Agricultural Mechanization and Structural Transformation in China

Ling Yao, University of Minnesota, yao00057@umn.edu

Selected Paper prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC; July 23-25, 2023

Agricultural Mechanization and Structural Transformation in China

Ling Yao*

January 2023

Abstract

Mechanization is an essential step toward agricultural modernization, especially in labor-abundant developing countries. Whether agricultural mechanization promotes structural transformation, however, remains an empirical question. This paper measures the effect of subsidized agricultural mechanization on non-agricultural employment. Using a unique machinery subsidy and purchase dataset from China, I construct local exposure to common subsidy policy as a shift-share instrument for mechanization. Two-stage least squares estimates at different administrative levels show no statistically significant change in the growth of formal sector non-agricultural employment due to subsidized mechanization within the same administrative unit. Moreover, rural individuals, especially men, are more likely to stay in agriculture in areas with faster mechanization. The results imply that subsidized mechanization plays a limited role in perpetuating structural transformation. Instead, these findings suggest that mechanization acts as a complement to the shrinking labor force in agricultural production.

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

---

*University of Minnesota, yao00057@umn.edu.

[1]I am grateful to Marc Bellemare, Terry Hurley, Metin Çakır, David Samuels, 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, and participants at the UMN Trade and Development Seminar for their helpful suggestions and comments. All errors and omissions are mine.

1 Introduction

In the past few decades, Asian economies have undergone substantial industrialization, driving a rising wage and a growing demand for agricultural mechanization.[1] While the adoption of machinery is well-explained by the theory of induced innovation (Hayami & Ruttan, 1971), the interaction between agricultural mechanization and non-agricultural labor markets remains understudied.

On the one hand, given the speed of structural transformation,[2] the speed of agricultural mechanization may be less than optimal due to various market constraints, including supply side factors, such as the lack of equipments, spare parts and skilled mechanics, as well as demand side factors, such as unfavorable relative factor prices and market access conditions (Prabhu Pingali, 2007). The use of machinery is labor-saving technology and farmers' adoption decision is subject to risk preferences, the indivisibility of the technology, the cost of learning and credit market imperfections (Just & Zilberman, 1988). In the case of China, farmland allocation is highly fragmented, limiting the economies of scale of machinery operation (Wang, Yamauchi, Huang, & Rozelle, 2020). Rural land ownership is collective and farmers cannot use land as a collateral. Without proper access to credit, farmers shun lump-sum investments such as large agricultural machinery purchase (Kumar, Turvey, & Kropp, 2013; Lohmar, Gale, Tuan, & Hansen, 2009). Under insufficient machanization, migrating farmers are left with the dilemma to choose between abandoning their cropland and turning down profitable non-agricultural job opportunities to stay in farming, neither are welfare

[1] See, for example, China (Wang, Yamauchi, & Huang, 2016; Wang, Yamauchi, Otsuka, & Huang, 2016), India (Afridi, Bishnu, & Mahajan, 2020), Bangladesh (Hassan & Kornher, 2019), Vietnam (Caunedo & Keller, 2021), Indonesia (Yamauchi, 2016) and Nepal (Takeshima, Prasad, Mahendra, Poudel, & Kumar, 2015).

[2] Here I follow the definition of structural transformation as the transition of an economy from low productivity economic activities to high productivity activities, which is characterized by the movement of labor from agricultural sector to non-agricultural sectors.

optimizing. In an effort to sustain domestic food production and promote agricultural modernization, the Chinese government started to subsidize machinery purchase in 2004. The expansion of custom machinery services (Yang et al., 2013) , along with the development of land rental institutions (Ito, Bao, & Ni, 2016), has also contributed to substantial mechanization growth. The total machinery power nearly doubled in 15 years, going from 640.3 million KW in 2004 to 1,027 million KW in 2019.[3]

On the other hand, machinery purchase subsidies work by reducing the relative price of capital, ceteris paribus, naturally raising the concern of labor displacement. Studies of automation in general have explored consequences beyond the capital-labor substitution. Autor (2019) pointed out that automation also complements cognitive capabilities of workers and create new employment opportunities. Moreover, whether technological progress harms overall employment depends on both the direct effect within the sector and the indirect effect on other sectors (D. Autor & Salomons, 2018). To which extent has mechanization released the constraints in farming households' labor allocation decisions to facilitate structural transformation and whether it has gone forward to displace labor depend on the skills of farm labor, the nature of the adopted machinery types and the labor demand non-agricultural sectors.

This paper studies the impact of agricultural mechanization on structural transformation. Specifically, I test whether agricultural mechanization speed up China's structural transformation by increasing employment in non-agricultural sectors in China.

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 (county or prefecture city).[4] By

[3]Source: China National Bureau of Statistics.

[4]In the Chinese administrative division, provinces are one level higher than prefecture cities, and prefecture cities are one level higher than counties.

interacting the common shock of machinery subsidy with local machinery use patterns, I construct local exposure to the shock as a shift-share instrumental variable for mechanization. In terms of Borusyak, Hull and Jaravel (2020)'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 local and period fixed effects.

To precisely measure the local mechanization level, I assemble a unique dataset of universal subsidized machinery purchase records. The purchase records are aggregated to the county and prefecture city levels by machinery type. Then I construct the shift-share instrumental variables by combining lagged machinery purchase patterns with annual machinery subsidy schemes from provincial government publications. I use county and prefecture city statistic yearbooks data to construct labor market outcome measures.

The empirical results show that subsidized mechanization does not promote local formal employment in non-agricultural sectors, except in the construction sector. Descriptive statistics show that in the meantime, agricultural mechanization in the local area is associated with a higher probability that rural individuals stay in agriculture to work on family farm, especially for male and individuals below 60 years old. It is not correlated with the probability of working in non-agricultural jobs for the same population. These results run counter to the conventional wisdom in agricultural and development economics that agricultural mechanization will release labor from agricultural sector and to feed a growing non-agricultural sector. In this context, it is more likely that mechanization has attracted rural residents who had migrated to work in informal-sector jobs back in rural areas to work agriculture.

This study is directly linked to the farming system evolution theory (Binswanger, 1986; Prabhu Pingali, 2007; P. Pingali, Bigot, & Binswanger, 1987). Summarizing 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. In line with the theoretical predictions, existing empirical studies have found mixed evidence (Daum & Birner, 2020). While some studies found reduced or increased labor use after mechanization, others did not find statistically significant correlation. Most of these studies relied on cross-sectional data and limited identification strategies. 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). Binswanger (1986) specifically warned against the labor displacement effect of mechanization stimulated by subsidy. The sustainability and efficacy issues with implementing various versions of mechanization subsidy are well-documented in literature (Houssou et al., 2013), whereas evidence on labor market outcomes is thin.

This study complements 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. This study also differs from most related empirical studies in that mechanization is measured using regional purchase data rather than by household level machinery ownership or usage, considering that regional custom machinery service plays a significant role in smallholder-dominated agriculture (Yang et al., 2013). Most related studies look at capital-labor substitution within agriculture. I extend the view to employment outcomes in non-agricultural sectors, helping to illustrate a fuller picture of labor dynamics during agricultural mechanization. With panel data and a shift-share instrumental variable design, this study aims at estimating parameters whose interpretation is causal.

The rest of this paper is organized as follows. Section 2 describes the data source and overviews mechanization, the subsidy policy and employment structure in China, highlighting on key descriptive statistics. Section 3 discusses the identification strategy and presents evidences for evaluating the strategy. Section 4 presents the results and Section 5 concludes.

2 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 the transition in employment structure in China. Description of data collected for this study and other supportive statistics are woven into the corresponding subsections.

2.1 Agricultural machinery purchase subsidy policy in China

China's machinery purchase subsidy started in 2004 and experienced several rounds of expansion and adaptation. It has transformed from a pilot program in a small group of selected counties to become a nation-wide 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 province level governments announce a catalog of subsidized machinery items, specifying key features of eligible items and the 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 for the sake of transparency. These purchase and subsidy records are collected and aggregated to the county and prefecture city levels to measure changes in 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.[5] The subsidy catalogs are also collected for contructing 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 the 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 3.1.1, it is worthwhile to evaluate factors behind the subsidy determination process. A typical provincial

[5]There is a limit on the number of units each household can purchase with subsidy per year. For most provinces the limit is 5 and it is rarely binding. The limit for organizations such as cooperatives is higher, usually 10 or 15.

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. 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. In designing the annual catalog of subsidized machinery, provincial governments follow the guidelines from the central government, which updates every three years. The guideline specifies all the candidate machinery items in two categories, general items and special items . The general items are widely-used machinery items and must be included in provincial catalogs. The special items are elective and are proposed by groups of provincial governments that share a common need for these items. The central government also sets the maximum subsidy from the central budget for each size of each item. Provincial governments can choose sizes to be included in the catalogs but cannot invent new size specifications. Provincial government then assign the implemented subsidy level to each item and size, with the option to supplement the subsidy of 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 consider 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.

2.2 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 developing countries 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.

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.

Mechanization level is uneven across subregions. 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. 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.

The descriptive analysis in this section 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 signals including the 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 has the advantage of having 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.

2.3 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 Statistical Yearbooks (county level) 2016 - 2019*.[6] The yearbooks cover around 2100 county level units out of total 2844 county level units in China. The missing ones are urban center districts and counties in remote areas. In the yearbooks, primary industry includes agriculture, forestry, livestock and fishery. Secondary industry includes manufacturing, energy and water supply and construction.[7] 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. Although the county level regressions involve higher numbers of observations and have higher statistical power, it is more prone to the spatial spillover issues discussed in Section 4.2. I supplement the data using secondary and tertiary industry employment that is reported for prefecture, a higher administration level than the counties. The prefecture city level yearbooks cover around 290 prefecture level units out of the total 333 in China. The missing ones are minority autonomous prefectures in remote areas. This allows me to run two sets of identical regressions with mechanization measures constructed to the county and city level. Moreover, the city level data contains employment by specific sectors. I repeat the regressions with the

[6]The yearbooks report information in the previous year so the data is for years 2015 - 2018. Data for year 2019 is available but is not used for concerns about outliers due to the COVID-19 pandemic.

[7]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.

dependent variable as change in employment in manufacturing, construction, wholesale/retail, hotel/food service and logistics, respectively, as these are the main sectors that rural labor enter when they leave agriculture.

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 number 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.

In order to supplement the administrative data, I also use household survey data to look at whether agricultural mechanization leads to *individual* occupational changes. I take the rural subsample (individuals are entitled to operate farmland by household registration) from the China Health and Retirement Longitudinal Survey (CHARLS) waves 2015 and 2018. The survey uses a multi-stage method to sample individuals above 45 years old. The survey obtains information about the occupation of the main survey respondent and their spouse. Among them, for each working individual I construct a binary variable to measure whether the occupation is agriculture or non-agriculture and a binary variable to represent whether they live in a rural area. Unlike many other household level studies of the effect of mechanization, I match households with the mechanization level in the residing city instead of using household level measures of mechanization. This is suitable for the study context as the majority of farms are small and rely on machinery rental and custom service in the neighboring area to avoid the financial burden of purchasing their own machines.

3 Research design

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

$$Y_{pct} = \alpha_0^m + \alpha_1^m mechanization_{pct}^m + \alpha_2^m X_{pct} + \epsilon_{pct}^m, \tag{1}$$

where Y_{pct} is the percentage of secondary and tertiary industry employment in county c of province p in year t. X_{pct} is a set of control variables. The key explanatory variable $mechanization_{pct}^m$ is the amount of machinery power of machinery category m at disposal for agricultural production in the same county in the same year. Measuring the machinery stock level is difficult, if not impossible, since it requires knowledge about which machinery units historically purchased are still in service. To match the flow measurement of mechanization available in the data, I take the first difference of all the variables and use the subsidized machinery purchase ($purchase_{pct}$) as a proxy for the change in mechanization level ($\Delta mechanization_{pct}^m$). The first-differenced model is as below:

$$\Delta Y_{pct} = \beta_0^m + \beta_1^m \Delta mechanization_{pct}^m + \beta_2^m \Delta X_{pct}^m + \varepsilon_{pct}^m. \tag{2}$$

The extent to which subsidized machinery purchase reflects the change in machinery stock relies on two factors that can be viewed as measurement error of the key explanatory variable: the unobserved quantity of unsubsidized machinery purchase and machinery wear-out. Since the wear and tear rate of existing machinery is a physical process that is subject to machinery engineering, it can be seen as orthogonal to the subsidized mechanization and should not undermine the consistency of the estimator. The unsubsidized machinery purchase quantities, on the other hand, may be positively correlated with the subsidized mechanization as counties with higher demand for machinery would purchase more of both subsidized and unsubsidized machinery items. Under such case, the estimated effect of subsidized mechanization would be an inflated measure of the effect of mechanization in general. Given the wide range of subsized machinery types and the easy accessibility of the subsidy (as described in Section 2.1), unsubsidized mechanization should only play a minor role in the current context and the estimates in this paper still provide meaningful information about the effect of mechanization.

3.1 The shift-share instrument

A prevalent empirical challenge of the model represented by Equation (2) is reverse causality. Counties 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, Takeshima, & Zhang, 2020). Moreover, county-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. Urban sprawling may also change the nature of agriculture, e.g. from land-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 (counties). 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 counties within a province, each county is exposed to the subsidy at a different level due to the difference in the local machinery usage pattern. For example, mountainous counties 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, counties with flatter terrain benefit more when large tractors are subsidized more heavily. 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 government do not strategically design the subsidy scheme based on county-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 county-level unobservable shocks to employment structure.

The directed acylic graph (DAG) for the identification strategy is shown in Figure 7. The shift-share instrumental variable for change in mechanization level is exposure to the subsidy, which is composed of both subsidy levels and the market shares, a product in which subsidy levels determined at the province level serve as the source of identification. As formalized in the next subsection, the market shares of the machinery sizes are allowed to be correlated with the outcome in an unobserved way.

3.1.1 Validity assumption

I follow (Borusyak et al., 2020) to formulate the shift-share instrumental variable and lay out the exact identifying assumptions. The shift-share instrumental variable Z^m_{pct} for the quantity of subsidized machinery purchase ($purchase_{pct}$) is the inner product of the county level market shares ($share_{jpct}$) and the province-year level subsidy rate ($subsidy_{jpt}$) of each machinery size (indexed by j) within the category m, where the market shares used for year t are measured in the baseline year of 2015 or in the previous year, and the subsidy rate is the ratio of the subsidy level and the national average retail price in the previous year.

$$Z^m_{pct} = \sum_{j=1}^{J} share^m_{jpct} \times subsidy^m_{jpt}$$

The key identifying assumption is that Z^m_{pct} is uncorrelated with the error term in Equation (2):

$$E\left[\sum_{t=1}^{T}\sum_{c=1}^{C}\left(\sum_{j=1}^{J} share^m_{jpct} \times subsidy^m_{jpt}\right) \times \varepsilon_{pct}\right] = 0. \qquad (3)$$

As per (2020), the moment condition in Equation (3) can be re-arranged to a shock level (in the case of this study the province-year-machinery size level) equation, such that

$$E\left[\sum_{t=1}^{T}\sum_{p=1}^{P}\sum_{j=1}^{J} share^m_{jpt} \times subsidy^m_{jpt} \times \bar{\varepsilon}_{jpt}\right] = 0, \qquad (4)$$

where $share^m_{jpt} = \frac{1}{C_p}\sum_{c=1}^{C_p} share^m_{jpct}$ represents the average market share of machinery size j within the machinery category m among the counties in province p in year t. $\bar{\varepsilon}_{jpt} = \frac{\sum_{c=1}^{C_p} share^m_{jpct} \times \varepsilon_{pct}}{share^m_{jpt}}$ is a share-weighted error term summed to the province-year-machinery size level. In words, Equation (4) states that the subsidy rates should not be correlated with the county-level unobservable shocks to secondary and tertiary industry employment when they are both weighted by the machinery market shares. An important aspect of this orthogonality condition is that the market shares of machinery only serve as weights, and they are allowed to be correlated with the error terms.[8]

The moment condition in Equation (4) is satisfied if the following condition holds

$$E[subsidy^m_{jpt}|share^m_{jpt}, \bar{\varepsilon}_{jpt}] = \mu \quad \forall j, p, t. \tag{5}$$

Equation (5) is the identifying assumption. It 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. A relaxation of this assumption allows the subsidy rates to be clustered at the province-machinery size level. In practice, provinces often consistently exclude certain machinery sizes out of the subsidy scheme. 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. Fortunately since these province preference patterns are time-invariant, they can be controlled for by province fixed effects. In other words, the subsidy rates for each machinery size can be re-centered at the province historical average subsidy and have the common expectation of zero. The modified Assumption 1 is then

$$E[subsidy^m_{jpt}|share^m_{jpt}, \bar{\varepsilon}_{jpt}] = \mu_{jp} \quad \forall t. \tag{6}$$

Similarly, the expected subsidy levels may change over time in correspondence with central government periodical policy guideline changes. Therefore, I also include fixed effects for the two three-year policy planning periods 2015-2017 and 2018-2020. Instead of adding the fixed effects dummy variables in the regressions[9], I incorporate the fixed effects by residualizing the

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

[9] Adding the fixed effects dummy variables is cumbersome for two reasons: First, the

variables. In other words, I regress a variable v on a group of province and period fixed dummy variables and use the residual $\tilde{v}$ in replace of the original variable. The structural form of the final model is:

$$\widetilde{\Delta Y}_{pct} = \gamma_0^m + \gamma_1^m \widetilde{purchase}_{pct}^m + \gamma_2^m \widetilde{\Delta X}_{pct}^m + e_{pct}^m, \tag{7}$$

$$\widetilde{purchase}_{pct}^m = \delta_0^m + \delta_1^m \sum_{j=1}^{J} share_{jpct}^m \times \widetilde{subsidy}_{jpt}^m + \delta_2^m \widetilde{\Delta X}_{pct}^m + u_{pct}^m. \tag{8}$$

3.2 Falsification tests

The modified identifying Assumption in Equation (5) is crucial. It requires that the provincial governments will not adjust the subsidy levels of some counties' preferred machinery sizes in response to the underlying structural transformation trend of these counties. 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 its province historical average (represented by μ_{jp}) is random and should not be correlated with the unobservable structural transformation shocks in counties 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 county 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 84 county level administrative

regression is at the county-year level while the subsidy levels are at the machinery size-province-year level. Adding county dummies cannot properly recenter the subsidy levels at the province mean. Second, since the shift-share instrumental variable is an interaction between the subsidy and the shares, purging out the fixed effects from the subsidy levels requires interacting the fixed effects dummies with the shares before being added to the regressions. With both province fixed effects and period fixed effects, the system of interaction terms quickly becomes messy to keep track of.

units.[10] Precisely targeting the potential machinery demand caused by the underlying structural transformation trends in all the counties is challenging, if not impossible. As discussed in Section 2.1, 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[11] and when certain machinery types becomes outdated.

Nonetheless, the subsidy rates assignment process is not random, and major common structural transformation shocks that increases machinery demand may still incentivize provincial governments to adjust the subsidy scheme accordingly. To scrutinize the exclusion restriction, I conduct falsification tests that regress a county-level proxy for the unobserved confounders on the shift-share IV. The proxy is the change in the number of industrial firms in the county that reflects the change in demand for non-agricultural labor. The results are reported in Table 1. Among the six machinery types, only the shift-share instrumental variable based on grain harvester subsidy and market shares is negatively correlated with the number of industrial firms, conditional on county and period and fixed effects. This means that provincial governments might lower subsidy levels of certain grain harvester sizes when there is faster industrialization in counties that use most of these grain harvester sizes. The rationale behind such potential subsidy adjustment is unclear, but the demand for harvest machines seem to be the least responsive to subsidy incentives, as shown in the first-stage regression F-statistics reported in Appendix B Table A1. Therefore, the main regression result for harvest machines may not be as reliable as for other machinery types. No evidence for potential validity restriction violation is found for

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

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

other machinery types.

3.3 Assessing regularity conditions

Borusyak et al.,(2020) pointed out that a shift-share IV regression is equivalent to a regular instrumental variable regression at the shock assignment level. In the case of this study, the shock assignment is at the province-year-machinery size level. The variables for the equivalent shock-level regression are aggregated from the county level variables while weighted by the market shares. This point of view reveals two regularity conditions required for the consistency of the shift-share IV estimates, which are evaluated below following the procedures suggested by Borusyak et al. (2020).

3.3.1 Effective sample size

First, there should be sufficient variation in the shift-share instrumental variable. 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 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 county level. The more counties 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 county 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 counties 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 2 summarizes the market shares for each machinery type. Across the machinery types, the inverse of the Herfindahl index based on the previous year market shares varies from 120.9 to 484.3, with the corresponding number of counties varies from 1,028 to 2,295. On the last two rows of Table 2, the largest province average market share does not exceed 2%, verifying that there is no market-dominating machinery size and that the machinery market structure varies sufficiently to ensure the large sample properties of the 2SLS estimates. The Herfindahl index constructed from city level market shares reflect similar effective sample sizes (Table 3).

3.3.2 Shock independence

The statistical inference of the 2SLS estimates depends on the correlation among the shocks. Ideally, the residualized subsidy rates would be mutually uncorrelated. If they are correlated within clusters, then cluster-robust inference should be applied. Indeed, the subsidy rates for different machinery sizes in each provincial annual subsidy scheme could be correlated. Since provincial governments need to implement the policy within annual subsidy budget, they may raise the subsidy rate for some machines while lowering it for others. If the goal of the provincial government is to improve modernization of a certain agricultural production stage, they may also raise the subsidy rates for all sizes under a machinery type at the same time. The inter-class correlation analysis reported in Table 4 verifies the latter. The correlation coefficient of machinery subsidy rates within provinces varies from 0.141 to 0.328, depending on the machinery type. Within provinces, the subsidy rate is positively correlated across years for each machinery size (row 3), showing stickiness to past subsidy schemes.

4 Results

4.1 County level results

Table 5 presents county level 2SLS estimates for the effect of subsidized mechanization on the change in the number of employees in the secondary and tertiary industries. Each column is a regression for one machinery type and the model specifications are the same across columns. I incorporate the county fixed effects by re-centering the outcome variable and the endogenous variable at the county mean. The province level subsidy rates are re-centered at the province history mean before interacted with the county level market shares. The period fixed effects are incorporated by further re-centering the transformed variables at the period mean. I control for the sum of shares to account for cases where no purchase was recorded in some counties in the previous year.

The first stage F-statistics show that the subsidy offers strong machinery purchase incentive for tractors, rotary tillers, seeders, transplanters and grain dryers. The demand for harvesters is not as responsive to local exposure to the subsidy, as indicated by the low F-statistics. This is possibly due to the wide spread of cross-regional machinery harvest services (Yang et al., 2013). The subsidized machinery purchase for each machinery type is measured in different units as reported in Table 2. For instance, in column 1 of Table 5, the coefficient of -0.501 means that every 1,000 horsepower new tractor purchase would lead to a reduction of 501 employees in the secondary and tertiary industries. However, none of the estimates for the machinery purchase coefficients is statistically significant based on the cluster-robust standard errors.

4.2 Spatial spillover effects

The results in Table 5 speaks for employment structure outcome within the county without considering the fact that labor can move across county boundaries. Mechanization in one county can potentially lead to labor migration that contribute to non-agricultural employment in other counties. According to a report based on the National Population Census, only 20.6% discrepancies between household registration location and residency location are within counties, 12.8% of migration happens across counties and within prefecture cities, 25% across prefecture cities and within province, 41.6% across provinces (Duan, Lv Lidan, & Cheng, 2019). Using prefecture city level data instead of county level data can partially absorb the spillover effect. To further reduce the spillover issue caused by inter-city and inter-province migration, I exclude the provincial capital cities and five provinces that are the most popular migration destinations[12] and repeat the regressions reported in Table 5. It is important to note that without information on inter-city and inter-province migration and employment, the estimated effect is still only a portion of the underlying total effect on structural transformation. If the estimated effect is statistically significant and positive, we can safely infer that mechanization also has positive effect on non-agricultural employment accompanied by inter-city and inter-province migration.

The results for prefecture city level regressions are presented in Table 6. The shift-share instrumental variable is weaker because the city level market structure reflects a mixture of local machinery demand. Therefore, the city level market structure in the previous year is less predicative of machinery demand in the current year. For machinery types with relative high first stage F-statistics, including tractors and rotary tillers, the coefficients for machinery purchase are not statistically different from zero.

The prefecture city level data provides information on employment by specific sectors. To look more deeply into the employment structure, I repeat the regressions using change in employment in different sectors as dependent variable. Since the shift-share instrument is relatively strong only for tractors, I focus on mechanization in terms of new tractor purchase. Table 7 shows that across sectors, only employment in construction responds positively to local agricultural mechanization. For every 1000 horsepower increase in tractor, there will be 152 more employees in the construction sector. The positive relationship between employment in construction and

[12] Author's calculation based on the *China Population Census Yearbook 2020* shows that the top 5 destinations are Guangzhou, Zhejiang, Shanghai, Jiangsu and Beijing, which receive 61% of total inter-province migration population.

agricultural employment is not surprising. The type of agricultural labor that can be replaced by machinery is heavy manual labor, which is also the most needed in construction. Moreover, administrative data in construction employment is less prone to the issue of missing informal employment compared to other sectors because construction workers are required to register with a labor dispatch firm if they are not directly hired by the construction firm, making them visible by the official.[13]

5 Mechanism

The above empirical analysis suggests that agricultural mechanization has limited effect on employment in non-agricultural sector. This null effect can be explained by various mechanisms. For instance, mechanization may have indeed replaced some agricultural labor, who cannot be absorbed by formal non-agricultural sectors and end up in informal sectors jobs and unemployment. Alternatively, the agricultural wage may already be so low relative to non-agricultural wage that labor have been leaving agriculture voluntarily even in the absence of machinery. The introduction of machinery only allowed the remaining labor to keep up with agricultural production. To overcome the limitation of administrative data in drawing insight on the mechanism, I take a different angle by using the CHARLS survey data to explore individual employment choices under mechanization. Specifically, I regress different dummy variables representing employment outcomes on the annual machinery purchase per unit of farmland in the residing city, while controlling for province and year fixed effects to absorb province and year-specific unobservable factors. The regression formula is

$$D_{it} = \theta_0 + \theta_1 purchase_{ct} + \theta_2 \sigma_p + \theta_3 \tau_t + v_{it} \tag{9}$$

[13]Since year 2005, an order by the Ministry of Housing and Urban-Rural Development (*Guanyu jianli he wanshan laowu fenbao zhidu fazhan jianzhu laowu qiye de yijian*) forbids unregistered individual from organizing construction workers.

where D_{it} is the dummy variable for respondent i working in agriculture, working in non-agriculture jobs, working as hired agricultural labor, working for family farm or living in the rural area in survey year t. $purchase_{ct}$ is the subsidized tractor power purchase per hectare of farmland in the respondent's residing prefecture city during survey year t. σ_p and τ_t are province and year fixed effects, respectively. The parameter of interest if $theta_1$, which represents the conditional correlation between the probability of employment choices and the speed of mechanization in the residing city. $purchase_{ct}$ is rescaled by dividing by 100 so that θ_1 represents change in percentage point.

The results of the fixed effect analysis points toward the second mechanism. Figure 8 shows that mechanization is positively correlated with the probability of working in agriculture, especially for males. For each horsepower increase per hectare of farmland, men in rural families that are entitled to farmland are 36% more likely to work in agriculture. When they stay in agriculture, they are more likely to work on their family farm than work as hired agricultural labor. Figure 9 and Figure 10 shows that the probability of working on family farm is higher by 33% and working as hired agricultural labor is higher by 7% when there is 1 horsepower increase in tractor per hectare of farmland in the same city. Females in rural families that are entitled to farmland are also more likely to work in agriculture, mostly in family farm when mechanization increase, but to the less extent than men. This positive correlation is stronger for respondents below 60 years old respondents than respondents above 60 years old. Figure 11 shows that there is a negative association between mechanization and working in non-agricultural jobs, but the correlation is not statistically significant. Note that this sample is mostly above 45 years old and is not the forefront generation in rural-to-urban migration. Nevertheless, the lack positive correlation between non-agricultural employment, whether in formal and informal sector, with mechanization coincides with the main regression results.Mechanization is also positively correlated with living in the rural area, especially for respondents 60 year and older, possibly because their children can take care of them in their rural home more easily when they are working in the family farm.

6 Conclusion

The implication of capital-labor substitution is a perennial theme, one going hand-in-hand with technical change. This study provides empirical evidence on the impact of agricultural mechanization on non-agricultural employment. I find no evidence that subsidized mechanization increases non-agricultural employment in formal sectors in the same county. Higher administrative level regressions that reduce spillover effect also do not show evidence of increased local non-agricultural employment due to mechanization, except in the construction sector. The null finding is reasonable given that relative agricultural wage has been continuously dropping as non-agricultural wage increases, fueling an ongoing structural transformation even in the absence of mechanization.

In a secondary analysis, the conditional correlation coefficients show no evidence of capital-labor substitution, but rather indicate that machinery is more of a complement to labor. The fact that employment in agriculture and the speed of mechanization increased simultaneously means that mechanization boosted agricultural production that slowed down during the exodus of young labor and attracted labor back into agriculture. Combing the main regression results, one can surmise that the workers who returned to agriculture is mostly likely to be the vulnerable informal sector workers and the unemployed. Although I cannot fully rule out the possibility that machinery substituted the least-skilled labor, it is safe to posit that subsidized agricultural mechanization in China is overall welfare-improving.

The main limitation of this study is the lack of data on informal sector employment. According to the latest estimate, informal sector employment accounts for around 31.3% - 32.6% of total urban employment, which is 40% the size of migrant labor force (Chen, Huang, Huang, & Yang, 2021). If mechanization squeezed out some agricultural labor, they are mostly likely to land in informal sector jobs that are unstable and lack social security protection. Understanding the dynamics between mechanization and informal employment is essential for measuring the welfare effect of mechanization and informing policy-makers about the need to improve the social safety net.

References

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. (2020). Gendering Technological Change: Evidence from Agricultural Mechanization. *IZA Discussion Paper*, (13712).

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

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

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. (2020). *Quasi-Experimental Shift-Share Research Designs*. National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w24997

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

Diao, X., Takeshima, H., & Zhang, X. (2020). *An Evolving Paradigm of Agricultural Mechanization Development - How much can Africa learn from Asia?* (p. 548). Retrieved from http://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/134095/filename/134306.pdf

Duan, C., Lv Lidan, & Cheng, M. (2019). *(2018)* (pp. 11–11). Beijing: NBS / UNFPA / UNICEF Joint Data Project. Retrieved from NBS / UNFPA / UNICEF Joint Data Project website: https://china.unfpa.org/zh-Hans/report/20190816

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

Hassan, F., & Kornher, L. (2019). *Lets get mechanized – labor market implications of structural transformation in Bangladesh*.

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

Houssou, N., Diao, X., Cossar, F., Kolavalli, S., Jimah, K., & Aboagye, P. O. (2013). Agricultural mechanization in Ghana: Is specialized agricultural mechanization service provision a viable business model? *American Journal of Agricultural Economics*, *95*(5), 1237–1244. https://doi.org/10.1093/ajae/aat026

Ito, J., Bao, Z., & Ni, J. (2016). Land rental development via institutional innovation in rural Jiangsu, China. *Food Policy*, *59*, 1–11. https://doi.org/10.1016/j.foodpol.2015.12.005

Just, R. E., & Zilberman, D. (1988). The effects of agricultural development policies on income distribution and technological change in agriculture. *Journal of Development Economics*, *28*(2), 193–216. https://doi.org/10.1016/0304-3878(88)90058-2

Kumar, C. S., Turvey, C. G., & Kropp, J. D. (2013). The Impact of Credit Constraints on Farm Households: Survey Results from India and China.

Applied Economic Perspectives and Policy, *35*(3), 508–527. https://doi.org/10.1093/aepp/ppt002

Lohmar, B., Gale, F., Tuan, F., & Hansen, J. (2009). China's Ongoing Agricultural Modernization:Challenges Remain After 30 Years of Reform. *Economic Information Bulletin, Issue 51*, (51).

Pingali, Prabhu. (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* (Vol. 64, p. 391). https://doi.org/10.2307/144242

Takeshima, H., Prasad, R., Mahendra, A., Poudel, N., & Kumar, A. (2015). *Farm Household Typologies and Mechanization Patterns in Nepal Terai: Descriptive Analysis of the Nepal Living Standards Survey*. (December).

Wang, X., Yamauchi, F., & Huang, J. (2016). Rising wages, mechanization, and the substitution between capital and labor: Evidence from small scale farm system in China. *Agricultural Economics (United Kingdom)*, *47*(3), 309–317. https://doi.org/10.1111/agec.12231

Wang, X., Yamauchi, F., Huang, J., & Rozelle, S. (2020). What constrains mechanization in Chinese agriculture? Role of farm size and fragmentation. *China Economic Review*, *62*(October 2017), 101221. https://doi.org/10.1016/j.chieco.2018.09.002

Wang, X., Yamauchi, F., Otsuka, K., & Huang, J. (2016). Wage Growth, Landholding, and Mechanization in Chinese Agriculture. *World Development*, *86*, 30–45. https://doi.org/10.1016/j.worlddev.2016.05.002

Yamauchi, F. (2016). Rising real wages, mechanization and growing advantage of large farms: Evidence from Indonesia. *Food Policy*, *58*, 62–69. https://doi.org/10.1016/j.foodpol.2015.11.004

Yang, J., Huang, Z., Zhang, X., Reardon, T., Yang, J., Huang, Z., … 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*.

Tables

Table 1: Falsification test with county and period fixed effects - county level regressions

	Dependent variable:					
	Number of industrial firms					
	Tractors	Rotary tiller	Seeder,transplanter	Grain harvester	Corn harvester	Grain dryer
	(1)	(2)	(3)	(4)	(5)	(6)
Shift-share IV	6.197	5.059	−7.714	−14.943***	−2.772	12.989
	(4.932)	(7.083)	(12.264)	(5.693)	(10.114)	(11.730)
Sum of shares (0/1)	0.100	0.812*	1.072*	0.983**	0.398	0.870
	(0.448)	(0.437)	(0.577)	(0.466)	(0.658)	(0.715)
Observations	4,946	4,675	4,113	4,634	2,283	3,328
R^2	0.0001	0.0003	0.001	0.002	0.0001	0.001
Adjusted R^2	-0.0003	-0.0001	0.0002	0.001	-0.001	0.0002
Residual Std. Error	25.652	25.276	22.208	21.115	21.770	23.710
F Statistic	0.216	0.721	1.332	3.640**	0.110	1.278

Notes: This table reports reduced-form regressions using the shift-share instrumental variable. The dependent variable is the change in the number of industrial firms that generate more than 20 million operating revenues from their primary business, as a proxy for the change in labor demand from non-primary industries. The reported standard errors in the parenthesis are heteroskedasticounty-robust and are clustered at the county level. The fixed effects are included by within-transformations of the variables. *p<0.1; **p<0.05; ***p<0.01.

Table 2: Summary of subsidy levels and county 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 counties (C)	2295	2114	1844	2060	1028	1418
Subsidy (yuan)						
Mean	24424.5	2307.3	5526	12951.6	31769.9	28972.7
SD	15501.2	1304.9	3939.5	8397.6	16240.5	15344
Interquartile range	4886.7	264.2	906	3997.8	9650.5	5556.7
Purchased quantity						
Unit (in 1000)	Horsepower	Width (m)	Rows	Speed (kg/s)	Rows	Capacounty (ton/day)
Mean	10	139.5	0.3	0.2	0.1	0.1
SD	19.9	286.8	0.9	0.4	0.2	0.3
Interquartile range	11.4	142.5	0.2	0.1	0.1	0.1
1/Herfindal index of market shares						
Base year share	454.8	180.3	247.2	197.5	130.1	102.5
Lagged share	484.3	188.9	279.2	153.7	120.9	146.2
Largest province average market share(%)						
Base year share	0.6	0.9	1.2	1.2	1.5	1.7
Lagged share	0.9	1.3	1.1	1.3	1.9	2.1

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

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	288	294	155	273
Subsidy (yuan)						
Mean	24424.5	2307.3	5526	12951.6	31769.9	28972.7
SD	15501.2	1304.9	3939.5	8397.6	16240.5	15344
Interquartile range	4886.7	264.2	906	3997.8	9650.5	5556.7
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.

Table 4: 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	0.305	0.328	0.117	0.141	0.270	0.147
SE	0.063		0.033	0.043	0.078	0.036
Size						
ICC	0.406	0.419	0.528	0.583	0.401	0.607
SE	0.050		0.053	0.065	0.081	0.054

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

Table 5: 2SLS estimates with county and period fixed effects - county level regressions

	Dependent variable:					
	Change in the number of employee in the 2nd and 3rd industries (1000 people)					
	Tractor	Rotary tiller	Seeder,transplanter	Grain harvester	Corn harvester	Grain dryer
	(1)	(2)	(3)	(4)	(5)	(6)
Subsidized machinery purchase	−0.501	−0.029	−6.276	−32.634	446.203	31.542
	(0.660)	(0.032)	(26.007)	(60.911)	(496.969)	(26.054)
Sum of shares (0/1)	−6.057***	−4.828**	−4.197***	−5.492**	0.377	−1.852
	(2.340)	(2.207)	(1.591)	(2.486)	(3.254)	(2.504)
Observations	3,783	3,377	3,227	3,601	1,821	2,584
R^2	-0.008	-0.010	-0.001	-0.010	-1.658	-0.010
Adjusted R^2	-0.009	-0.010	-0.001	-0.011	-1.661	-0.011
Residual Std. Error	43.640	45.520	46.537	44.406	54.876	44.329
First stage F-stat	100.649	29.820	25.761	9.441	0.771	25.182

Notes: This table reports the two-stage least squares estimates on the effect of subsidized machinery purchase on the change in secondary and tertiary employment with county level data in year 2016 - 2018 from China. The shift-share instrumental variable measures the county's exposure to the machinery subsidiy policy depending on the market structure in the previous year. The reported standard errors in the parenthesis are heteroskedasticounty-robust and are clustered at the county level. The fixed effects are included by within-transformations of the variables. Craigg-Donald Wald F statistics for instrumental variable relevance are reported in the last row. *p<0.1; **p<0.05; ***p<0.01.

Table 6: 2SLS estimates with city and period fixed effects - city level regressions

	Dependent variable:					
	Change in the number of employees in the 2nd and 3rd industries (1000 people)					
	Tractors	Rotary tiller	Seeder,transplanter	Grain harvester	Corn harvester	Grain dryer
	(1)	(2)	(3)	(4)	(5)	(6)
Subsidized machinery purchase	-0.208	0.005	15.778*	18.781	-42.779	-1.408
	(0.440)	(0.030)	(8.951)	(81.589)	(690.982)	(17.506)
Sum of shares (0/1)	9.283*	11.012**	10.231***	12.111	21.575***	5.772*
	(5.045)	(4.684)	(3.841)	(7.568)	(5.995)	(3.495)
Observations	464	410	460	455	273	440
R^2	-0.017	0.010	-0.182	-0.013	-0.106	0.002
Adjusted R^2	-0.021	0.005	-0.187	-0.017	-0.114	-0.002
Residual Std. Error	47.421	49.060	51.112	47.412	58.343	47.817
First stage F-stat	33.416	11.861	4.987	2.115	0.022	4.112

Notes: This table reports the two-stage least squares estimates on the effect of subsidized machinery purchase on the change in secondary and tertiary employment with city level data in year 2016 - 2018 from China. The shift-share instrumental variable measures the city's exposure to the machinery subsidiy policy depending on the market structure in the previous year. The reported standard errors in the parenthesis are heteroskedasticity-robust and are clustered at the city level. The fixed effects are included by within-transformations of the variables. Craigg-Donald Wald F statistics for instrumental variable relevance are reported in the last row. *p<0.1; **p<0.05; ***p<0.01.

Table 7: 2SLS estimates with city and period fixed effects - city level regressions

	Dependent variable:				
	Change in the number of employee by sectors (1000 people)				
	Manufacturing	Construction	Wholesale and retail	Hotel and food service	Logistic
	(1)	(2)	(3)	(4)	(5)
Subsidized tractor purchase (1000 horsepower)	-0.158	0.152^{**}	0.013	-0.067	0.002
	(0.158)	(0.072)	(0.025)	(0.083)	(0.015)
Sum of shares (0/1)	4.434^{*}	3.021^{***}	1.797^{***}	-0.414	0.474^{**}
	(2.427)	(0.985)	(0.636)	(0.700)	(0.194)
Observations	464	464	464	464	464
R^2	-0.063	-0.079	0.017	-0.068	0.005
Adjusted R^2	-0.068	-0.084	0.013	-0.072	0.001
Residual Std. Error	19.563	11.946	5.867	7.414	3.021
First stage F-stat	33.434	33.434	33.434	33.434	33.434

Notes: This table reports the two-stage least squares estimates on the effect of subsidized tractor purchase on the change in employment in different sectors with city level data in year 2016 - 2018 from China. The shift-share instrumental variable measures the city's exposure to the machinery subsidiy policy depending on the market structure in the previous year. The reported standard errors in the parenthesis are heteroskedasticity-robust and are clustered at the city level. The fixed effects are included by within-transformations of the variables. Craigg-Donald Wald F statistics for instrumental variable relevance are reported in the last row. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Graphs

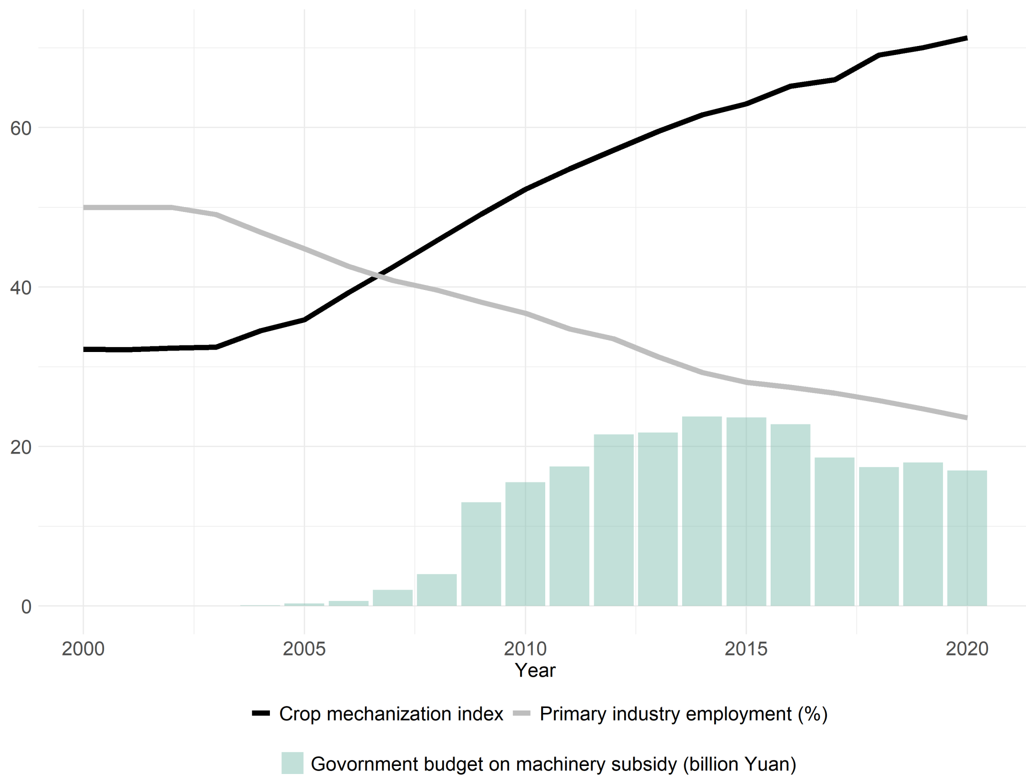

Figure 1: Crop mechanization index and subsidy budget

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.

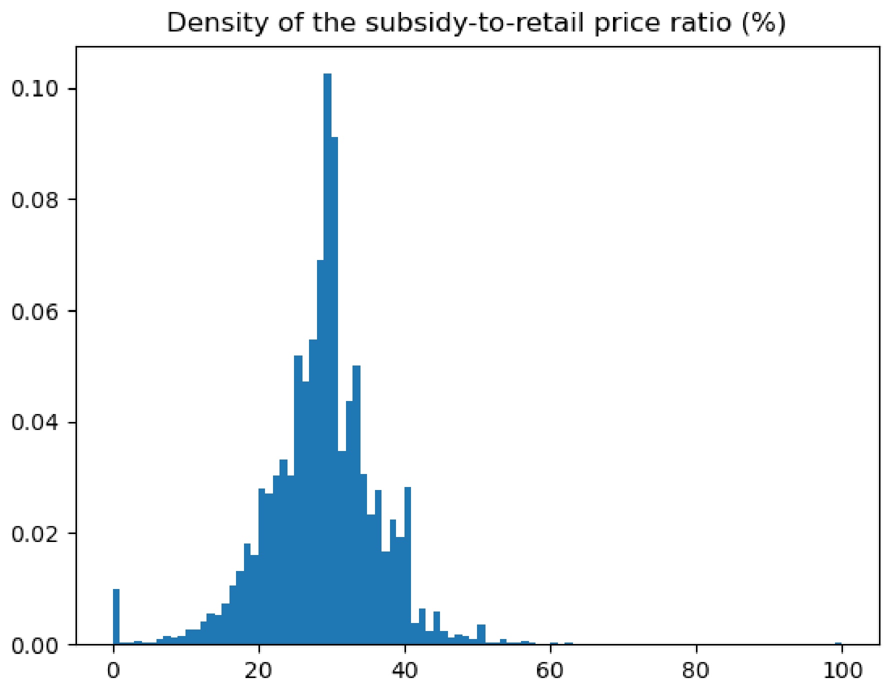

Figure 2: Density of the subsidy-to-retail price ratio (%)
Notes: The author's computation based on the subsidized machinery purchase data.

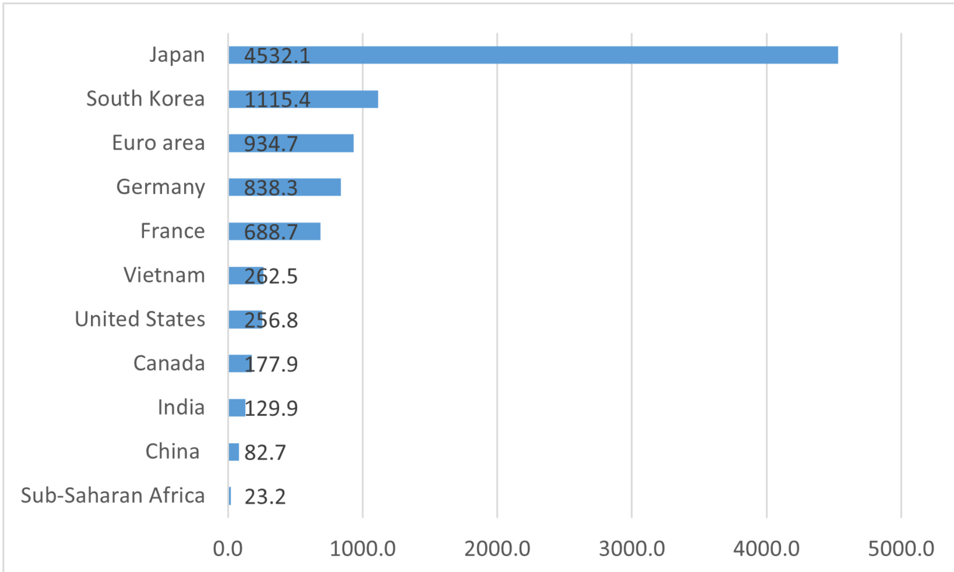

Figure 3: Agricultural machinery, tractors per 100 sq. km of arable land in year 2000

Notes: Data from World Bank. Data for Sub-Saharan Africa is for year 1987.

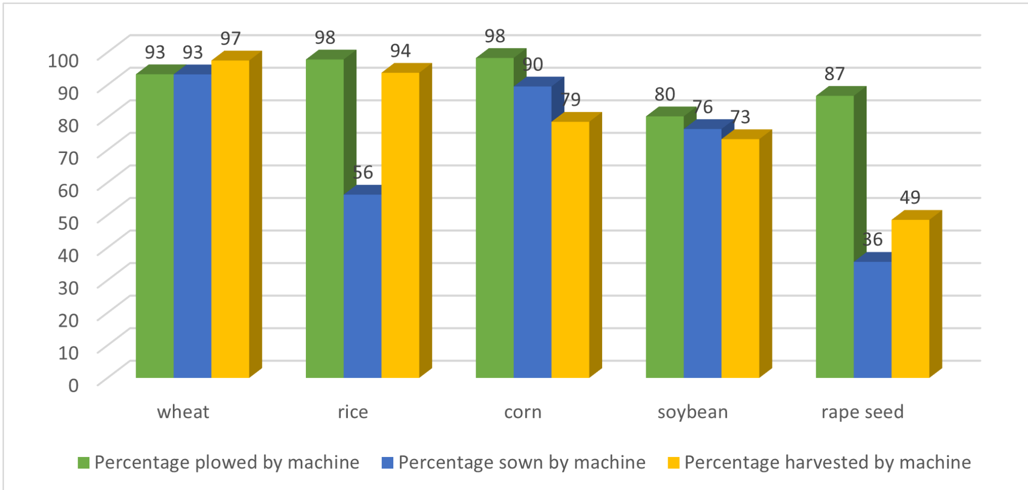

Figure 4: Percentage of mechanized area by crop and production stage in year 2020

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.

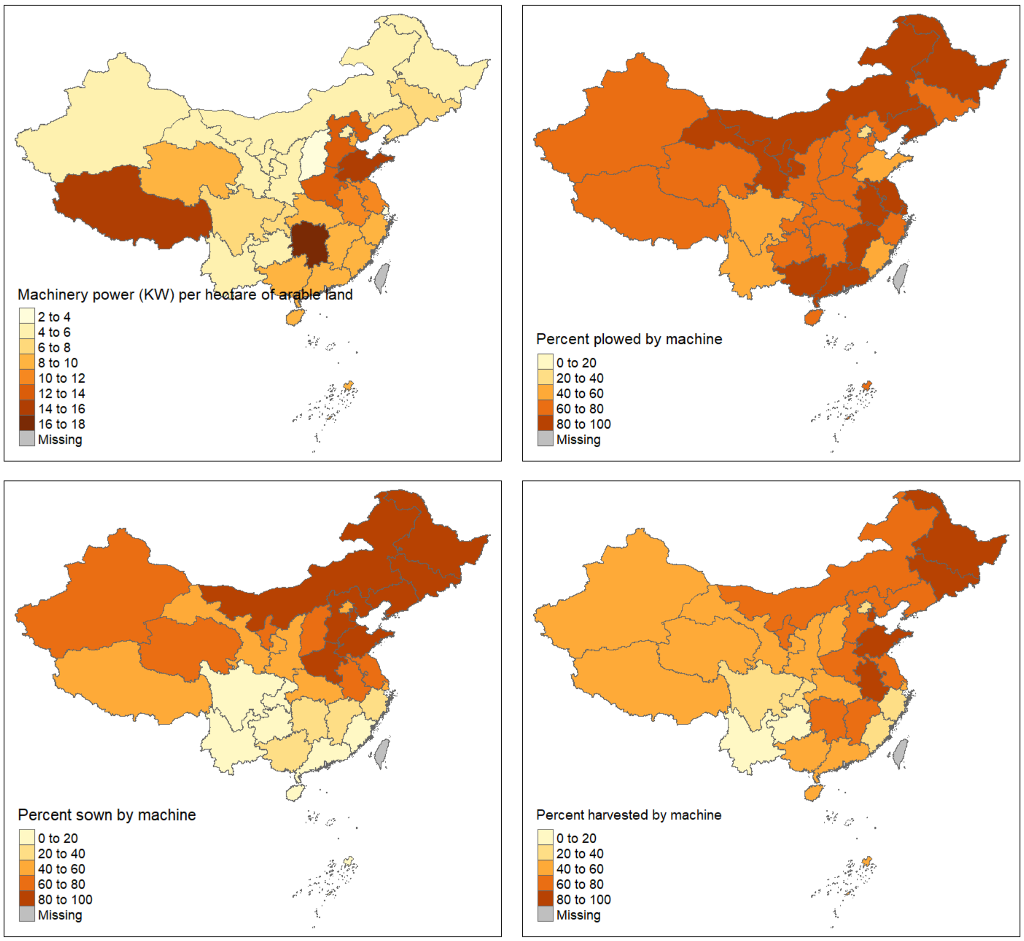

Figure 5: Mechanization by province in 2020

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.

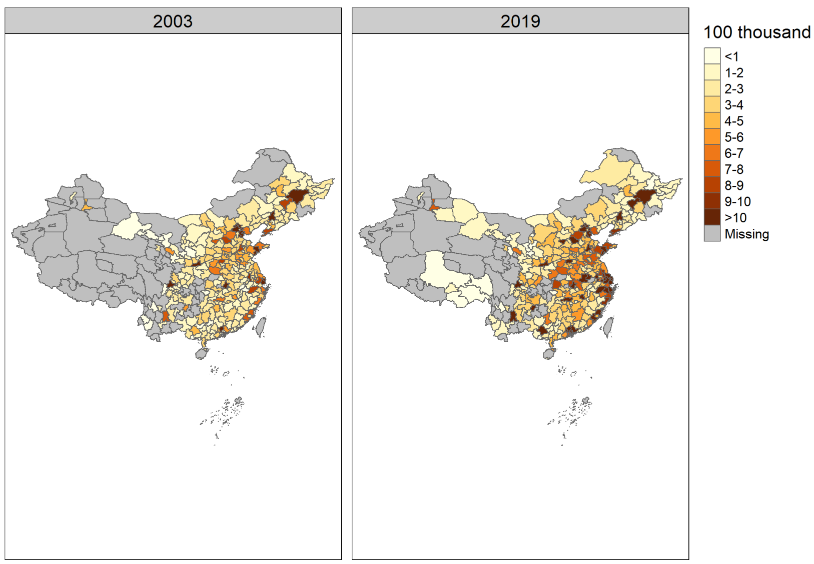

Figure 6: Number of employees in the second and tertiary industries
Notes: Data from *China City Statistical Yearbooks* 2014 - 2020. The primary sector includes agriculture, forestry, animal husbandry and fishery. The secondary industry includes mining, manufacturing, energy, water supply and construction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

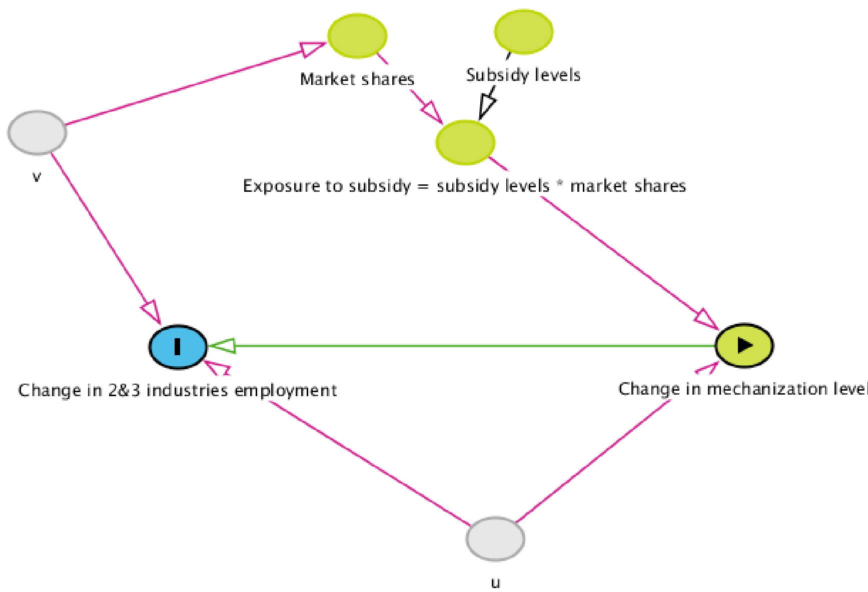

Figure 7: Directed acyclic graph of the shift-share IV design

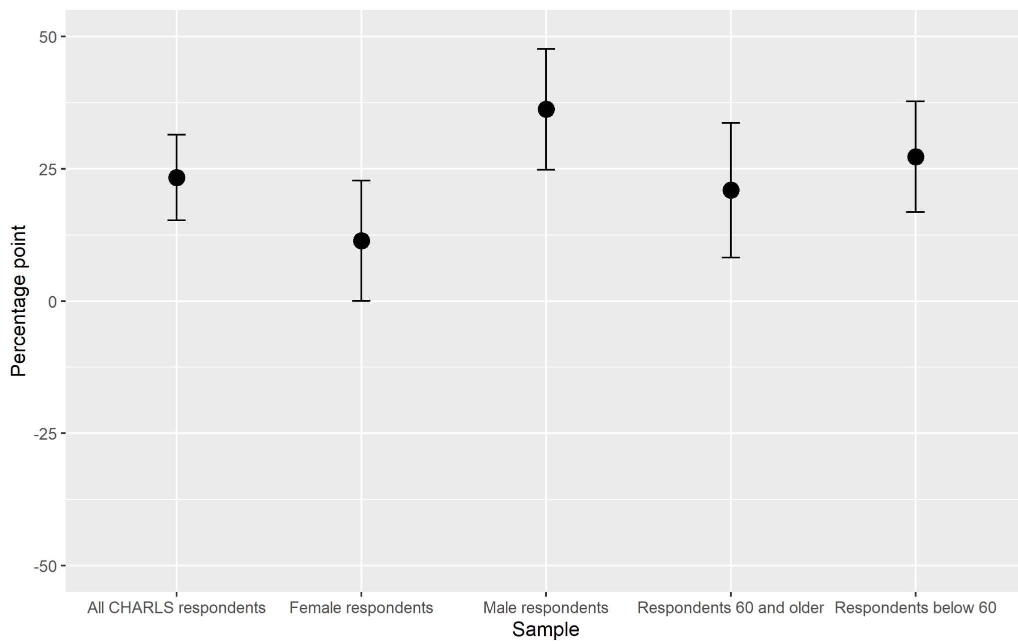

Figure 8: Correlation between probability of working in agriculture and tractor purchase in residing city (horsepower/hectare of farmland)
Notes: Based on the rural subsample of the CHARLS survey. Correlation conditional on province and year fixed effects. 95% confidence interval.

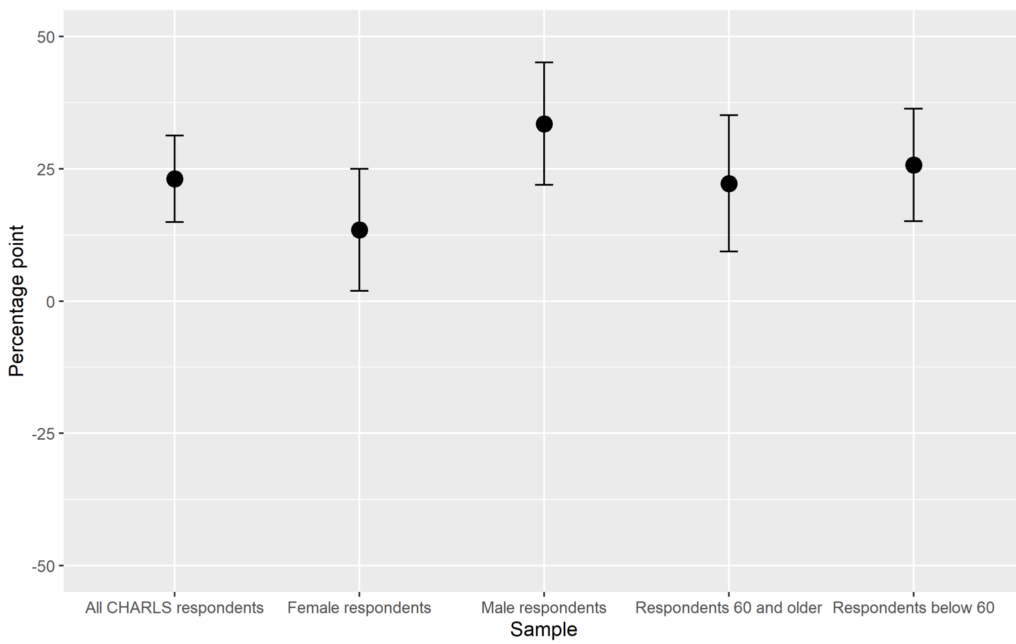

Figure 9: Correlation between probability of working on own farm and tractor purchase in residing city (horsepower/hectare of farmland)
Notes: Based on the rural subsample of the CHARLS survey. Correlation conditional on province and year fixed effects. 95% confidence interval.

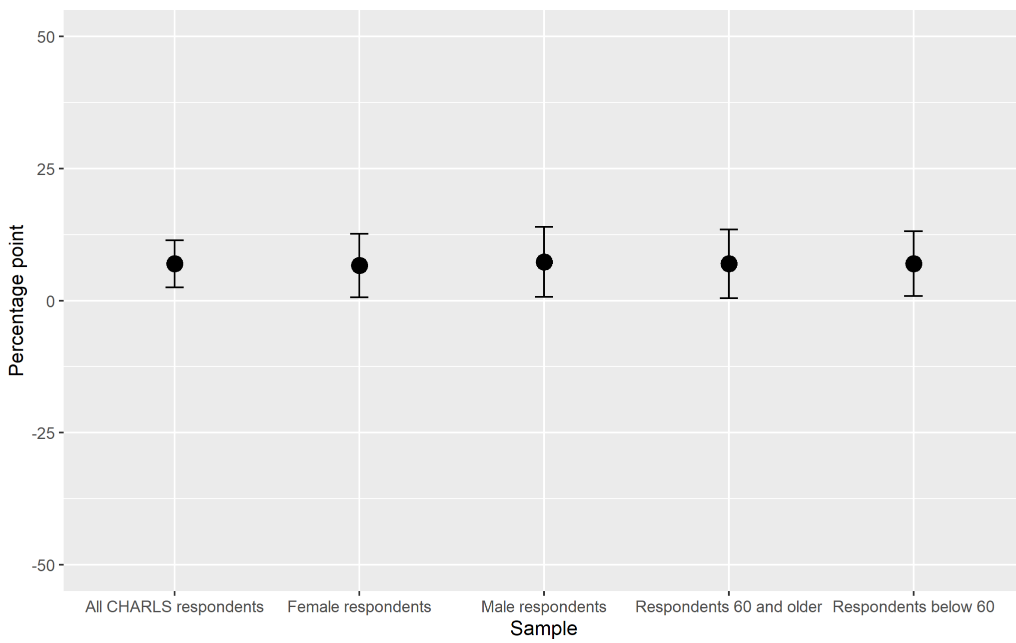

Figure 10: Correlation between probability of working as hired agricultural worker and tractor purchase in residing city (horsepower/hectare of farmland)

Notes: Based on the rural subsample of the CHARLS survey. Correlation conditional on province and year fixed effects. 95% confidence interval.

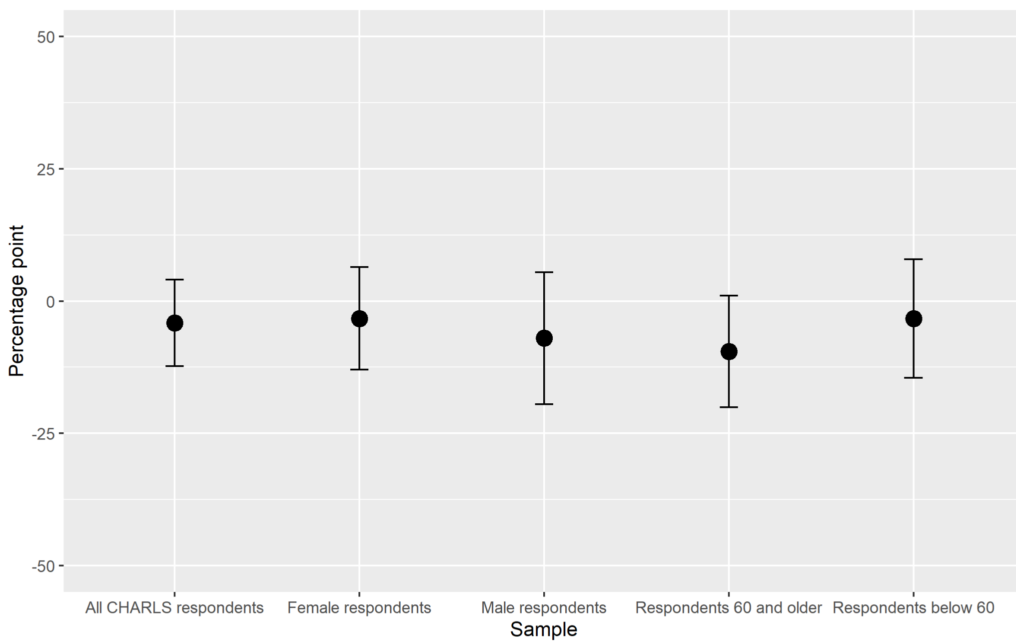

Figure 11: Correlation between probability of working in nonag jobs and tractor purchase in residing city (horsepower/hectare of farmland)
Notes: Based on the rural subsample of the CHARLS survey. Correlation conditional on province and year fixed effects. 95% confidence interval

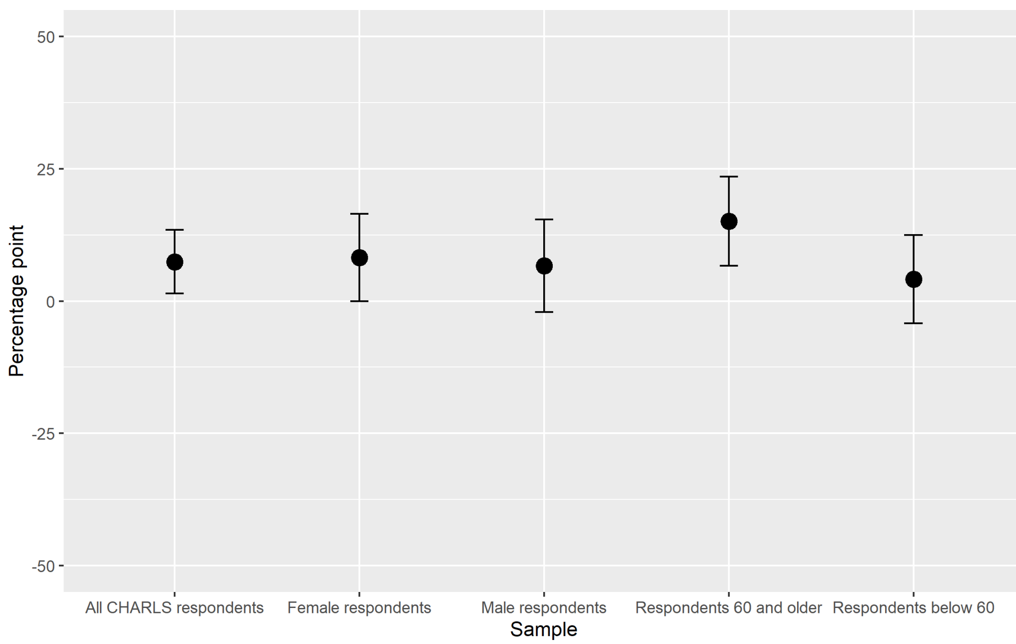

Figure 12: Correlation between probability of living in rural and tractor purchase in residing city (horsepower/hectare of farmland)
Notes: Based on the rural subsample of the CHARLS survey. Correlation conditional on province and year fixed effects. 95% confidence interval.

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 regression results

First stage results

Table A1: First stage estimates with county and period fixed effects - county level regressions

	Dependent variable:					
	Machinery purchase					
	Tractor	Rotary tiller	Seeder,transplanter	Grain harvester	Corn harvester	Grain dryer
	(1)	(2)	(3)	(4)	(5)	(6)
Shift-share IV	24.411***	440.075***	1.030***	0.144***	−0.054	0.701***
	(2.853)	(70.229)	(0.232)	(0.048)	(0.051)	(0.148)
Sum of shares (0/1)	−0.742***	−17.525***	0.014*	−0.020***	−0.005**	−0.053***
	(0.153)	(3.592)	(0.007)	(0.004)	(0.002)	(0.009)
Observations	3,783	3,377	3,227	3,601	1,821	2,584
R^2	0.027	0.011	0.008	0.007	0.001	0.017
Adjusted R^2	0.027	0.011	0.008	0.006	-0.0001	0.016
Residual Std. Error	6.248	140.723	0.373	0.144	0.100	0.257
First stage F-stat	100.649	29.820	25.761	9.441	0.771	25.182

Notes: This table reports the first stage estimates of the shift-share instrumental variable with county level data in year 2016 - 2018 from China. The shift-share instrumental variable measures the county's exposure to the machinery subsidiy policy depending on the market structure in the previous year. The endogenous variable is subsidized machinery purchase. The reported standard errors in the parenthesis are heteroskedasticcounty-robust and are clustered at the county level. The fixed effects are included by within-transformations of the variables. Craigg-Donald Wald F statistics for instrumental variable relevance are reported in the last row. The shift-share instrument is constructed using market shares in the previous year. *p<0.1; **p<0.05; ***p<0.01.

Table A2: First stage estimates with city and period fixed effects - city level regressions

	Dependent variable:					
	Machinery purchase					
	Tractor	Rotary tiller	Seeder,transplanter	Grain harvester	Corn harvester	Grain dryer
	(1)	(2)	(3)	(4)	(5)	(6)
Shift-share IV	189.654^{***}	$2,580.414^{***}$	4.243^{***}	0.997^{*}	-0.084	2.278^{**}
	(33.424)	(775.628)	(1.528)	(0.592)	(0.410)	(0.887)
Sum of shares (0/1)	-8.557^{**}	-114.519^{***}	-0.012	-0.085^{**}	-0.002	-0.082^{*}
	(3.352)	(35.847)	(0.070)	(0.040)	(0.011)	(0.046)
Observations	464	410	460	455	273	440
R^2	0.076	0.036	0.011	0.007	0.0001	0.010
Adjusted R^2	0.072	0.031	0.007	0.002	-0.007	0.006
Residual Std. Error	28.805	473.401	1.331	0.753	0.424	1.001
First stage F-stat	33.416	11.861	4.987	2.115	0.022	4.112
F Statistic	19.012***	7.511***	2.503*	1.550	0.011	2.221

Notes: This table reports the first stage estimates of the shift-share instrumental variable with city level data in year 2016 - 2018 from China. The shift-share instrumental variable measures the city's exposure to the machinery subsidiy policy depending on the market structure in the previous year. The endogenous variable is subsidized machinery purchase. The reported standard errors in the parenthesis are heteroskedasticity-robust and are clustered at the city level. The fixed effects are included by within-transformations of the variables. Craigg-Donald Wald F statistics for instrumental variable relevance are reported in the last row. The shift-share instrument is constructed using market shares in the previous year. *p<0.1; **p<0.05; ***p<0.01.