Agricultural Mechanization and Structural Transformation

Ling Yao*

September 2024

Click here for the latest version.

Abstract

This paper measures the effect of subsidized agricultural mechanization on non-agricultural employment. Combining provincial agricultural machinery subsidy catalogs with purchase records in China, I construct city-level exposure to the subsidy as a shift-share instrument for mechanization. I find no evidence that mechanization increases non-agricultural employment in the local area or in migration destinations, except in certain sectors in migration destinations. Individual-level analysis further reveals that in more mechanized areas men return to agriculture while women withdraw from non-agricultural jobs without participating more in farming. This study shows that mechanization plays a limited role in accelerating structural transformation. In the meantime, it reinforces gender inequality in agriculture.

Keywords: Agricultural mechanization; structural transformation; gender; shift-share instrumental variable

*University of Minnesota, yao00057@umn.edu.

[1]I am grateful to Marc Bellemare, Terry Hurley, Metin Çakır, David Samuels, Paul Glewwe, Anne Fitzpatrick, Ruiqing Miao, Colette Salemi, Chris Boyd, Matthew Bombyk, Stephen Pitts, SongYi Paik, Jhih-Yun Liu, Adriana Castillo Castillo, Berenger Djoumessi Tiague, Yanxu Long, Amanda Agan, Caroline Krafft, Yang Song and participants at the 2023 AAEA annual meeting for their helpful suggestions and comments. All errors and omissions are mine.

1 Introduction

In the past few decades, labor has been moving out of agriculture to the non-agriculture sectors in low- and middle-income countries (LMICs) at an unprecedented speed, raising concerns over who will produce food for the world even in labor-abundant countries. Such rapid structural transformation can explain the adoption of labor-saving agricultural technologies, according to the induced innovation theory (Hayami & Ruttan, 1971). Conversely, it is also possible that labor-saving technologies fuel structural transformation by releasing labor from agriculture, according to the labor-push hypothesis (Alvarez-Cuadrado & Poschke, 2011; Gallardo & Sauer, 2018).

Does the labor-push hypothesis apply in current LMICs that are experiencing agricultural mechanization? First, agricultural mechanization may not provide a push to structural transformation if it lags behind the outflow of farm labor. The adoption of machinery in LMICs is slow due to various market failures on both the supply and demand sides (Diao et al., 2020; Pingali, 2007), such as the low level of crop intensification, the lack of custom hiring services market among smallholders and insufficient knowledge on machinery use and repair. Furthermore, labor-augmenting technologies are not necessarily labor-saving (Bustos et al., 2016). Whether agricultural mechanization has displaced labor also depends on the nature of mechanized tasks, machinery types and the skills of farm labor (Hamilton et al., 2022; Pingali et al., 1987), as well as labor demand in non-agricultural sectors (D. Autor et al., 2019; Autor & Salomons, 2018).

This paper investigates whether local agricultural mechanization increases employment in non-agricultural sectors locally and in migration destinations. To control for unobserved confounders between labor market outcomes and mechanization, I exploit the variation in mechanization that comes from a machinery subsidy scheme determined at a higher administrative level (province) than the local observed units (cities). By interacting the common shock of machinery subsidy with local machinery use patterns, local exposure to the shock is constructed as a shift-share instrumental variable for mechanization. In terms of Borusyak, Hull, and Jaravel (2022)'s shift-share IV formalization, the identification of the local average treatment effect hinges on the plausible exogeneity of the machinery subsidy scheme to local level unobservable confounders after controlling for city and time fixed effects. To precisely measure the local mechanization level, I assemble a dataset of universal subsidized machinery purchase records within cities by machinery type, then construct the shift-share instrumental variables by combining lagged machinery purchase patterns with annual machinery subsidy schemes from provincial government publications.

This study looks at the labor market outcomes of mechanization in three different ways. First, with administrative employment data, I show that mechanization has no effect on the local secondary and tertiary sectors as a whole and only increases employment in the wholesale and retail sectors. The result is robust with units of observation defined at a more granular level. Second, to capture occupational change accompanied by labor migration, I match migration destinations with origins and find that mechanization in the origin cities boosts employment in the manufacturing, wholesale and retail sectors in the destination city, but the effect on all non-agricultural sectors is negligible. Lastly, household-level analysis with survey data shows that mechanization has asymmetric impacts on men and women's labor allocation. In more mechanized areas, working-age men are more likely to work in agriculture but not working-age women. In the meantime, women's participation in off-farm jobs falls.

This study is directly linked to the farming system evolution theory (Binswanger, 1986; Pingali et al., 1987; Pingali, 2007). Drawing from historical evidence, Binswanger (1986) pointed out that the employment outcome of mechanization depends on land scarcity, non-agricultural labor demand and the profitability of alternative techniques. The nature of mechanized operations also shapes labor market outcomes. Pingali (2007) observed that power-intensive machinery such as milling and tilling machinery often does not lead to labor displacement. Autor (2019) pointed out that automation also complements cognitive capabilities of workers and creates new employment opportunities. In line with the theoretical predictions, existing empirical studies have found mixed evidence. As summarized by Daum and Birner (2020), while some studies found reduced or increased labor use after mechanization, others did not find a statistically significant correlation. More recently, evidence from randomized control trials shows that mechanization improves household welfare by releasing family labor from supervision to non-agricultural jobs (Caunedo & Keller, 2021). In the context of China, researchers found a simultaneous increase in machinery service usage and non-farm employment (Zheng et al., 2022). Mechanization is also associated with out-migration which harms the local economy (Zou et al., 2024). Studies with a gender perspective found that mechanization reduces women's farm labor use in India (Afridi et al., 2023) and increases their non-farm employment in China (Ma et al., 2024).

This study adds to the existing literature by providing empirical evidence related to the predictions of the farming system evolution framework in a context where relatively abundant labor is moving fast out of agriculture, and mechanization is stimulated by government subsidy. While it is important to look at both the direct effect within the sector and the indirect effect on other sectors when evaluating whether technological progress harms overall employment (Autor & Salomons, 2018), most studies on agricultural mechanization look only within the agricultural sector. I extend the view to employment outcomes in non-agricultural sectors, helping to illustrate a fuller picture of labor dynamics during agricultural mechanization. Most of the early studies, as well as the most recent study (Ma et al., 2024) relied on cross-sectional data and limited identification strategies. With panel data and a shift-share instrumental variable design, this study aims to give parameter estimates a causal interpretation. This study also differs from most related empirical studies in that mechanization is measured at the local level rather than by household machinery ownership. The former is more appropriate in the context of smallholder-dominated agriculture, in which mechanization is mainly achieved through hiring custom machinery service (Yang et al., 2013).

The rest of this paper is organized as follows. Section 2 describes the data and gives and overview of mechanization, the subsidy policy, and employment structure in China, highlighting key descriptive statistics. Section 3 discusses the identification strategy and evaluates its validity. Section 4 to 6 present the results at different levels and Section 7 concludes.

2 Theoretical framework

With a simple two-sectors model, Bustos et al.(2016) show that the effect of a labor-augmenting technical change in agriculture (e.g. mechanization) on structural transformation depends on the nature of agricultural production. When labor and land are strong technical complements, a labor-augmenting technical change releases labor from agriculture to non-agricultural sectors because it reduces the marginal product of labor in agriculture. In this case, mechanization in agriculture is labor-saving. Whereas if labor and land are not strong technical complements, labor moves in the opposite direction. In addition, mechanization is also land-augmenting since it improves the timeliness and the quality of cultivation. Bustos et al.(2016) show that a land-augmenting technology increases the marginal product of labor in agriculture and induces a reallocation of labor back into agriculture.

Taking a closer look, technical change often generates heterogeneous effects on labor of different education levels (Autor et al., n.d.) and gender (Afridi et al., 2023). Indeed, it has long been recorded that the division of farm tasks is often gender-based. Women are less strong than men on average, and so they are less represented in high human energy-expenditure activities in rural settings (Pitt et al., 2012). In the pre-industrial era, the use of certain strength-intensive production technologies, such as ploughing (Alesina et al., 2013) and irrigation (Fredriksson & Gupta, 2023), creates long-lasting social norms of men working in the field and women staying at home and attending household tasks. While in areas with more precision-intensive agriculture, such as tea production, women play a greater role in farm production (Qian, 2008). Agricultural machinery favors men's participation in farm production since it requires heavy physical strength to work with. Moreover, there is a gender gap in the diffusion of new information and technology. Women are less likely to receive agricultural information from social networks (Beaman & Dillon, 2018) and are less trusted as a source of information on a new technology (BenYishay et al., 2020). Women may be excluded from the process of mechanization due to similar gender barriers. For example, they may not get formal and informal machinery operation and maintenance training from others as male farmers do. They may be perceived as less knowledgeable and skilled around machinery. This prevents women from improving their labor productivity on the farm and causes them to allocate labor away from agriculture.

To shed light on the heterogeneous labor market outcome of men and women during agricultural mechanization, I model the agricultural sector in an open local economy with functioning labor and land market. The output of the agriculture sector depends on the input of land (T) and effective labor (L). Assume the production function takes a constant elasticity of substitution technology and constant returns to scale of the following form:

$$F(L,T) = A_N \left[\gamma L^{\frac{\sigma-1}{\sigma}} + (1-\gamma)\left(\phi_T(M)T\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where γ is the share parameter that represents the relative importance of inputs, and σ is the elasticity of substitution between land and labor. A_N is Hicks-neutral technological change that augments the productivity of all inputs proportionally. Mechanization (M) augments land through an increasing function ϕ_T, which magnifies the land input.

The effective labor (L) is comprised of both female labor (L_f) and male labor (L_m). Assume it takes the following form:

$$L = \left[\alpha L_f^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha)\left(\phi_L(M)L_m\right)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}}, \tag{2}$$

where α reflects the relative importance of female versus male labor, and ϵ is the elasticity of substitution between female and male labor. The function ϕ_L indicates that male labor is augmented by mechanization[1].

If we further make the following plausible Assumptions 1-3, then we can derive the testable Propositions 1-3 about the impact of mechanization on agricultural labor allocation by gender.

Assumption 1: Female and male labor are substitutes, i.e. $\epsilon > 1$.

Assumption 2: It is easier to substitute between types of labor than between labor and land, i.e. $\epsilon > \sigma$.

Assumption 3: As the use of machinery increases, its augmenting effect on labor grows faster than on land, i.e. $\frac{\partial \frac{\phi_L(M)}{\phi_T(M)}}{\partial M} > 0$.

Proposition 1: Under Assumption 1, mechanization increases the equilibrium male-to-female labor input ratio in agriculture ($\frac{L_m}{L_f}$).

Proof: The marginal product of female labor is

$$MP_{Lf} = A_N \Delta^{\frac{1}{\sigma-1}} \cdot \gamma L^{-1/\sigma} \cdot \Omega^{\frac{1}{\epsilon-1}} \cdot \alpha L_f^{-1/\epsilon}, \tag{3}$$

and the marginal product of male labor is

$$MP_{Lm} = A_N \Delta^{\frac{1}{\sigma-1}} \cdot \gamma L^{-1/\sigma} \cdot \Omega^{\frac{1}{\epsilon-1}} \cdot (1-\alpha)\phi_L^{\frac{\epsilon-1}{\epsilon}}(M) L_m^{-1/\epsilon}, \tag{4}$$

where $\Delta = \gamma L^{\frac{\sigma-1}{\sigma}} + (1-\gamma)\left(\phi_T(M)T\right)^{\frac{\sigma-1}{\sigma}}$ and $\Omega = \alpha L_f^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha)\left(\phi_L(M)L_m\right)^{\frac{\epsilon-1}{\epsilon}}$ for the simplicity of notations.

[1]This setup is an adaptation of Afridi et al. (2023)'s model, in which mechanization enters the model as a Hicks-neutral technological change (A_N).

For interior solutions, the first-order profit maximization condition requires that the ratio of the marginal product of female and male labor equals to the ratio of their off-farm wages: $\frac{MP_{Lf}}{MP_{Lm}} = \frac{w_f}{w_m}$. Plugging in equations (3) and (4) we have

$$\frac{\alpha L_f^{-1/\epsilon}}{(1-\alpha)\phi_L^{\frac{\epsilon-1}{\epsilon}}(M)L_m^{-1/\epsilon}} = \frac{w_f}{w_m}. \tag{5}$$

Simplifying we have

$$\frac{L_m}{L_f} = \left(\frac{(1-\alpha)w_f}{\alpha w_m}\right)^{\epsilon} \phi_L^{\epsilon-1}(M). \tag{6}$$

Taking the first derivative with respect to M:

$$\frac{\partial(L_m/L_f)}{\partial M} = (\epsilon-1)\left(\frac{(1-\alpha)w_f}{\alpha w_m}\right)^{\epsilon} \phi_L^{\epsilon-2}(M). \tag{7}$$

The right-hand side expression in Equation (7) is positive under Assumption 1, meaning that when female and male labor are substitutes, an increase in mechanization (M) increases the male-to-female labor ratio. **Q.E.D.**

Proposition 2: Under Assumptions 1-2, mechanization reduces the equilibrium female labor input per unit of land ($\frac{L_f}{T}$).

Proof: The marginal product of land is

$$MP_T = A_N \Delta^{\frac{1}{\sigma-1}} \cdot (1-\gamma) \cdot \phi_T^{\frac{\sigma-1}{\sigma}}(M) T^{-1/\sigma}. \tag{8}$$

For interior solutions, the first-order profit maximization condition requires that the ratio of the marginal product of female labor and land equals the ratio of female wage and land rent(τ): $\frac{MP_{Lf}}{MP_T} = \frac{w_f}{\tau}$. Plugging in equations (3) and (8) we have,

$$\frac{\gamma L^{-1/\sigma} \cdot \Omega^{\frac{1}{\epsilon-1}} \cdot \alpha L_f^{-1/\epsilon}}{(1-\gamma) \cdot \phi_T^{\frac{\sigma-1}{\sigma}}(M) T^{-1/\sigma}} = \frac{w_f}{\tau}. \tag{9}$$

From equation (6) we know that $L_m = (\frac{(1-\alpha)w_f}{\alpha w_m})^\epsilon \phi_L^{\epsilon-1}(M) L_f$. Using this to substitute out L_m in L and Ω we have

$$L^{-1/\sigma} \cdot \Omega^{\frac{1}{\epsilon-1}} = L_f^{\frac{1-\epsilon/\sigma}{\epsilon}} \left[\alpha + (1-\alpha) \left(\phi_L(M) \frac{(1-\alpha)w_f}{\alpha w_m} \right)^{\epsilon-1} \right]^{\frac{1-\epsilon/\sigma}{\epsilon-1}}. \tag{10}$$

Plugging equation (10) back to equation (9) we have

$$\frac{w_f}{\tau} = \frac{\gamma\alpha \left[\alpha + (1-\alpha) \left(\phi_L(M) \frac{(1-\alpha)w_f}{\alpha w_m} \right)^{\epsilon-1} \right]^{\frac{1-\epsilon/\sigma}{\epsilon-1}} L_f^{-1/\sigma}}{(1-\gamma) \cdot \phi_T^{\frac{\sigma-1}{\sigma}}(M) T^{-1/\sigma}}, \tag{11}$$

which can be rearranged to

$$\frac{L_f}{T} = (\frac{\tau}{w_f})^\sigma \cdot \frac{(\gamma\alpha)^\sigma \left[\alpha + (1-\alpha) \left(\phi_L(M) \frac{(1-\alpha)w_f}{\alpha w_m} \right)^{\epsilon-1} \right]^{\frac{\sigma-\epsilon}{\epsilon-1}}}{(1-\gamma)^\sigma \phi_T^{\sigma-1}(M)}. \tag{12}$$

Equation (12) tells us that an increase in mechanization affects $\frac{L_f}{T}$ through both labor augmentation ϕ_L in the numerator and land augmentation ϕ_T in the denominator. Under Assumption 1 ($\sigma > 1$), an increase in M increases the denominator and reduces $\frac{L_f}{T}$. Under Assumption 2 ($\epsilon > \sigma$), an increase in M decreases the numerator and also reduces $\frac{L_f}{T}$. Moreover, the decrease in $\frac{L_f}{T}$ is stronger if the female labor market wage (w_f) is higher. **Q.E.D.**

Proposition 3: Under Assumptions 1-3, mechanization increases the equilibrium male labor input per unit of land ($\frac{L_m}{T}$).

Proof: For interior solutions, the first-order profit maximization condition requires that the ratio of the marginal product of male labor and land equals the ratio of male wage and land rent: $\frac{MP_{Lm}}{MP_T} = \frac{w_m}{\tau}$. Plugging in equations (4) and (8) we have

$$\frac{\gamma L^{-1/\sigma} \cdot \Omega^{\frac{1}{\epsilon-1}} \cdot (1-\alpha)\phi_L^{\frac{\epsilon-1}{\epsilon}}(M) L_m^{-1/\epsilon}}{(1-\gamma)\cdot \phi_T^{\frac{\sigma-1}{\sigma}}(M) T^{-1/\sigma}} = \frac{w_m}{\tau}. \tag{13}$$

Again from the equation (6) we can get $L_f = \left(\frac{\alpha w_m}{(1-\alpha)w_f}\right)^{\epsilon} \phi_L^{1-\epsilon}(M) L_m$. Using this to substitute out L_f in L and Ω we have

$$L^{-1/\sigma} \cdot \Omega^{\frac{1}{\epsilon-1}} = L_m^{\frac{1-\epsilon/\sigma}{\epsilon}} \phi_L^{\frac{1-\epsilon/\sigma}{\epsilon}}(M) \left[\alpha \left(\frac{(1-\alpha)w_f}{\alpha w_m}\right)^{\epsilon-1} \phi_L^{1-\epsilon}(M) + (1-\alpha)\right]^{\frac{1-\epsilon/\sigma}{\epsilon-1}} \tag{14}$$

Plugging equation (14) back into equation (13) we have

$$\frac{w_m}{\tau} = \frac{\gamma(1-\alpha)\left[\alpha\left(\frac{(1-\alpha)w_f}{\alpha w_m}\right)^{\epsilon-1} \phi_L^{1-\epsilon}(M) + (1-\alpha)\right]^{\frac{1-\epsilon/\sigma}{\epsilon-1}} \phi_L^{\frac{\epsilon-\epsilon/\sigma}{\epsilon}}(M) L_m^{-1/\sigma}}{(1-\gamma)\cdot \phi_T^{\frac{\sigma-1}{\sigma}}(M) T^{-1/\sigma}}, \tag{15}$$

which can be rearranged to

$$\frac{L_m}{T} = (\frac{\tau}{w_m})^{\sigma} \cdot \left[\frac{\gamma(1-\alpha)}{1-\gamma}\right]^{\sigma} \cdot \left[\alpha\left(\frac{(1-\alpha)w_f}{\alpha w_m}\right)^{\epsilon-1} \phi_L^{1-\epsilon}(M) + (1-\alpha)\right]^{\frac{\epsilon/\sigma-1}{1-\epsilon}} \cdot \left[\frac{\phi_L(M)}{\phi_T(M)}\right]^{\sigma-1}. \tag{16}$$

Equation (16) tells us that under Assunption 1 and 3 ($\sigma > 1$ and $\frac{\partial \frac{\phi_L(M)}{\phi_T(M)}}{\partial M} > 0$), an increase in M has a positive effect on $\frac{L_m}{T}$ through the last component. Under Assunption 2 ($\epsilon > \sigma$), an increase in M has a positive effect on $\frac{L_m}{T}$ through the second last component. **Q.E.D.**

3 Background and data

This section explains the design of the agricultural machinery purchase subsidy policy, which is essential for understanding the source of identification in the shift-share instrumental variable design. It then introduces the development of agricultural mechanization and the transition in employment structure in China. Description of data collected for this study and other supportive statistics are woven into the corresponding subsections.

3.1 Agricultural machinery purchase subsidy policy in China

China's machinery purchase subsidy started in 2004 and underwent several rounds of expansion and adaptation. It has transformed from a pilot program in a small group of selected counties to become a nationwide agricultural development policy. The subsidy is mainly financed by the central government, and local governments are allowed to contribute voluntarily. As shown in Figure 1, the central government budget increased from 70 million yuan (around 10 million US dollars) in 2004 to 23.64 billion yuan (around 3.5 billion US dollars) in 2015.

In the earlier years of implementation, the machinery purchase subsidy only applied to a catalog of manufacturers and prototypes selected by local governments. The subsidy came in the form of reduced retail price for machinery customers who had applied for the subsidy and obtained approval before purchase. Manufacturers were compensated for eligible sales annually. Not surprisingly, such arrangements led to noticeable local government rent-seeking behavior and inflated machinery price. Starting in 2015, the current model was established to minimize market distortion and corruption. The subsidy became neutral to manufacturers and brands, allowing for market competition. Every spring the provincial governments announce a catalog of subsidized machinery items, specifying key features of eligible items and the

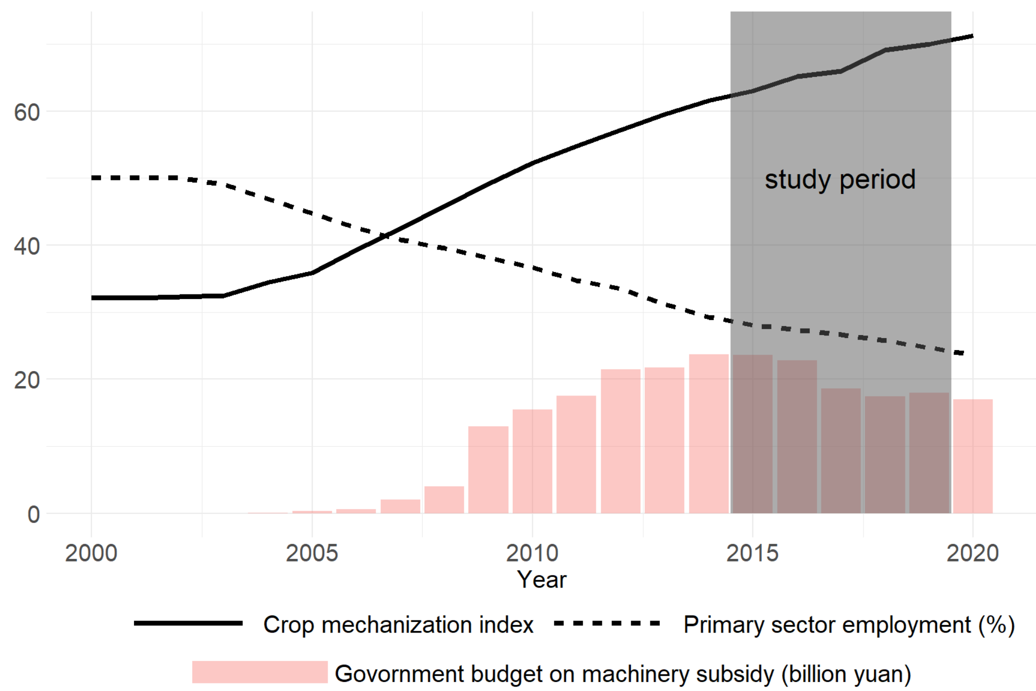

Figure 1: Crop mechanization, employment and subsidy budget

Notes: The subsidy budget data is from government announcements (http://www.gov.cn). The crop mechanization index data is from *China Agricultural Machinery Industry Yearbooks.* The index is a weighted average of the percentage of crop area plowed, sown and harvest by machinery across crops. The primary sector includes agriculture, forestry, animal husbandry and fishing. The primary sector employment data is from the National Statistics Bureau.

corresponding subsidy level. Manufacturers then register their products that satisfy the specified features in the catalog. Since all agricultural producers have become eligible for the subsidy, the requirement for pre-purchase approval was lifted. After purchase, customers report the transaction to the county level governments and receive the subsidy directly from the central government once the transaction is verified.

Provincial governments are required to publish all subsidized purchases on their websites for the sake of transparency. I collected these purchase and subsidy records and aggregated them to the county and city levels to measure changes in the local mechanization level. The amount of subsidized machinery purchase can largely reflect changes in mechanization because the subsidy covers a comprehensive range of implements covering all stages of production, and it does not screen off applicants.[2] The subsidy catalogs are used to contruct the shift-share instrumental variable. Details of the data compilation process and the variable construction process are reported in Appendix A.

Since the key identification assumption of the shift-share instrumental variable is that provincial subsidy levels of different machinery items and sizes are not confounded by county level unobservables when weighted by market shares, as discussed in detail in Section 3, it is worthwhile to evaluate factors behind the subsidy determination process. A typical provincial subsidy catalog covers machinery implements of all crop production stages, including land preparation, sowing and planting, crop management, harvesting and primary processing. It also includes tractors and livestock husbandry equipments. Within each item, the subsidy level varies by size range.[3] In de-

[2]There is a limit on the number of units each household can purchase with subsidy per year. For most provinces the limit is five and it is rarely binding. The limit for organizations such as cooperatives is higher, usually 10 or 15.

[3]For instance, in Heilongjiang province in 2020, the subsidy levels for crawler tractors are 38,100 yuan ($5720), 46,800 yuan ($7026), 55,800 yuan ($8378), 56,700 yuan ($8513) for 100-110 horsepower, 110-120 horsepower, 120-130 horsepower, 130-140 horsepower, respectively.

signing the annual catalog of subsidized machinery, provincial governments follow the guidelines from the central government, which updates every three years. The guideline specifies the required and elective machinery items to be subsidized and the maximum central government subsidy for each item and size. Provincial governments then choose items and sizes and assign the implemented subsidy level, with the option to supplement certain items and sizes with local budget. Beyond the central government guidelines, factors behind provincial governments' decision of the implemented subsidy levels are unobservable to researchers. At minimum, each provincial government considers a convoluted group of factors. In the majority of cases, the implemented subsidy does not deviate far away from the central government maximum. In fact, 54% of the implemented subsidy is set at the central government maximum and 90% of the implemented subsidy is above 80% of the central government maximum. On average of the purchase records, the subsidy accounted for 29% of the retail price of the purchased items. The distribution of the subsidy-retail price ratio is shown in Figure 2.

3.2 Agricultural mechanization in China

Before the mechanization policy, the agricultural mechanization level in China ranked in the lower-middle tier globally. As in Figure 3, the number of tractors by total arable land was lower than in other Asian LMICs such as India and Vietnam and was higher than in Sub-Saharan African countries. Japan has the highest number of tractors relative to arable land among all countries. Tractors used in Japan are compact to accommodate relatively small farm sizes, contributing to the large number far above other developed countries. The density of tractors in China was less than 10% of Japan or South Korea although the average farm size is comparable.

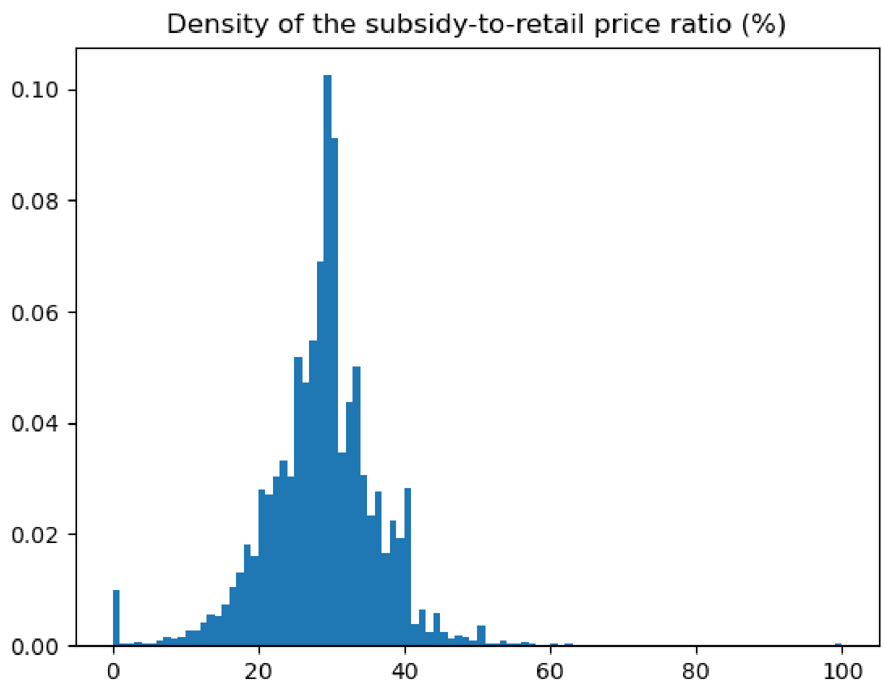

Figure 2: Density of the subsidy-to-retail price ratio (%)

Notes: The author's computation based on the subsidized machinery purchase data.

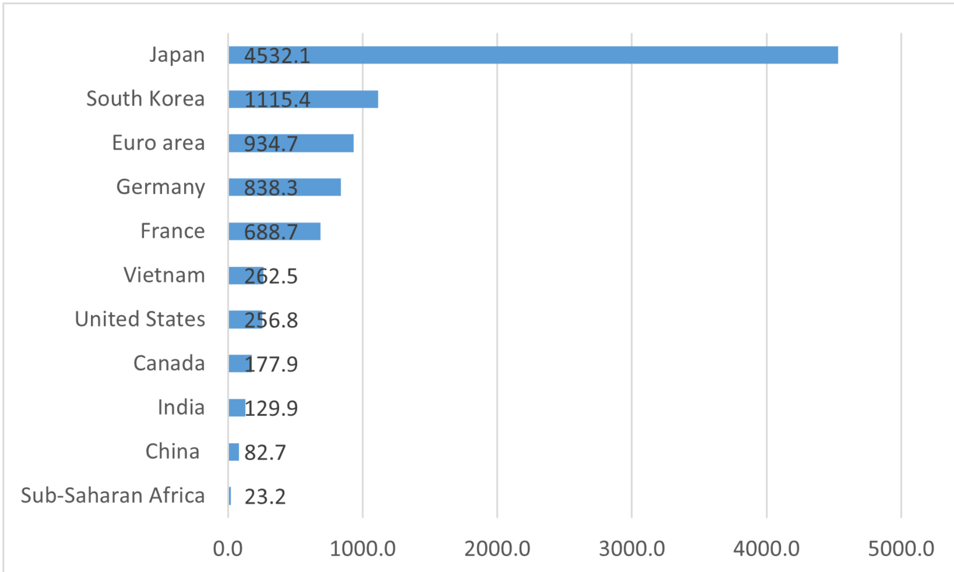

Figure 3: Agricultural machinery, tractors per 100 sq. km of arable land in year 2000
Notes: Data from World Bank. Data for Sub-Saharan Africa is for year 1987.

Under the subsidy, mechanization in China has undergone substantial growth. Total machinery power doubled from 525.7 million KW in 2000 to 1055.5 million KW in 2020. The mechanization index in Figure 1 measures the weighted average share of plowing, sowing and harvesting tasks completed by machines versus labor among major crops. Before the start of the subsidy in 2004, the portion of farming tasks completed by machinery was stable around one third. After 2004, the growth of the mechanization index mirrored the increase in the subsidy budget and the total machinery power, reaching over 70% in 2020.

Among the top 5 major crops, wheat is the most mechanized. As shown in Figure 4, more than 90% of the wheat sown acreage is plowed, sown and harvested by machinery. Other crops face technical bottlenecks at different production stages, as reflected in lagged mechanization rate compared to the land preparation stage. Rice has similar mechanization rate as wheat at the land preparation and harvesting stages, but only half of the sowing and transplanting is completed by machines. Corn and soybean are least mechanized at the harvesting stage due to the technical difficulty in separating grain from kernel without damage. Rapeseed has the lowest sowing and harvest mechanization rate because handling tiny seeds requires high precision.

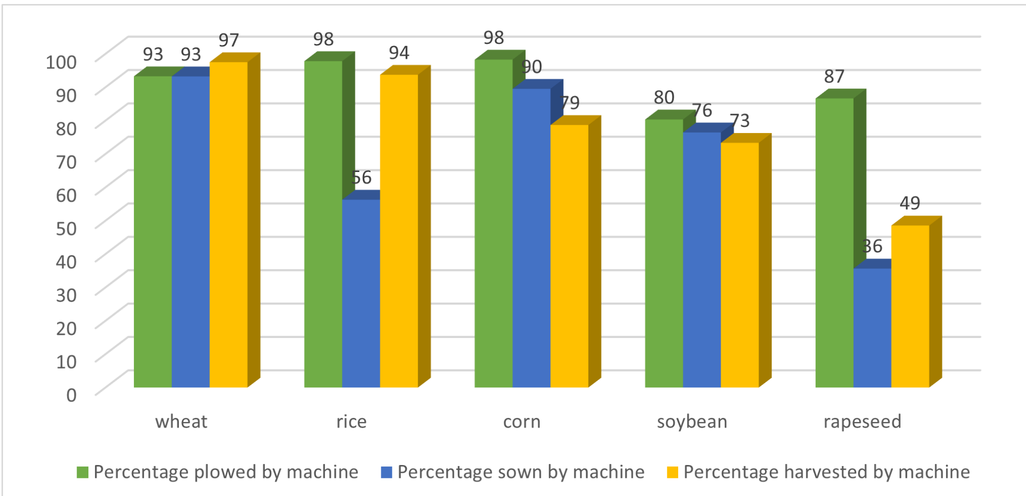

Figure 4: Percentage of mechanized area by crop and production stage in year 2020

Notes: Author's calculation based on sown area data from the National Bureau of Statistics and machinery plowed, sown and harvested area from the *China Agricultural Machinery Industry Yearbook 2021.* The numbers for corn come directly from *China Agricultural Machinery Industry Yearbook 2021.* For other crops, the mechanization rate for plowing could be slightly underestimated since a portion of acreage is under no-till practice or does not require plowing.

Due to China's diverse geography, mechanization level is uneven across regions. When mechanization is measured by the portion of tasks completed by machinery, the spatial distribution roughly follows the landscape pattern. The Northeast Plain and the North Plain are naturally suitable for large-size machinery operations, allowing the economies of scale. Whereas the rugged landscape in the southwest Sichuan Basin and the Yunnan–Guizhou Plateau makes it challenging for machinery to operate and be transported. At the land preparation stage, the wide adoption of small-sized machinery such as mini-cultivators has narrowed the mechanization gap created by geographic limitations (as shown in the upper-right panel in Figure 5). At the sowing and harvesting stages, well-designed, durable and economic small-sized machinery implements are still in short supply. As shown in the bottom row in Figure 5, mechanization in sowing and harvesting in rugged areas lags behind other regions. Nevertheless, when mechanization is measured by machinery power per hectare of arable land, the wealthier eastern provinces have a higher farm capital to arable land ratio (as shown in the upper-left panel in Figure 5) with the exception of Tibet and Qinhai Autonomous Regions, where arable land is scarce and the capital-intensive livestock husbandry is the major form of agriculture.

Figure 5 reveals that simplified and aggregated mechanization measures such as the number of machinery units or the total machinery power per unit of arable land mask important variations related to machinery size and the labor-substituting/complementing nature of different machinery types. Studies relying solely on such data may fail to delineate the true relationship between mechanization and the investigated outcome. The data collected for this study have the advantage of detailed information on each new machinery purchased under the subsidy policy, including machinery type and technical specifications, allowing me to tailor mechanization measures when aggregating to the regional level. For instance, in the case of tractors, I aggregate the horsepower of the purchased units. But for combine harvesters, I aggregate by the weight harvested per second. I run regressions separately for machines related to different production stages.

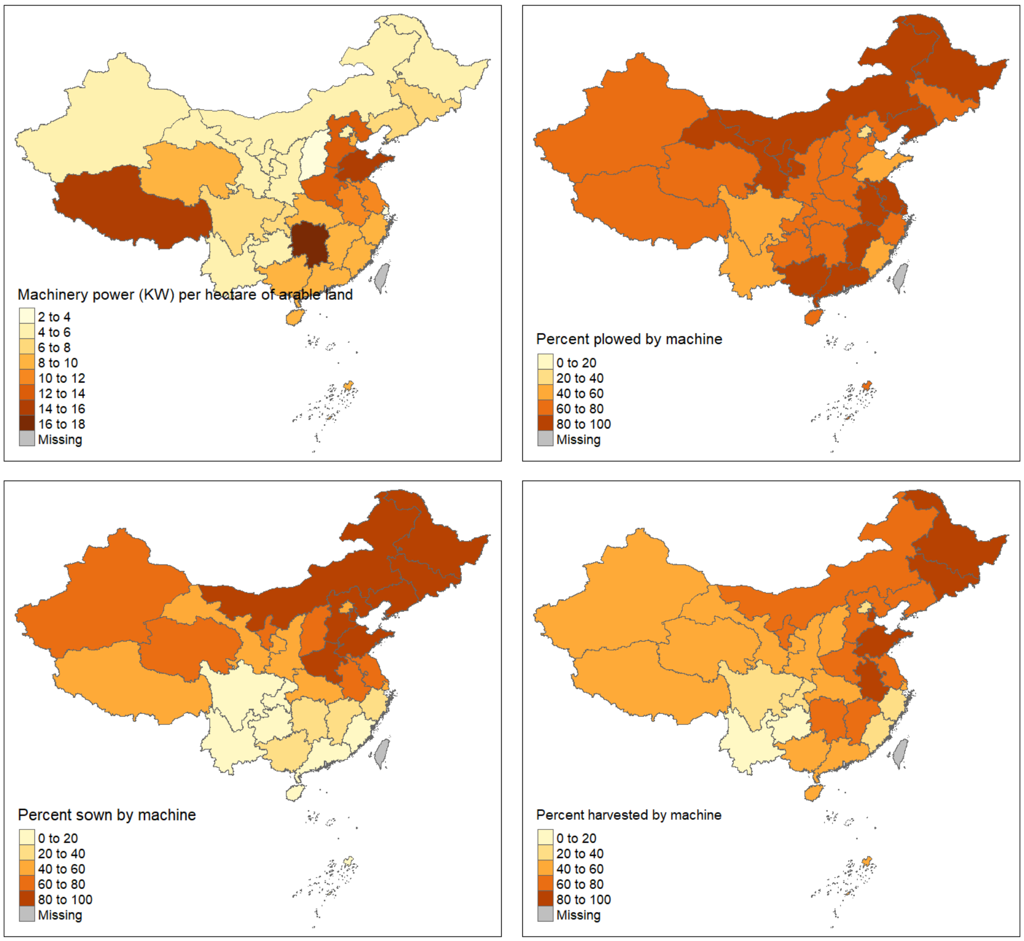

Figure 5: Mechanization by province in 2020

Notes: Author's calculation. Machinery power data, machine plowed, machine sown and machine harvested area data are from the China Agricultural Machinery Industry Yearbook 2021. Arable land data and total sown area date are from the 2006 Agricultural Census.

3.3 Agricultural labor and structural transformation in China

I measure employment at different levels using administrative data. The benchmark regressions are based on reported secondary and tertiary industry employment from the *China City Yearbooks 2016 - 2020*, with information for years 2015-2019.[4] The yearbooks cover around 290 level units out of the total 333 in China. The missing ones are minority autonomous prefectures in remote areas. In the yearbooks, primary industry includes agriculture, forestry, livestock and fishery. Secondary industry includes manufacturing, energy and water supply and construction.[5] Tertiary industry is defined as all industries other than primary and secondary industries, including wholesaling and retailing, transportation, logistics and postal service, accommodation and food service, information technology and finance.

The source of employment information in the yearbooks is corporate unit registrations at the Human Resources and Social Security Bureau and the Administration for Market Regulation. Compared to other data sources with similar employment structure measures, such as the China National Population Census, the yearbooks are less prone to reporting error in terms of classifying employment by sectors, making it more suitable for the purpose of this study. Since the information comes only from registered employers, however, it inevitably omits employment in informal sectors, especially those working in logistics, retail and restaurants (Yue, 2005). Informal sector jobs are frequent choices of migrant labor. Missing employment information in these sectors leads to underestimation of potential employment structure effects. This caveat is important to consider when interpreting the results.

The administrative data shows that from 2003 to 2019, the average num-

[4]The Chinese government halted publishing city level employment data since 2020.

[5]For historical reasons, mining is also classified as secondary industry in Chinese administrative data. To be consistent with international standards, I dropped mining when constructing the secondary and tertiary industry employment variable.

ber of non-primary sector employment per city increased from 355 thousand to 575 thousand, by 62%. Figure 6 shows that the eastern and coastal cities experienced the largest growth.

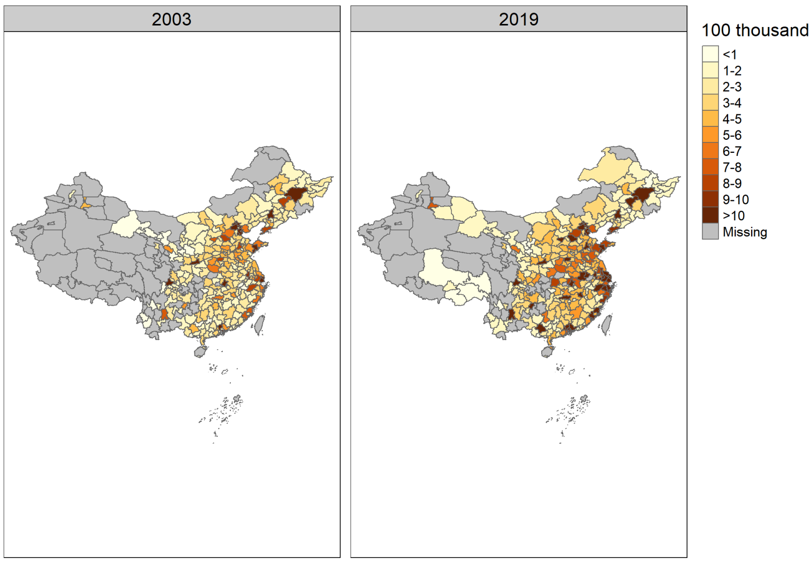

Figure 6: Number of employees in the second and tertiary industries
Notes: Data from *China City Statistical Yearbooks* 2014 - 2020. The primary sector includes agriculture, forestry, animal husbandry and fishery. The secondary industry includes mining, manufacturing, energy, water supply and construction sectors. The tertiary industry includes all service sectors.

4 Research design

The goal of this study is to estimate the causal effect of agricultural mechanization on local employment. It can be expressed as α_1^m in the regression

$$Y_{pct} = \alpha_0^m + \alpha_1^m mechanization_{pct}^m + \epsilon_{pct}^m, \tag{17}$$

where Y_{pct} is the total number of labor employed in the secondary and the tertiary industries in city c of province p in year t. The key explanatory variable $mechanization^m_{pct}$ is the amount of machinery power of category m purchased in the same city in the same year.

4.1 The shift-share instrument

A prevalent empirical challenge of the model represented by Equation (17) is the endogeneity of mechanization. Cities with faster structural transformation attract more local agricultural labor to non-agricultural sectors, inducing a higher demand for agricultural machinery. A higher level of industrialization means workers from farming households can send more remittance back to fund agricultural machinery purchase (Diao et al., 2020). Moreover, city-specific unobservable time-variant factors could contribute to both structural transformation and agricultural mechanization. For example, the transformation of farm land to industrial land creates more industrial employment opportunities and reduces agricultural production, thus the need for agricultural machinery. Urban sprawl may also change the nature of agriculture: e.g. from machinery-intensive crop production to labor-intensive recreational agriculture.

In order to identify the causal effect of mechanization, mechanization must be exogenized. The identification strategy in this paper is motivated by the source of plausibly exogenous shock: the machinery subsidy scheme determined at a higher level (province) than the observed units (cities). Each machinery subsidy scheme specifies a catalog of machinery types and sizes and their corresponding subsidy levels. While the machinery subsidy scheme applies equally to all cities within a province, each city is exposed to the subsidy at a different level due to their difference in the local machinery usage pattern. For example, mountainous cities where small-sized tractors have a greater market share than large-sized tractors benefit more when the provincial government decides to subsidize small tractors more heavily. In comparison, cities with flatter terrain benefit more when large tractors are subsidized more heavily.

The shift-share instrumental variable Z^m_{pct} for mechanization is the inner product of the city level market shares ($share_{jpc,t-1}$) in the previous year and the province-year level subsidy ($subsidy_{jpt}$) in the current year of each machinery size (indexed by j) within the category m.

$$Z^m_{pct} = \sum_{j=1}^{J} share^m_{jpc,t-1} \times subsidy^m_{jpt}$$

To the extent that machinery purchase responds to the subsidy incentive, the 2SLS yields an estimate for the local average treatment effect (LATE) of mechanization on non-agricultural employment.

4.1.1 Relevance

I evaluate the relevance of the shift-share instrumental variable across six machinery categories: tractors, rotary tillers, seeders/transplanters, grain harvesters, corn harvesters and grain dryers. I use the market shares of machinery sizes in the previous year because it provides higher relevance compared to the market shares in the baseline year of 2015. Table 1 shows that the Montiel-Plueger robust F statistic (Olea & Pflueger, 2013) for the instrumental variable is 34.56 for tractors, the highest among all categories. Each 1000 yuan increase in subsidy boosts 763 horsepower of increase in tractor power in a city. The F statistics for other categories are below or around 10, not granting consistent second stage estimates. With data at the county level, which is a more disaggregated administrative unit, the F statistics is 78.4 for tractors (Table 5), with other machinery categories showing limited response to the subsidy.

Table 1: First stage estimates with machinery purchase subsidy exposure as the shift-share instrumental variable for machinery purchase

	Dependent variable: Machinery purchase (in 1000)					
	Tractor Horsepower	Rotary tiller Width (m)	Seeder,transplanter Rows	Grain harvester Speed (kg/s)	Corn harvester Rows	Grain dryer Capacity (ton/day)
	(1)	(2)	(3)	(4)	(5)	(6)
Shift-share IV = market shares × subsidy (1000 yuan)	0.763***	115.775	0.028	0.010***	0.002	−0.007**
	(0.141)	(141.604)	(0.018)	(0.004)	(0.002)	(0.003)
Montiel-Pflueger robust F stat. for weak instrument	34.567	0.768	2.255	9.664	1.047	8.076
Observations	643	563	607	611	346	526
R^2	0.061	0.003	0.017	0.071	0.004	0.169
Adjusted R^2	0.055	-0.004	0.010	0.064	-0.007	0.163
Residual Std. Error	49.123	1,284.600	2.680	1.010	0.849	1.737

Notes: This table reports the first stage estimates of the shift-share instrumental variable with city level data in year 2015 - 2019 from China. The shift-share instrumental variable measures the city's exposure to the machinery subsidy policy depending on the market shares in the previous year. The endogenous variable is the subsidized machinery purchase. The reported standard errors in the parenthesis are heteroskedasticity-robust and are clustered at the city level. The city FE is included by first-differencing the variables and the shocks, as suggested by Borusyak et al. (2022). The year FE are included using dummy variables. *p<0.1; **p<0.05; ***p<0.01.

Based on the first stage results, I select tractor purchase as the main indicator of mechanization in the subsequent analysis. Tractor is a staple machinery that is indispensable in all types of agricultural mechanization scenarios. It provides drive for other machinery implements and is needed for transportation.

4.1.2 Validity

Figure 7 illustrates the identification strategy in a directed acyclic graph (DAG). The goal is to identify the causal effect of mechanization on non-agricultural employment at the city level, and the challenge is to circumvent the unobservable confounder u that is correlated with both variables, e.g. changes in land use. The subsidy levels determined at the province level serve as the source of exogenous variation in mechanization that is orthogonal to the city-level unobservables u and v, conditional on city and year fixed effects. To improve relevance, the subsidy levels are interacted with city-level market shares to construct the shift-share instrumental variable for mechanization. Since the market share only serve as weights, they can be correlated with the outcome through the unobserved v.

The critical source of identification stems from the variations in the subsidy level across the machinery types and sizes within each provincial annual subsidy scheme. The exclusion restriction of the shift-share instrument is satisfied as long as the provincial governments do not strategically design the subsidy scheme based on city-level unobservable shocks to employment structure. In other words, within provinces, all the machinery types and sizes have the same expected subsidy rate, and any deviation from the expected rate is due to factors that are irrelevant to city-level unobservable shocks to employment structure.

The key identifying assumption is that Z^m_{pct} is uncorrelated with the error

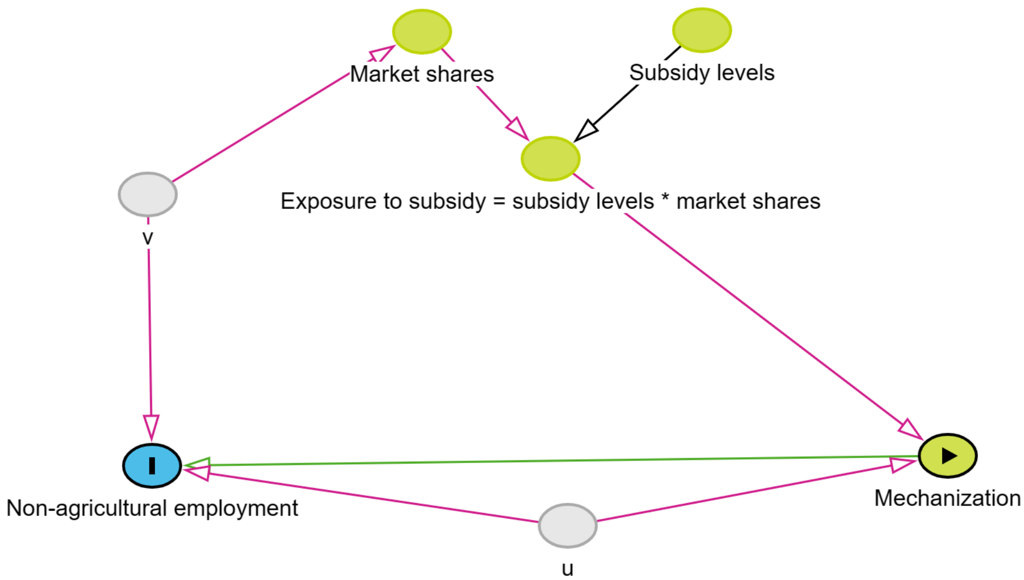

Figure 7: Directed acyclic graph of the shift-share IV design

term in Equation (17):

$$E\left[\sum_{t=1}^{T}\sum_{c=1}^{C}\left(\sum_{j=1}^{J} share^{m}_{jpc,t-1} \times subsidy^{m}_{jpt}\right) \times \epsilon_{pct}\right] = 0. \tag{18}$$

Per Borusyak et al.(2022), the market shares of machinery only serve as weights, and they are allowed to be correlated with the error terms.[6] To see this more clearly, re-arrange the moment condition in Equation (18) to a shock level (in the case of this study, the province-year-machinery size level)

[6] An alternative identification strategy is to rely on the exogeneity of the shares, as formalized by Goldsmith-Pinkhman, Sorkin and Swift (2020). See discussions in Adão, Kolesár & Morales (2019) and Borusyak et al. (2022). The exogenous market shares assumption are less likely to hold in the context of this study than the exogeneity of the shocks (subsidy levels).

equation, such that

$$E\left[\sum_{t=1}^{T}\sum_{p=1}^{P}\sum_{j=1}^{J} share_{jp,t-1}^{m} \times subsidy_{jpt}^{m} \times \bar{\epsilon}_{jpt}\right] = 0, \tag{19}$$

where $share_{jp,t-1}^{m} = \frac{1}{C_p}\sum_{c=1}^{C_p} share_{jpc,t-1}^{m}$. $\bar{\epsilon}_{jpt} = \frac{\sum_{c=1}^{C_p} share_{jpc,t-1}^{m} \times \epsilon_{pct}}{share_{jp,t-1}^{m}}$ is a share-weighted error term summed to the province-year-machinery size level. In words, Equation (19) states that the subsidy rates should not be correlated with the city-level unobservable shocks to secondary and tertiary industry employment when they are both weighted by the machinery market shares.

The moment condition in Equation (19) is satisfied if the following condition holds

$$E[subsidy_{jpt}^{m} | share_{jp,t-1}^{m}, \bar{\varepsilon}_{jpt}] = \mu \quad \forall j, p, t. \tag{20}$$

Equation (20) states that conditional on the province average market shares of machinery sizes and the share-weighted unobservables, all machinery sizes should have the same expected subsidy level in each province in each year.

In practice, provinces often consistently set the subsidy for certain machinery sizes to be low or at zero. For example, Northeastern provinces with large farm sizes never subsidized small tractors because small tractors are not useful in the region. In this case the expected subsidy rates for small tractors in these provinces are always zero and is different from large tractors. To account for such time-invariant local feature, I control for city fixed effects. Following the suggestions of Borusyak et al.(2022) and Millimet & Bellemare (2023), I use first differencing rather than dummy variables since the market shares are from the previous year and are time-variant. To construct the transformed shift-share instrumental variable, the subsidy levels are first-differenced and then multiplied to the shares. In addition, I control for year fixed effects to account for year-specific adjustments to the subsidy and employment trend.

The identifying assumption requires that the provincial governments will not adjust the subsidy levels of some cities' preferred machinery sizes in response to the underlying structural transformation trend in these cities. For example, when the province of Heilongjiang are deciding on the subsidy rate of crawler tractor of horsepower 100-120, the deviation of the subsidy rate from the previous year should not be correlated with the unobservable structural transformation shocks in cities that are the main markets of this tractor type. A violation of the exclusion restriction would require provincial governments to make sophisticated calibration in the subsidy scheme design aimed narrowly at offsetting city level shocks to employment structure. This is unlikely because the provincial government has to balance over a wide range of factors when designing the subsidy scheme. On average, each province-level administrative units contains around 10-20 city level and 84 county level administrative units.[7] Precisely targeting the potential machinery demand caused by the underlying structural transformation trends in all the cities is challenging, if not impossible. As discussed in Section 3.1, provincial governments design the annual subsidy schemes under the central government's national guidelines, which specify the eligible types and sizes of machinery and the maximum subsidy for each machinery type and size. They are required to reduce the subsidy when the subsidy is too high relative to the retail price[8] and when certain machinery types becomes outdated.

4.1.2.1 Falsification test To scrutinize the exclusion restriction, I conduct falsification tests that regress a city-level proxy for the unobserved confounders on the shift-share IV for tractors purchase. The proxy is the number of industrial firms in the previous year that reflects the demand for

[7]As of the end of 2017, there are 34 province-level administrative units, 334 city-level units and 2851 county-level administrative units in China.

[8]There is no strict line for overly high subsidy but 30% of the current retail price is commonly used as a reference.

non-agricultural labor when the subsidy policy is designed. The results reported in Table 2 show that across all model specifications (i.e. OLS, time fixed effect, first-differencing, and time fixed effect with first-differencing), the coefficient for the shift-share instrumental variable is not statistically different from zero, which means that there is no support for the hypothesis that the subsidy levels are assigned in response to potential labor demand from non-agricultural sectors.

Table 2: Falsification test results

	Dependent variable: Number of industrial firms in the previous year			
	OLS	Time FE	First Difference	Time FE + FD
	(1)	(2)	(3)	(4)
Shift-Share IV	2.269	1.101	−0.487	−0.394
	(3.973)	(4.307)	(0.885)	(0.890)
Observations	639	639	636	636
R^2	0.001	0.003	0.001	0.005
Adjusted R^2	-0.001	-0.003	-0.0001	-0.001
Residual Std. Error	918.607	919.790	162.869	162.947

Notes: This table reports the falsification test for the shift-share instrumental variable using the number of industrial firms (including manufacturing and mining) as a proxy for non-agricultural labor demand in the city when the subsidy levels are assigned by the provincial governments. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

4.1.3 Regularity condition: effective sample size

Borusyak et al.(2022) pointed out that there should be sufficient variation in the shift-share instrumental variable to ensure consistency of the estimates. In the canonical shift-share instrumental variable setting such as Autor, Dorn & Hanson (2013), the shock is common for all observations in the same time period so that the cross-sectional variation mainly comes from the shares. In this study, the variation in the shift-share IV comes from both the subsidies that varies at the province-year level and the market shares that vary at the city level. The more cities in a province vary in machinery market structure, the more they differ in their exposure to the same subsidy policy, and the higher the effective sample size. The variation in city market structure can be measured by the Herfindahl index of province average market shares $\sum_{j,p,t}(\frac{share^m_{jpt}}{\sum_{j,p,t} share^m_{jpt}})^2$. When cities differ in market structure, there is no dominating machinery size at the province level and the Herfindahl index converges to 0 as the number of observations goes to infinity. Table 3 summarizes the market shares for each machinery type. For tractors, the inverse of the Herfindahl index based on the previous year market shares is 421.9, much greater than the satisfactory level of 20 (Borusyak et al., 2022) to ensure that there is enough variation in the shares. On the last rows of Table 3, the largest province average market share is 1.1%, verifying that there is no market-dominating tractor size and that the tractor market structure varies sufficiently to ensure the large sample properties of the 2SLS estimates.

Table 3: Summary of subsidy levels and city level market shares

Machinery type (m)	**Tractors**	**Rotary tiller**	**Seeder and trans-planter**	**Grain and rape har-vester**	**Corn har-vester**	**Grain dryer**
Number of sizes (J)	41	9	30	24	11	10
Number of cities (C)	299	287	290	294	155	273
Subsidy (1000 yuan)						
Mean	24.4	2.3	5.5	13	31.8	29
SD	28.8	3.3	10.1	14.4	26.6	35
Interquartile range	33.5	2.2	4.5	20.1	39.9	37.1
Purchased quantity						
Unit (in 1000)	Horsepower	Width (m)	Rows	Speed (kg/s)	Rows	Capacity (ton/day)
Mean	77.2	1024.5	2.1	1.2	0.8	0.7
SD	120.3	1521.2	4.4	2.1	1.4	1.4
Interquartile range	92	1202.3	2	1.2	0.9	0.8
1/Herfindal index of market shares						
Base year share	460.8	166.8	253	199.4	138	108.3
Lagged share	421.9	182.2	278.1	155.2	128.7	155.6
Largest province average market share(%)						
Base year share	0.6	1.2	0.9	1.1	1.3	1.4
Lagged share	1.1	1.2	1	1.2	1.6	1.4

Notes: Based on machinery purchase and subsidy data in 24 provinces in years 2015 to 2018.

5 Results

5.1 Local employment outcomes

Table 4 presents 2SLS estimates for the effect of subsidized tractor purchase on the number of non-agricultural employees by sectors in the same city. Column 1 shows the result for the aggregated employment in all non-agricultural sectors. According to the National Statistics Bureau Rural Migrant Workers Monitoring Survey Report, in 2015 the top 5 sectors that rural labor enter when they leave agriculture include manufacturing, construction, wholesale/retail, hotel/food service and logistics. Columns 2-6 show the results for employment in these 5 sectors, respectively. To focus on cities that are directly affected by local agricultural mechanization, I dropped 24 highly urbanized cities with little agriculture and huge migration inflow.[9]

A 1000 increase in tractor horsepower in the city decreases non-agricultural employment by 72 people, although it is not statistically different from 0 with a cluster-robust standard error of 0.378. From Table 3 we know that that on average a city experiences 77.2 thousand increase in horsepower annually during the study period. The 95% confidence interval of the estimated effect at the mean is between a reduction of 64.8 thousand people and an increase of 53.4 thousand people, which is wide compared to the average non-agricultural employment of 372.3 thousand in each city.

Among the top 5 sectors, the wholesale and retail sectors is most likely to have experienced a positive employment effect and the logistics sector is most likely to have experienced a negative employment effect. Again, none of the estimates are statistically significant at the 95% confidence level.

[9] Excluded cities are Shanghai, Beijing, Guangzhou, Shenzhen, Chengdu,Chongqin, Hangzhou, Wuhan, Suzhou, Xi'an, Nanjing, Changsha, Tianjin, Zhengzhou, Dongguan, Qingdao, Kunming, Ningbo, Hefei, Foshan, Shenyang, Wuxi, Xiamen, Dalian

Table 4: 2SLS estimation results for the effect of mechanization on employment by sector

	Dependent variable: Employment by sector (1000 people)					
	All non-ag sectors	Manufacturing	Construction	Wholesale and retail	Hotel and food service	Logistics
	(1)	(2)	(3)	(4)	(5)	(6)
Tractor purchase (1000 horsepower)	−0.072	0.178	0.004	0.031	−0.050	−0.075
	(0.378)	(0.168)	(0.075)	(0.021)	(0.061)	(0.077)
Mean employment (1000 people)	372.323	91.154	61.990	17.532	4.589	14.323
Effect at the mean	−5.725	14.218	0.337	2.451	−3.977	−5.988
	[−64.868, 53.419]	[−12.041, 40.476]	[−11.352, 12.026]	[−0.846, 5.749]	[−13.596, 5.643]	[−18.064, 6.088]
Montiel-Pflueger robust F stat	34.567	34.567	34.567	34.567	34.567	34.567
Observations	643	643	643	643	643	643

Notes: This table reports the two-stage least squares estimates on the effect of subsidized tractor purchase on employment in all secondary and tertiary sectors and the 5 sectors with the largest rural workers representation with city level data in year 2015 - 2019 from China. 24 Highly urbanized cities with little agriculture and large migration inflow are excluded. The shift-share instrumental variable measures the city's exposure to the machinery subsidy policy depending on the market shares in the previous year. The endogenous variable is the subsidized machinery purchase. The reported standard errors in the parenthesis are heteroskedasticity-robust and are clustered at the city level. The city fixed effects are included by first-differencing the variables and the shocks, as suggested by Borusyak et al. (2022). The year fixed effects are included as dummy variables. *p<0.1; **p<0.05; ***p<0.01.

The standard errors reported in Table 4 are clustered at the city level, which may not be most appropriate for the shift-share IV design. Adjusting for clustering is necessary when the sample is cluster-sampled or the treatment assignment is clustered (Abadie et al., 2023). The latter case applies to the current study design. There are three potential ways that the shift-share instrumental variable is clustered: The first source of clustering is addressed in Adão et al.(Adão et al., 2019). Cities in a same province with similar market shares are exposed to the same subsidies, and tend to have similar residual values. Second, there may be correlation between subsidy levels across different machinery sizes. Third, there may be serial correlation and correlation across provinces in subsidy levels. To examine the pattern of clustering in the subsidy, I calculated the inter-class correlation coefficient (ICC) of the subsidy levels in Table A1 in Appendix B. The numbers in row 1 shows that subsidy levels are loosely correlated within provinces. The numbers in row 3 shows that within the provinces-size cells, the subsidy levels are highly correlated across years.

Shock-level regressions provide a natural way to produce inference result that accounts for the inherent non-standard clustering in shift-share IV designs (Borusyak et al., 2022). Using the the shock-level variables and weights defined in Appendix C, I estimate the 2SLS regression and adjust the standard errors for the province-size level clustering. The point estimates in Appendix C Table A2 are close to the benchmark results in Table 4. For the estimated effect on overall non-agricultural employment, the standard errors is reduced by more than a half, yet it is still statistically insignificant.

5.1.1 Robustness: county level regressions

The 2SLS regression is repeated using county level mechanization and employment data. With county level data, the consistency of the 2SLS estimator is better guaranteed with larger number of observations and a stronger first-stage. The downside of using more disaggregated data is that it captures structural change from a more narrow scope. In line with city level results, column 1 of Table 5 shows no evidence that mechanization changes non-agricultural employment in the local economy.

Table 5: 2SLS estimation results for the effect of mechanization on employment by machinery types - county level

	Dependent variable: Non-agricultural employment (1000 people)					
	Tractor Horsepower	Rotary tiller Width (m)	Seeder,transplanter Rows	Grain harvester Speed (kg/s)	Corn harvester Rows	Grain dryer Capacity (ton/day)
	(1)	(2)	(3)	(4)	(5)	(6)
Machinery purchase (in 1000)	−0.221	−0.451	−118.535	23.677	199.957	−127.237
	(0.364)	(0.326)	(125.750)	(49.286)	(319.741)	(249.332)
F stat for weak instrument	78.379	2.586	1.804	15.024	2.614	0.7056
Observations	3,964	3,263	2,808	3,314	1,707	1,797
Adjusted R^2	0.0003	-10.339	-5.833	-0.030	-1.556	-2.171
Residual Std. Error	35.092	124.640	98.500	37.009	51.236	76.175

Notes: This table reports the two-stage least squares estimates on the effect of subsidized machinery purchase on employment in all secondary and tertiary sectors with county level data in year 2015 - 2019 from China. The shift-share instrumental variable measures the county's exposure to the machinery subsidy policy depending on the market shares in the previous year. The endogenous variable is the subsidized machinery purchase. The reported standard errors in the parenthesis are heteroskedasticity-robust and are clustered at the county level. The fixed effects are included by first-differencing the variables and the shocks, as suggested by Borusyak et al. (2022). *p<0.1; **p<0.05; ***p<0.01.

5.1.2 Robustness: the Andrews and Armstrong (2017) estimator

The first stage F statistic in this study ranges from 19.4 to 78.4, depending on the level of data aggregation. The range falls below the recently established standard of 104.7 proposed by Lee et al. (2022) for detecting weak instrument. Moreover,the city level regression results are based on 643 observations, which is not a particularly large sample by recent standards.

To test the robustness of the benchmark results to potential threats from small sample and weak instrument, I report the Andrews and Armstrong (2017)'s unbiased estimator which out-performs the 2SLS estimator when the instrument is weak regardless of the sample size. The basis for using the Andrews and Armstrong (2017) unbiased estimator is that machinery subsidy can not discourage mechanization so that the sign of the first-stage coefficient is unambiguously positive.

The results reportd in Table {shock_level_u} are similar to those in Table 4 and are more precisely estimated. It further confirms the null effect of mechanization on non-agricultural employment in the local area.

Table 6: Shock-level Andrews and Armstrong (2017) Unbiased Estimator results

	Dependent variable: Employment by sector (1000 people)					
	All non-ag sectors	Manufacturing	Construction	Wholesale and retail	Hotel and food service	Logistics
	(1)	(2)	(3)	(4)	(5)	(6)
Tractor purchase (1000 horsepower)	-0.0646	0.128	0.0195	0.0252	-0.0529	-0.0725^{***}
	(0.164)	(0.201)	(0.068)	(0.016)	(0.034)	(0.024)
Observations	1125	1125	1125	1125	1125	1125

Notes: This table reports the Andrews and Armstrong (2017) unbiased estimator with variables re-constructed to the shock-level. The city fixed effects are included by first-differencing the variables and the year fixed effects are inluded by re-centering at the year mean. The standard errors are robust to clustering at the province-size level.

5.2 Migration destination employment outcomes

Migration is an important mitigation strategy during occupational change. In China, only 20.6% change in residence are within counties. 12.8% of migration happens across counties and within cities, 25% across cities and within province and 41.6% across provinces (Duan et al., 2019). Without accounting for migration, structural change is under-estimated.

To capture the potential effect of mechanization on occupational changes accompanied by migration outside the city, I use the 2017 China Migrants Dynamic Survey data to match migration destination cities with the top 5 origin cities of migrants. The survey was conducted using the PPS method to sample district and then communities in 32 province level units in China. Within each sampled communities, 50 migrant households or 100 migrants are sampled at random. For each of the 330 cities where migrants are sampled, the top 5 origins are identified based on the number of sampled migrants, which is the proxy for the migration flow. The destination city's exposure to agricultural mechanization is calculated as the weighted average of tractor purchase in the top 5 origins, with migration flow as weights. Similarly, the destination city's exposure to the mechanization subsidy is the weighted average of the shift-share instrument for each of the 5 origin cities.

Table 7 reports the results for non-agricultural employment in migration destination cities. Again, column 1 shows the result for the aggregated employment in all non-agricultural sectors and columns 2-6 show the results for the 5 sectors that absorb the most rural labor. Echoing the results reported in the last section, agricultural mechanization in the migration origin cities does not have statistically significant positive effect on overall non-agricultural employment in destination cities. Looking into specific sectors, the manufacturing sectors employs 501 more people when the origin cities on average increase tractor power by 1000 horsepower. The employment in the wholesale/retail sectors also increases by 81 person. It is worthy to note that coefficients for the construction and logistics sectors are negative in both the local regressions and the migrant destination regressions, although they are not statically significant. Labor from these sectors is equipped with skills of driving and operating machines. It is not surprising workers who have experience in either of the two sectors are most likely to be attracted by the new opportunities in agriculture opened up by mechanization.

Table 7: 2SLS estimation results for the effect of mechanization on employment by sector in migration destination cities

	Dependent variable: Employment by sector in the migtation destination city (1000 people)					
	All non-ag sectors	Manufacturing	Construction	Wholesale and retail	Hotel and food service	Logistics
	(1)	(2)	(3)	(4)	(5)	(6)
Weighted average of tractor purchase (1000 horsepower) in the top 5 migration origin cities	0.220	0.501***	−0.088	0.083**	0.028	−0.132
	(0.339)	(0.185)	(0.105)	(0.042)	(0.017)	(0.088)
Montiel-Pflueger robust F stat	22.309	22.309	22.309	22.309	22.309	22.309
Observations	414	414	414	414	414	414

Notes: This table reports the two-stage least squares estimates on the effect of subsidized machinery purchase in the top 5 migration origin cities on the secondary and tertiary employment in the corresponding migration destination cities during 2016 - 2019 in China. The average machinery purchase in the top 5 origin cities is weighted using the migration flow. The reported standard errors in the parenthesis are heteroskedasticity-robust. Since time-variant lagged shares are used to construct the shift-share instrumental variable, the destination city fixed effects are included by first-differencing the variables and the shocks, as suggested by Borusyak et al. (2022). *p<0.1; **p<0.05; ***p<0.01.

6 Unpacking the null effect result with individual level data analysis

There are two potential explanations that the above analysis does not detect a significant effect of mechanization on non-agricultural employment. First, even if mechanization is labor-saving, the labor "push" effect may be weak because the mechanization lags behind the existing outflow of labor due to adoption barriers. In fact, when the supply of agricultural labor is low, subsidized machines increase the marginal product of labor and thus the demand for labor in agriculture, as shown in Section 2. Second, administrative employment data in China does not capture changes in the buffering zone between agricultural and non-agricultural employment, which includes unemployment, informal sector employment[10] and the choice to stay at home and take care of children and elderly.

To explore the change in individual occupations and potential heterogeneity across gender, I look into individual employment outcomes using the 2015 and 2018 China Health and Retirement Longitudinal Survey data. The survey asks respondents about their occupations, whether it is in the formal or informal sector. This survey data is the best fit for the analysis because the survey has a well-designed method to track respondents when they change residence, thus minimizing bias introduced by sample attrition related to occupational change. I limit the sample to the subset of households with access to farmland due to their village membership and individuals below 70 years old. Due to the sampling design, most respondents are above 45 years old. It is widely known in China that few rural young people consider farming as a career. Thus, the current analysis only applies to the rural population that would consider working in agriculture at the margin. The survey data is matched to the city level mechanization data by the household residence.

[10]Informal sector employment is estimated to account for around 31.3% - 32.6% of total urban employment, which is 40% the size of migrant labor force (Chen et al., 2021).

In a linear probability model, a group of dummy variables representing employment outcomes are regressed on the annual tractor purchase per unit of farmland in the residing city, while controlling for city and year fixed effects to absorb city and year-specific unobservable factors. In addition, to show heterogeneous effect across demographic groups, the tractor purchase is interacted with a categorical variable indicating four demographic groups: working-age (< 61 years old) men, working-age women, elderly men (above 60 years old) and elderly women.

The top row in Table 8 shows the effect of mechanization on employment for the default group working-age men. The lower panel shows the average marginal effect by demographic groups, respectively. Column 1 of Table 8 shows that an increase in 1 tractor horsepower per ha in the residing city is associated with a 15.3% increase in the probability of working in agriculture for working-age men, mainly due to increased probability (16.3%) of working at their own farm. In the meantime, there is no drop in the probability of working in non-agriculture jobs, implying that the returning male labor are those without an advantage in the non-agricultural job market or those who are seasonally employed in a non-agricultural job[11]. For working-age women, it is associated with a 19% decrease in the probability of working in non-agricultural jobs, but not an increase in the probability of working in agriculture. In addition, elderly women are less likely to engage in farming activities in more mechanized areas.

The above results shows that mechanization has different implications for rural men and women. The aforementioned first explanation applies to men as machines augment their productivity in agriculture and attract them back to the farms. The second explanation applies to women. Their reduced participation in the labor market reflects two aspects of the underlying dynamics within households: 1. Enhanced division of tasks based on gender,

[11] Most respondents have passed their prime age as manual labor and have lower education levels compared to the younger generation.

with men specialize in farm production and women specialize in domestic chores; 2. Increased consumption of leisure due to higher farm productivity and household income. This is consistent with Afridi et al. (2023)'s finding in India, where mechanized tilling led to a fall in women's farm labor use without an increase in their non-farm sector employment. The gendered impact of technological change stems from social norms as well as land tenure institutions that exclude women.

The individual analysis shows that it is unlikely that mechanization has a positive effect on non-agricultural employment in the context of China. There may be a negative effect through reduced women's participation in the non-agricultural job that is hard to detect with noisy and incomplete administrative employment data. This finding seems to be at odds with Caunedo et al.(2021)'s finding that machinery custom service vouchers reduces worker supervision needs and non-agricultural income in the state of Karnataka in India, but it is important to note that hired agricultural labor is less common and mostly seasonal in China. Although the finding here contradicts those in recent studies in China Zheng et al. (2022), the shift-share instrumental variable study design and the panel data lend more credibility to the current study.

Table 8: Linar probability model regression results for individual employment status

	Dependent variable: Binary variables for			
	Working in agriculture	Working in agriculture, hired	Working on own farm	Working in non-agriculture jobs
	(1)	(2)	(3)	(4)
Tractor purchase per ha of agricultural land (horsepower)	0.153***	−0.051	0.163***	0.086
	(0.051)	(0.036)	(0.051)	(0.094)
Tractor purchase × elderly women	−0.416***	−0.003	−0.390***	−0.046
	(0.125)	(0.058)	(0.125)	(0.097)
Tractor purchase × elderly men	−0.083	−0.017	−0.097	−0.011
	(0.093)	(0.051)	(0.094)	(0.092)
Tractor purchase × working age women	−0.156*	0.029	−0.171*	−0.276**
	(0.089)	(0.047)	(0.095)	(0.132)
Elderly women	0.010	−0.022	0.004	−0.499***
	(0.030)	(0.014)	(0.030)	(0.028)
Elderly men	0.039	0.001	0.036	−0.329***
	(0.028)	(0.013)	(0.027)	(0.029)
Working age women	0.009	−0.004	0.008	−0.258***
	(0.023)	(0.009)	(0.023)	(0.037)
Average marginal effect				
Working-age men	0.153***	−0.051	0.163***	0.086
	(0.051)	(0.036)	(0.052)	(0.094)
Elderly women	−0.263**	−0.054	−0.228*	0.039
	(0.115)	(0.052)	(0.124)	(0.062)
Elderly men	0.071	−0.069	0.066	0.074
	(0.090)	(0.044)	(0.096)	(0.081)
Working-age women	−0.003	−0.022	−0.009	−0.190**
	(0.078)	(0.058)	(0.091)	(0.089)
Observations	10,782	10,782	10,781	10,781
R^2	0.064	0.037	0.061	0.194
Adjusted R^2	0.055	0.028	0.053	0.186
Residual Std. Error	0.454 (df = 10682)	0.255 (df = 10682)	0.462 (df = 10681)	0.426 (df = 10681)

Notes: City fixed effect and year fixed effect are included in all four regressions. The reported standard errors in the parenthesis are heteroskedasticity-robust and are clustered at the city level. The individual employment status information is from the CHARLS survey waves 2015 and 2018. The observations are from a sample subset of rural households with access to farm land and individuals below 70 years old. Individuals no more than 60 years old are classified as working age. Individuals above 60 years old are classified as elderly. *p<0.1; **p<0.05; ***p<0.01.

7 Conclusion

This study provides empirical evidence on the impact of agricultural mechanization on structural transformation using a shift-share instrumental variable design based on the machinery purchase subsidy policy in China. I do not find evidence that agricultural mechanization increases non-agricultural employment. The finding is robust to various scopes of data and alternative estimation methods. Moreover, individual level analysis indicates that mechanization has gendered effects. It enhances working-age men's role in the family farm and reduces working-age women's participation in non-agricultural jobs. Therefore, agricultural mechanization may have a negative effect on non-agricultural employment that is obscured by the limitations of administrative employment data.

This study suggests that policymakers who are concerned about the employment effect of agricultural mechanization need to look both within and above agriculture. First, agricultural mechanization promotion policies may not significantly increase labor supply to the secondary and tertiary sectors, even in labor-abundant scenarios. Instead, it can help to overcome barriers to adoption, increase labor productivity and help to rejuvenate the rural economy. Second, employment effects on non-agricultural sectors can travel beyond the local area and emerge in a different form in migration destinations with different economic structures. Therefore, cross-regional information is needed to evaluate the full consequence of mechanization. Lastly, within agriculture, mechanization complements the skills, resources and the conventional roles of men, but not women. This provides additional evidence that the "feminization" of agriculture during structural transformation is a myth(de Brauw et al., 2008; Kawarazuka et al., 2022). Measures need to be taken to make technological progress more inclusive for women, such as providing them with training, improving their access to credit and empowering them with visions of the future of agriculture.

A major limitation of this study is that of external validity. China is a special context with high speed growth in the secondary and tertiary sectors and relatively abundant job opportunities outside of agriculture. Thus, the force of "labor pull" is strong enough to nullify the force of "labor push." The policy implication of this study does not readily transfer to countries with stagnant economic growth. Second, due to the limitation of data, I cannot directly test the relationship of agricultural mechanization with unemployment, informal sector employment and intra-household time allocation. Future studies can further investigate these intriguing topics with a gender perspective. A follow-up study can also look into the link between mechanization, cropland abandonment and agricultural production.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When Should You Adjust Standard Errors for Clustering?*. *The Quarterly Journal of Economics*, *138*(1), 1–35. https://doi.org/10.1093/qje/qjac038

Adão, R., Kolesár, M., & Morales, E. (2019). Shift-Share Designs: Theory and Inference*. *The Quarterly Journal of Economics*, *134*(4), 1949–2010. https://doi.org/10.1093/qje/qjz025

Afridi, F., Bishnu, M., & Mahajan, K. (2023). Gender and mechanization: Evidence from Indian agriculture. *American Journal of Agricultural Economics*, *105*(1), 52–75. https://doi.org/10.1111/ajae.12315

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough *. *The Quarterly Journal of Economics*, *128*(2), 469–530. https://doi.org/10.1093/qje/qjt005

Alvarez-Cuadrado, F., & Poschke, M. (2011). Structural Change Out of Agriculture: Labor Push versus Labor Pull. *American Economic Journal: Macroeconomics*, *3*(3), 127–158. https://doi.org/10.1257/mac.3.3.127

Andrews, I., & Armstrong, T. B. (2017). Unbiased instrumental variables estimation under known first-stage sign. *Quantitative Economics*, *8*(2), 479–503. https://doi.org/10.3982/QE700

Autor, D. H., Dorn, D., & Hanson, G. H. (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, *103*(6), 2121–2168. https://doi.org/10.1257/AER.103.6.2121

Autor, D. H., Katz, L. F., & Krueger, A. B. (n.d.). Computing inequality: Have computers changed the labor market? *Quarterly Journal of Economics*.

Autor, D., Mindell, D. A., & Reynolds, E. B. (2019). *The work of the future: Shaping technology and institutions* (pp. 20–21). MIT.

Autor, & Salomons, A. (2018). *Is automation labor-displacing? Productivity* *growth, employment and the labor share*. NBER.

Beaman, L., & Dillon, A. (2018). Diffusion of agricultural information within social networks: Evidence on gender inequalities from Mali. *Journal of Development Economics*, *133*, 147–161. https://doi.org/10.1016/j.jdeveco.2018.01.009

BenYishay, A., Jones, M., Kondylis, F., & Mobarak, A. M. (2020). Gender gaps in technology diffusion. *Journal of Development Economics*, *143*, 102380. https://doi.org/10.1016/j.jdeveco.2019.102380

Binswanger, H. (1986). Agricultural mechanization: A comparative historical perspective. *World Bank Research Observer*, *1*(1), 27–56. https://doi.org/10.1093/wbro/1.1.27

Borusyak, K., Hull, P., & Jaravel, X. (2022). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies*, *89*(1), 181–213. https://doi.org/10.1093/restud/rdab030

Bustos, P., Caprettini, B., & Ponticelli, J. (2016). Agricultural Productivity and Structural Transformation: Evidence from Brazil. *American Economic Review*, *106*(6), 1320–1365. https://doi.org/10.1257/aer.20131061

Caunedo, J., & Keller, E. (2021). Capital Obsolescence and Agricultural Productivity. *Quarterly Journal of Economics*, *136*(1), 505–561. https://doi.org/10.1093/qje/qjaa028

Chen, M., Huang, X., Huang, G., & Yang, Y. (2021). New urbanization and informal employment: Scale, pattern, and social integration. *Progress in Geography*, *40*(1), 50–60. https://doi.org/10.18306/DLKXJZ.2021.01.005

Daum, T., & Birner, R. (2020). Agricultural mechanization in Africa: Myths, realities and an emerging research agenda. *Global Food Security*, *26*(November 2019), 100393. https://doi.org/10.1016/j.gfs.2020.100393

de Brauw, A., Li, Q., Liu, C., Rozelle, S., & Zhang, L. (2008). Feminization of Agriculture in China? Myths Surrounding Women's Participation in Farming. *The China Quarterly*, *194*, 327–348. https://www.jstor.org/

stable/20192200

Diao, X., Takeshima, H., & Zhang, X. (2020). *An Evolving Paradigm of Agricultural Mechanization Development - How much can Africa learn from Asia?*

Duan, C., Lv Lidan, & Cheng, M. (2019). *(2018)* (pp. 11–11). NBS / UNFPA / UNICEF Joint Data Project.

Fredriksson, P. G., & Gupta, S. K. (2023). Irrigation and gender roles. *Journal of Development Economics*, *163*, 103076. https://doi.org/10.1016/j.jdeveco.2023.103076

Gallardo, R. K., & Sauer, J. (2018). Adoption of Labor-Saving Technologies in Agriculture. *Annual Review of Resource Economics*, *10*(Volume 10, 2018), 185–206. https://doi.org/10.1146/annurev-resource-100517-023018

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, *110*(8), 2586–2624. https://doi.org/10.1257/AER.20181047

Hamilton, S. F., Richards, T. J., Shafran, A. P., & Vasilaky, K. N. (2022). Farm labor productivity and the impact of mechanization. *American Journal of Agricultural Economics*, *104*(4), 1435–1459. https://doi.org/10.1111/ajae.12273

Hayami, Y., & Ruttan, V. (1971). *Agricultural development: An international perspective.* Baltimore, Md/London: The Johns Hopkins Press.

Kawarazuka, N., Doss, C. R., Farnworth, C. R., & Pyburn, R. (2022). Myths about the feminization of agriculture: Implications for global food security. *Global Food Security*, *33*, 100611. https://doi.org/10.1016/j.gfs.2022.100611

Lee, D. S., McCrary, J., Moreira, M. J., & Porter, J. (2022). Valid t-ratio Inference for IV. *American Economic Review*, *112*(10), 3260–3290. https://doi.org/10.1257/aer.20211063

Ma, W., Zhou, X., Boansi, D., Horlu, G. S. A., & Owusu, V. (2024). Adoption and intensity of agricultural mechanization and their impact on non-farm employment of rural women. *World Development*, *173*, 106434. https://doi.org/10.1016/j.worlddev.2023.106434

Millimet, D., & Bellemare, M. F. (2023). *Fixed effects and causal inference*.

Olea, J. L. M., & Pflueger, C. (2013). A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics*, *31*(3), 358–369. https://doi.org/10.1080/00401706.2013.806694

Pingali, P. (2007). Chapter 54 Agricultural Mechanization: Adoption Patterns and Economic Impact. In *Handbook of Agricultural Economics* (Vol. 3, pp. 2779–2805). Elsevier. https://doi.org/10.1016/S1574-0072(06)03054-4

Pingali, P., Bigot, Y., & Binswanger, H. P. (1987). Agricultural Mechanization and the Evolution of Farming Systems in Sub-Saharan Africa. In *Economic Geography* (Vol. 64). https://doi.org/10.2307/144242

Pitt, M. M., Rosenzweig, M. R., & Hassan, M. N. (2012). Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy. *American Economic Review*, *102*(7), 3531–3560. https://doi.org/10.1257/aer.102.7.3531

Qian, N. (2008). Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance. *The Quarterly Journal of Economics*, *123*(3), 1251–1285. https://www.jstor.org/stable/25098928

Yang, J., Huang, Z., Zhang, X., Reardon, T., Yang, J., Huang, Z., Zhang, X., & Reardon, T. (2013). *The Rapid Rise of Cross-Regional Agricultural Mechanization Services in China*. *95*(5), 1245–1251. https://doi.org/10.1093/ajae/a

Yue, X. (2005). Problems of current employment statistics in China. *Economic Research Journal*.

Zheng, H., Ma, W., Guo, Y., & Zhou, X. (2022). Interactive relationship between non-farm employment and mechanization service expenditure in rural China. *China Agricultural Economic Review*, *14*(1), 84–105. https:

//doi.org/10.1108/CAER-10-2020-0251

Zou, B., Chen, Y., Mishra, A. K., & Hirsch, S. (2024). Agricultural mechanization and the performance of the local Chinese economy. *Food Policy*, *125*, 102648. https://doi.org/10.1016/j.foodpol.2024.102648

Appendices

Appendix A: Machinery subsidy and purchase data compilation

The machinery subsidy and purchase data used in this research is collected from government websites. This appendix describes the process to compile downloaded raw data to organized data files, as well as the process to construct the county and city level mechanization measurements, market shares, local subsidy levels and the shift-share instruments.

Machinery subsidy catalogs are from provincial government announcements. These announcements are made in Spring annually and may be followed up with additions and minor adjustments later in the same year. All announcements are collected and compiled into one data file.

The subsidized machinery purchase records are collected from government websites. Each record entry contains information about the beneficiary's address and name, machinery type, quantity, manufacturer, purchase date, retailer, price and subsidy level. I geo-coded the beneficiary's address to the county level using an address parser and the Gaode map API in Python. The data is dated from year 2015 to 2020. Data for some provinces are not available for earlier years since these provincial governments no longer keep the earlier purchase records published online at the time of the data collection. All counties that are ever observed are colored in the following figure.

County and city level measures of mechanization are obtained by aggregating the purchase records. The measurement unit changes by machinery type to account for machinery size. For instance, tractors are measure in horsepower and seeders are measured in the number of rows they can sow in one pass. The market shares for each machinery size group are also contructed based on the size unit. For example, there are 39 size groups for tractors. From 20 to 200 horsepower, tractors are divided into size groups by horsepower intervals and category (wheeled vs. crawler, two wheels vs four wheels). The market shares of the 39 size groups are the shares of horsepower from each group of tractors purchased in a county/city in a year. The vector market shares is interacted with vector of subsidy levels of the same size groups to construct the shift-share instrumental variable.

The raw purchase records contains information about the machinery type and subsidy level, but not the machinery size. In order to label each purchase record with the machinery size information, I matched the records with the subsidy catalog in the same province in the same year by machinery type and subsidy level. In cases where multiple sizes are assigned with the same subsidy level, I take the average of the matched sizes.

Appendix B: Additional descriptive statistics

Table A1: Inter-Class Correlation of Machinery Subsidy Rates

Machinery type (m)	**Tractors**	**Rotary tiller**	**Seeder and trans-planter**	**Grain and rape har-vester**	**Corn har-vester**	**Grain dryer**
Province						
ICC	0.141	0.048	0.030	0.192	0.221	0.077
SE	0.039	0.017		0.042	0.070	0.017
Size						
ICC	0.626	0.823	0.827	0.551	0.542	0.771
SE	0.054	0.046		0.058	0.073	0.048

Notes: The subsidy rates are calculated by dividing subsidy levels by the national average retail in the previous year.

Appendix C: Shock level regressions

The shift-share instrumental variable estimator is equivalent to the second-stage coefficient from a shock-weighted shock-level regression (Borusyak et al., 2022). In the case of this study, the shock (subsidy) varies at the province-year-machinery size level. Therefore, $subsidy_{jpt}^{m}$ by itself can be used directly as the instrument for a shock-level mechanization variable that is constructed from aggregating machinery purchase across cities within provinces using market shares as weights. Specifically,

$$mechanization_{jpt}^{m} = \frac{\sum_{c=1}^{C_p} share_{jpc,t-1}^{m} mechanization_{pct}^{m}}{\sum_{c=1}^{C_p} share_{jpc,t-1}^{m}} \tag{21}$$

The shock-level outcome variable is also constructed from transforming the city-year level employment outcomes in the same way. The regression weight for each province-year-machinery size level observation is $share_{jp,t-1}^{m} = \frac{1}{C_p} \sum_{c=1}^{C_p} share_{jpc,t-1}^{m}$, which represents the average market share of machinery size j within the machinery category m among the cities in province p in year $t-1$.

The 2SLS regression results with shock-level variable are reported below.

Table A2: Shock-level 2SLS estimation results

	Dependent variable: Employment by sector (1000 people)					
	All non-ag sectors	Manufacturing	Construction	Wholesale and retail	Hotel and food service	Logistics
	(1)	(2)	(3)	(4)	(5)	(6)
Tractor purchase (1000 horsepower)	−0.0676	0.130	0.0182	0.0256	−0.0537	-0.0737^{***}
	(0.168)	(0.204)	(0.069)	(0.016)	(0.037)	(0.025)
Observations	1125	1125	1125	1125	1125	1125
First stage F statistic	19.42	19.42	19.42	19.42	19.42	19.42

Notes: This table reports the two-stage least squares estimates with variables reconstructed to the shock-level. The city fixed effects are included by first-differencing the variables and the year fixed effects are inluded by re-centering at the year mean. The standard errors are robust to clustering at the province-size level. Montiel-Pflueger robust F statistic in reported in the last row.