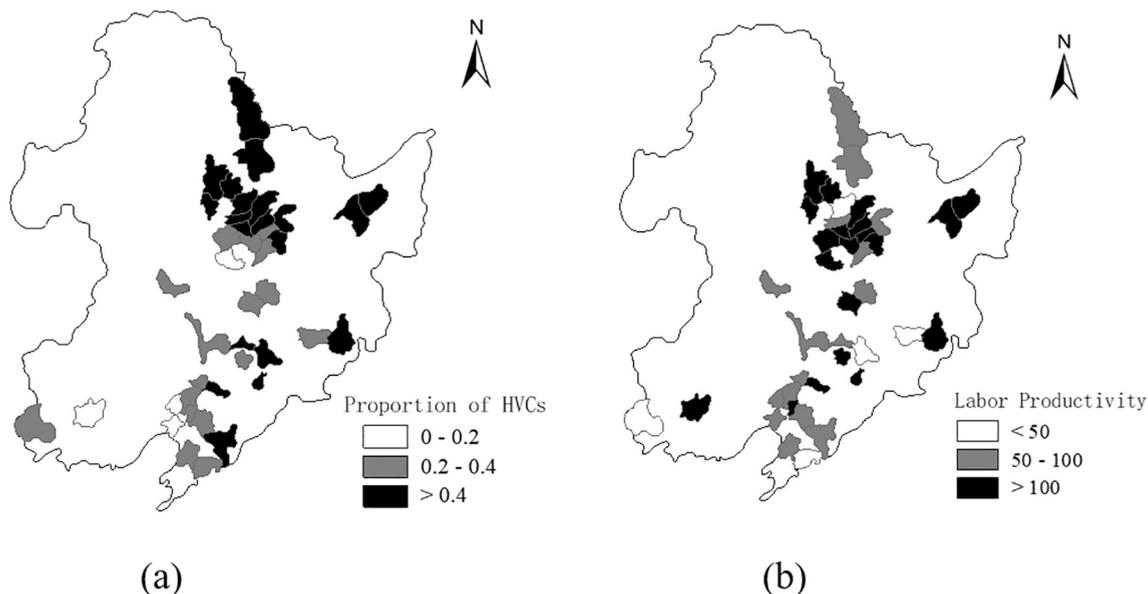


Fig. A.2. Spatial distribution of HVCs and labor productivity. Source: the 1934–1936 Manchurian rural survey.

Table A.1

Summary of crops in each surveyed village.

County name	(1) Area proportion of cash crop with largest area	(2) Sale proportion of cash crop with largest area	(3) Area proportion change of cash crop with largest area	(4) Area proportion of food crop with largest area	(5) Sale proportion of food crop with largest area	(6) Area proportion change of food crop with largest area
Hailun	32	80		16 <sup>a</sup>	0	
Wangkui	27	85		11 <sup>a</sup>	0	
Suihua	18	90	0	16 <sup>a</sup>	15	0
Qingcheng	30	65	15	19 <sup>a</sup>	1	9
Hulan	22	63	2	26 <sup>a</sup>	15	1
Bayan	33	87	6	15 <sup>b</sup>	11	1
Qinggang	19 <sup>A</sup>	82	8	17 <sup>a</sup>	23	9
Lanxi	8	92	1	24 <sup>a</sup>	19	3
Anda	13 <sup>A</sup>	88	11	21 <sup>b</sup>	32	9
Zhaozhou	6	81	5	20 <sup>b</sup>	23	1
Fuyu (1)	25	27	7	21 <sup>a</sup>	5	1
Fuyu (2)	33	77	8	22 <sup>a</sup>	3	5
Nehe	41	64	1	22 <sup>a</sup>	4	0
Baiquan	15	54	12	14 <sup>c</sup>	12	9
Mingshui	21 <sup>A</sup>	82	9	25 <sup>a</sup>	2	5
Keshan (1)	37	72	11	20 <sup>a</sup>	43	0
Keshan (2)	31	65	7	20 <sup>a</sup>	0	1
Keshan (3)	28	20	9	25 <sup>a</sup>	3	1
Longzhen	26	64	5	23 <sup>a</sup>	0	4
Aihui	16 <sup>B</sup>	46	1	18 <sup>a</sup>	0	1
Taonan	25	64	4	28 <sup>b</sup>	35	1
Huachuan	30	43	2	17 <sup>a</sup>	1	2
Fujin	16	30	1	14 <sup>b</sup>	20	0
Dunhua	18	25	4	23 <sup>c</sup>	11	4
Panshi	39	41	0	33 <sup>c</sup>	0	3
Yushu	33	82	0	22 <sup>a</sup>	4	4
Yanji (1)	40 <sup>C</sup>	29	5	40 <sup>d</sup>	29	5
Yanji (2)	69 <sup>C</sup>	14	6	69 <sup>d</sup>	14	6
Zhuanghe	12	4	0	60 <sup>c</sup>	4	0
Fengcheng	12	48	0	52 <sup>c</sup>	4	0
Liaoyang	15	53	2	53 <sup>b</sup>	19	5
Liaozhong	17	99	6	52 <sup>b</sup>	41	4
Gaiping	14	65	1	54 <sup>b</sup>	9	2
Xinmin	8	66	2	40 <sup>b</sup>	17	0
Lishu	27	57	1	33 <sup>b</sup>	6	6

(continued on next page)

Table A.1 (continued)

County name	(1) Area proportion of cash crop with largest area	(2) Sale proportion of cash crop with largest area	(3) Area proportion change of cash crop with largest area	(4) Area proportion of food crop with largest area	(5) Sale proportion of food crop with largest area	(6) Area proportion change of food crop with largest area
Xifeng	21	54	2	38 <sup>b</sup>	32	1
Hailong	37	71	2	24 <sup>b</sup>	18	3
Heishan	4	17	2	64 <sup>b</sup>	7	3
Panshan	6	34	2	69 <sup>b</sup>	24	3
Fengning	17	1	1	51 <sup>a</sup>	1	3
Ningcheng	6	78	2	44 <sup>b</sup>	11	1
Tieling	33	86	5	27 <sup>b</sup>	6	2
Faku	25	85	1	47 <sup>b</sup>	4	10
Yitong	37	60	5	18 <sup>a</sup>	3	15
Dehui	23	68	2	26 <sup>b</sup>	21	1
Average	23.0	59.0	4.1	30.9	12.0	3.3

Note: (1) Source: the 1934–1936 Manchurian rural survey. (2) Crops: column (1): A: perilla-seed, B: wheat, C: rice, others unmentioned are soybean; column (4): a: millet, b: sorghum, c: corn, d: rice. All figures are in %.

Table A.2

The effects of HVCs grown on agricultural performance and welfare with IV.

Panel A: First stage			
Dependent variable	HVCs grow (yes = 1)		
	(1)	(2)	(3)
Yield_ratio	0.635*** (0.145)	0.653*** (0.145)	0.544*** (0.164)
Controls	Y	Y	Y
Obs	1198	1170	836
F-value	19.241	20.291	11.043
R-squared	0.425	0.411	0.459

Panel A: Second stage			
Dependent variable	Profit (ln)	Expenditure (ln)	Income (ln)
	(1)	(2)	(3)
HVCs grow (yes = 1)	0.272** (0.136)	2.825*** (0.671)	0.267* (0.150)
controls	Y	Y	Y
Obs	1198	1170	836

Note: (1) The models are 2SLS. (2) The controls include household and village control variables and province dummy. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Table A.3

The effects of HVCs' proportion on agricultural performance and welfare (drop top and bottom 5%).

Dependent variable	Profit (ln)		Expenditure (ln)		Income (ln)	
	(1)	(2)	(3)	(4)	(5)	(6)
HVCs prop	0.095*** (0.019)	0.513*** (0.134)	0.259*** (0.072)	3.300*** (0.668)	0.059** (0.028)	0.304 (0.191)
Controls	Y	Y	Y	Y	Y	Y
Obs	1079	1079	1053	1053	749	749
R-squared	0.583	0.337	0.407		0.070	

Note: (1) The models are OLS in columns (1), (3), (5) and 2SLS in columns (2), (4), (6). (2) The controls include household and village control variables and province dummy. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Table A.4

The effects of HVCs' proportion on profit with quantile regression.

Dependent variable	Profit (ln)				
	0.1	0.3	0.5	0.7	0.9
Quantile	(1)	(2)	(3)	(4)	(5)
HVCs prop	0.102***	0.096***	0.098***	0.120***	0.090***

(continued on next page)

**Table A.4** (continued)

Dependent variable	Profit (ln)				
	0.1	0.3	0.5	0.7	0.9
Quantile	(1)	(2)	(3)	(4)	(5)
Controls	(0.018) Y	(0.014) Y	(0.017) Y	(0.013) Y	(0.023) Y
Obs	1198	1198	1198	1198	1198
Pseudo-R <sup>2</sup>	0.215	0.303	0.392	0.471	0.591

Note: (1) The controls include household and village control variables and province dummy. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\*p < 0.01 , \*\*p < 0.05 , \*p < 0.1.

**Table A.5**

The effects of HVCs' proportion on expenditure with quantile regression.

Dependent variable	Expenditure (ln)				
	0.1	0.3	0.5	0.7	0.9
Quantile	(1)	(2)	(3)	(4)	(5)
HVCs prop	0.358*** (0.068)	0.379*** (0.094)	0.345*** (0.095)	0.292*** (0.083)	0.190 (0.141)
Controls	Y	Y	Y	Y	Y
Obs	1170	1170	1170	1170	1170
Pseudo-R <sup>2</sup>	0.234	0.273	0.282	0.275	0.279

Note: (1) The controls include household and village control variables and province dummy. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\*p < 0.01 , \*\*p < 0.05 , \*p < 0.1.

**Table A.6**

The effects of HVCs' proportion on income with quantile regression.

Dependent variable	Income (ln)				
	0.1	0.3	0.5	0.7	0.9
Quantile	(1)	(2)	(3)	(4)	(5)
HVCs prop	0.007 (0.032)	0.002 (0.011)	0.009 (0.006)	0.021** (0.009)	0.068* (0.039)
Controls	Y	Y	Y	Y	Y
Obs	836	836	836	836	836
Pseudo-R <sup>2</sup>	0.131	0.030	0.013	0.034	0.101

Note: (1) The controls include household and village control variables and province dummy. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\*p < 0.01 , \*\*p < 0.05 , \*p < 0.1.

**Table A.7**

The effects of HVCs' proportion on inequality with IV.

Dependent variable	Gini_pro	Gini_exp	Gini_inc
	(1)	(2)	(3)
HVCs prop	0.117 (0.189)	-0.406** (0.167)	-0.443*** (0.139)
Controls	Y	Y	Y
Obs	45	45	26
R-squared	0.713	0.475	0.654

Note: (1) The controls include village control variables and province dummy. (3) Robust standard errors in parentheses. (4) Significance level: \*\*\*p < 0.01 , \*\*p < 0.05 , \*p < 0.1.

**A.2. Part B: Mathematical proof**

Here we firstly prove that  $\partial(w/r)/\partial A > 0$  , then sequentially prove  $\partial\mu/\partial A > 0$ . First of all, the constants  $Z_i$  ( $i=1, 2, 3, 4$ ) in Eq. (8) of the main text are,

$$\begin{aligned}
 Z_1 &= \left(\frac{1-\alpha_g}{\alpha_g}\right)^{\frac{\sigma(1-\alpha_g)}{1-\sigma}} \alpha_g^{\frac{1}{1-\sigma}}, Z_2 = \left(\frac{1-\alpha_c}{\alpha_c}\right)^{\frac{\sigma(1-\alpha_c)}{1-\sigma}} (P\alpha_c)^{\frac{1}{1-\sigma}} \\
 Z_3 &= \left(\frac{1-\alpha_g}{\alpha_g}\right)^{\frac{1-\sigma\alpha_g}{1-\sigma}} \alpha_g^{\frac{1}{1-\sigma}}, Z_4 = \left(\frac{1-\alpha_c}{\alpha_c}\right)^{\frac{1-\sigma\alpha_c}{1-\sigma}} (P\alpha_c)^{\frac{1}{1-\sigma}}
 \end{aligned}
 \tag{A1}$$

Since all the parameters are positive, we have  $Z_i > 0$ . For convenience, define  $a \equiv \frac{\sigma(1-\alpha_g)}{1-\sigma}$ ,  $b \equiv \frac{\sigma(1-\alpha_c)}{1-\sigma}$ ,  $c \equiv \frac{\sigma(\alpha_g-\alpha_c)}{1-\sigma} = b - a$ ,  $\frac{w}{r} = x$ , where  $a$ ,  $b$  and  $c$  are constants. Since we have assumed  $\alpha_g < \alpha_c$ , so we have  $c < 0$ . Therefore, Eq. (8) can be written as:

$$\frac{Z_1x^a + AZ_2x^b}{Z_3x^{a+1} + AZ_4x^{b+1}} = \frac{N}{L} \tag{A2}$$

Dividing both sides of the above equation by  $x^a$ , one can derive

$$Z_1 + AZ_2x^c = Z_3(N/L)x + AZ_4(N/L)x^{c+1} \tag{A3}$$

According to the law of implicit function derivation, computing the partial derivative with respect to A, one can derive

$$\frac{\partial x}{\partial A} = \frac{x^c[Z_4(N/L)x - Z_2]}{cAx^{c-1}[Z_2 - Z_4(N/L)x] - Z_3(N/L) - AZ_4(N/L)x^c} \tag{A4}$$

Since  $Z_4 = Z_2\left(\frac{\alpha_c}{1-\alpha_c}\right)$ , thus the numerator equals  $x^cZ_4\left(\frac{Nw}{rL} - \frac{\alpha_c}{1-\alpha_c}\right)$ . According to the property of Cobb-Douglas function,  $Nw$  and  $rL$  denote wage and land rent income, respectively. Since  $\alpha_c > \alpha_g$ , so the fraction of wage income in total output is higher when HVCs are grown. In the model we assume that farmers grow both cash and grain crops simultaneously, so we have  $\frac{\alpha_g}{1-\alpha_g} < \frac{Nw}{rL} < \frac{\alpha_c}{1-\alpha_c}$ . Hence, the numerator in Eq. (A4) is negative, as well as the first term in denominator ( $c < 0$ ). In addition, since other parameters are all positive, so the last two terms in denominator are both smaller than zero. Overall, we derive that,  $\partial x/\partial A > 0$ , that is,  $\partial(w/r)/\partial A > 0$ .

Next, we prove that  $\partial\mu/\partial A > 0$ . According to Eq. (7) in main text, we have

$$\mu = AP^{1-\varepsilon}x^c(1-\alpha_c)^{b+1}(1-\alpha_g)^{a+1} \tag{A5}$$

Taking the partial derivative with respect to A, one can derive

$$\frac{\partial\mu}{\partial A} = dx^{c-1}\left(x + Ac\frac{\partial x}{\partial A}\right) \tag{A6}$$

where  $d \equiv P^{1-\varepsilon}(1-\alpha_c)^{b+1}(1-\alpha_g)^{a+1} > 0$ . Since the first term in parenthesis is positive whereas the second term is negative, we need further analysis. Substituting  $\partial x/\partial A$  into Eq. (A7), one can derive after simplification,

$$x + Ac\frac{\partial x}{\partial A} = \frac{-Z_3(N/L)x - AZ_4(N/L)x^{c+1}}{cAx^{c-1}[Z_2 - Z_4(N/L)x] - Z_3(N/L) - AZ_4(N/L)x^c} \tag{A7}$$

Basing on the analysis above, we know that both numerator and denominator in Eq. (A7) are negative, thus,  $\partial\mu/\partial A > 0$  is proven. Q.E.D.

## References

- Adamopoulos, T., Brandt, L., Leight, J., & Restuccia, D. (2022). Misallocation, selection, and productivity: A quantitative analysis with panel data from China. *Econometrica*, 90(3), 1261–1282.
- Adamopoulos, T., & Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6), 1667–1697.
- Adamopoulos, T., & Restuccia, D. (2020). Land reform and productivity: A quantitative analysis with Micro data. *American Economic Journal: Macroeconomics*, 12(3), 1–39.
- Benjamin, D., & Brandt, L. (1997). Land, factor markets, and inequality in rural China: Historical evidence. *Explorations in Economic History*, 34(4), 460–494.
- Birthal, P. S., Hazrana, J., & Negi, D. S. (2020). Diversification in Indian agriculture towards high value crops: Multilevel determinants and policy implications. *Land Use Policy*, 91, Article 104427.
- Birthal, P. S., Roy, D., & Negi, D. S. (2015). Assessing the impact of crop diversification on farm poverty in India. *World Development*, 72, 70–92.
- Blanco, C., & Raurich, X. (2022). Agricultural composition and labor productivity. *Journal of Development Economics*, 158, Article 102934.
- Brandt, L. (1985). Chinese agriculture and the international economy, 1870-1930s: A reassessment. *Explorations in Economic History*, 22(2), 168–193.
- Brandt, L. (1987). Farm household behavior, factor markets, and the distributive consequences of commercialization in early twentieth-century China. *The Journal of Economic History*, 47(3), 711–737.
- Brandt, L. (1989). *Commercialization and agricultural development in east-Central China, 1870–1937*. Cambridge University Press.
- Bustos, P., Caprettini, B., & Ponticelli, J. (2016). Agricultural productivity and structural transformation: Evidence from Brazil. *American Economic Review*, 106(6), 1320–1365.
- Chari, A., Liu, E. M., Wang, S.-Y., & Wang, Y. (2020). Property rights, land misallocation, and agricultural efficiency in China. *The Review of Economic Studies*, 88(4), 1831–1862.
- Chen, C., Restuccia, D., & Santaaulàlia-Lllops, R. (2022). The effects of land markets on resource allocation and agricultural productivity. *Review of Economic Dynamics*, 45, 41–54.
- Eckstein, A., Chao, K., & Chang, J. (1974). The economic development of Manchuria: The rise of a frontier economy. *The Journal of Economic History*, 34(1), 239–264.
- Fafchamps, M. (1992). Cash crop production, food Price volatility, and rural market integration in the third world. *American Journal of Agricultural Economics*, 74(1), 90–99.
- Huang, P. C. C. (1985). *The peasant economy and social change in North China*. Stanford: Stanford University Press.
- Jin, S., Min, S., Huang, J., & Waibel, H. (2021). Falling price induced diversification strategies and rural inequality: Evidence of smallholder rubber farmers. *World Development*, 146, Article 105604.
- Kebede, H. A. (2022). Market integration and separability of production and consumption decisions in farm households. *Journal of Development Economics*, 158, Article 102939.
- Kong, J. (1986). *Dongbei Jingishi (economic history of Manchuria)*. Chengdu: Sichuan Renmin Press.
- Kung, J. K.-s., & Li, N. (2011). Commercialization as exogenous shocks: The effect of the soybean trade and migration in Manchurian villages, 1895–1934. *Explorations in Economic History*, 48(4), 568–589.
- Lagakos, D., & Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2), 948–980.

- Lei, H. (1981). *Dongbei de Douhuo Maoyi (Manchurian soybean trade)*. Taipei: Institute of History, Taiwan National Normal University.
- Li, D., & Li, N. (2017). Moving to the right place at the right time: Economic effects on migrants of the Manchuria plague of 1910–11. *Explorations in Economic History*, 63, 91–106.
- Lucas, R. E. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, 9(2), 508–523.
- Ma, D. (2008). Economic growth in the lower Yangzi region of China in 1911–1937: A quantitative and historical analysis. *The Journal of Economic History*, 68(2), 355–392.
- Makate, C., Angelsen, A., Holden, S. T., & Westengen, O. T. (2022). Crops in crises: Shocks shape smallholders' diversification in rural Ethiopia. *World Development*, 159, Article 106054.
- Martin, S. M., & Lorenzen, K. (2016). Livelihood diversification in rural Laos. *World Development*, 83, 231–243.
- Masanjala, W. H. (2006). Cash crop liberalization and poverty alleviation in Africa: Evidence from Malawi. 35(2), 231–240.
- Michler, J. D., & Josephson, A. L. (2017). To specialize or diversify: Agricultural diversity and poverty dynamics in Ethiopia. *World Development*, 89, 214–226.
- Montero, E. (2022). Cooperative property rights and development: Evidence from land reform in El Salvador. *Journal of Political Economy*, 130(1), 48–93.
- Myers, R. H. (1970). *The Chinese Peasant Economy: Agricultural Development in Hopei and Shantung, 1890–1949*. Cambridge: Harvard University Press.
- Myers, R. H. (1976). Socioeconomic change in villages of Manchuria during the Ch'ing and republican periods: Some preliminary findings. *Modern Asian Studies*, 10(4), 591–620.
- Ogutu, S. O., & Qaim, M. (2019). Commercialization of the small farm sector and multidimensional poverty. *World Development*, 114, 281–293.
- Otsuka, K. (2007). Chapter 51 efficiency and equity effects of land markets. In R. Evenson, & P. Pingali (Eds.), 3. *Handbook of agricultural economics* (pp. 2671–2703). Elsevier.
- Perkins, D. H. (1970). *Agricultural development in China, 1368–1968*. Edinburgh University Press.
- Rawski, T. G. (1989). *Economic growth in prewar China*. University of California Press.
- Rivera-Padilla, A. (2020). Crop choice, trade costs, and agricultural productivity. *Journal of Development Economics*, 146, Article 102517.
- Shi, Z. (2017). *Agricultural development in Qing China—a quantitative study, 1661–1911*. Leiden: Brill Academic Publishers.
- Tesfaye, W., & Tirivayi, N. (2020). Crop diversity, household welfare and consumption smoothing under risk: Evidence from rural Uganda. *World Development*, 125, Article 104686.
- Weinberger, K., & Lumpkin, T. A. (2007). Diversification into horticulture and poverty reduction: A research agenda. *World Development*, 35(8), 1464–1480.
- Zhang, J., Mishra, A. K., & Hirsch, S. (2021). Market-oriented agriculture and farm performance: Evidence from rural China. *Food Policy*, 100, Article 102023.