Agricultural Mechanization and Structural Change: Evidence from China

Hundanol Kebede*

December 11, 2024

Abstract

Using a unique measure of agricultural mechanization at county level, I study the link between agricultural mechanization and structural change in China. I use fluctuations in the international prices of crops and their revenue shares in counties to construct exogenous variation to mechanization and show that agricultural mechanization causes reallocation of output and employment from agriculture to manufacturing within counties. Consistent with this, both the number of manufacturing firms and their output increased significantly following mechanization. I estimate 7% aggregate productivity growth due to within county reallocation of labor between 2000-2010. Furthermore, mechanization caused large-scale emigration of workers aged 20-29.

Keywords: Aggregate productivity, Agricultural mechanization, Economic development, Structural transformation. JEL Codes: J43, O13, O14, O33, O47, Q16

*I would like to thank Sajal Lahiri for his insightful comments and suggestions. I also thank participants of seminar at SIU Carbondale, 2024 Pacific Conference for Development Economics (PacDev), SEA 94th Annual Meetings, and CEA2024 Online Conference for their helpful comments.
Southern Illinois University Carbondale, email: `hundanol.kebede@siu.edu`

1 Introduction

The role of agriculture in the process of economic development is one of the oldest and at the same time less understood topics in development economics. The dominant view is the dual economy view (Lewis, 1951; Ranis and Fei, 1961). According to this view, agriculture in low-income countries is a subsistence sector incapable of being a driving sector for sustained economic growth, and economic development requires reallocation of labor from the agricultural sector where the marginal product of labor is near zero to a modern sector where the combination of labor with capital increases the productivity of labor.[1]

A recent empirical work by Restuccia et al. (2008) shows that the major reason behind low aggregate productivity in poor countries, compared to rich countries, is that the poor countries have higher share of employment in agriculture coupled with low agricultural labor productivity.[2] The implication is that, while improving agricultural productivity gap could close the income gap between rich and poor countries partially, substantial reduction or elimination of the income gap requires reallocation of labor from the less productive agricultural sector to the more productive modern sectors in poor countries.

In the mean time, reallocation of labor from the agricultural sector to the modern sectors remains one of the key developmental challenges in low-income countries. In most countries in sub-Saharan Africa, South Asia and Latin America, between half and three-quarters of the population live on small scale farms. In such countries, structural transformation is not only required to increase aggregate productivity but also that due to increasing population pressure, climate change and environmental degradation agriculture is unlikely to be a sustainable source of livelihood for the mass of the population. This begs the question: How could these countries achieve faster reallocation of labor from agriculture to the modern sectors? The experience from the Green Revolution shows that technological improvements that increase agricultural productivity did not necessarily lead to structural change in many countries (Foster and Rosenzweig, 2004; Moscona, 2019). In this paper, I empirically study the role of a different kind of agricultural technology, i.e., agricultural mechanization, in

[1] A related but slightly different view is that the key to achieving successful industrialization lies in robust agricultural productivity growth Schultz (1953); Nurkse (1953); Rostow (1956). This view became widely adopted in neoclassical models of structural transformation (Kongsamut et al., 2001; Gollin et al., 2002, 2004; Ngai and Pissarides, 2007). However, other scholars such as Matsuyama (1992) argued that the role of agricultural productivity growth in industrialization critically depends on the openness of the economy. In an open economy environment, high agricultural productivity may lock a country into its comparative advantage agricultural sector which could be detrimental to structural change and sustained economic development in the long run.

[2] Adamopoulos and Restuccia (2021) show that such agricultural productivity gap between poor and rich countries cannot be explained by geography and soil quality. If countries produced crops across fields according to their potential yields using the same technology, the agricultural yield gap between rich and poor countries would decrease from 214% to just 5%.

driving structural transformation. Agricultural mechanization is different from the widespread adoption of high yield varieties or chemical fertilizers brought by the Green Revolution in that agricultural mechanization is a labor-saving technological change that could induce a large scale release of surplus labor from agriculture, thus providing the modern sectors with cheap labor.

I leverage county-level panel data that includes a unique measure of agricultural mechanization from China, a country that experienced both rapid rate of agricultural mechanization and dramatic structural transformation over the past three decades. My measure of agricultural mechanization is a county-level annual statistics on the total agricultural machine powers (in Kilowatts) which is obtained by aggregating all the agricultural machine powers used for ploughing, harvesting, threshing, pumping, etc. I combine this dataset with county-level panel data on output and employment disaggregated by sectors and rounds of census data since 1982.

Empirical study on the link between agricultural mechanization and structural transformation is complicated by at least two factors. First, data on accurate measure of agricultural mechanization and sectoral measures of economic activities at detailed subnational level is not widely available for most developing countries. Second, even when such data is available, exogenous source of variation to the rate of agricultural mechanization is difficult to find. Agricultural mechanization could be the driver of structural transformation by releasing surplus labor from agriculture and/or creating demand for manufacturing sector. On the contrary, rapid development of the manufacturing sector due to, say, expansion of international trade could drive agricultural mechanization by drawing workers from agriculture to manufacturing and thus raising the cost of labor to the agricultural sector.

I address the above mentioned endogeneity issue by constructing a plausibly exogenous source of variation to agricultural mechanization overtime and across counties that is informed by both theory and the context of Chinese agricultural sector. Agriculture in China is conducted predominantly by smallholder and subsistent farmers with an average farm size of 0.6 hectares. Because buying machines is not efficient at this scale, farmers who seek to mechanize their operations rely on rental machines. Rental machine reduces labor demand for farmers but it also involves significant fixed costs including the costs of search and scheduling, organization with neighbors, and transporting the machines from rental firms to the farm site. In such environment, a positive shock to farmers' revenue due to increases in prices of agricultural products would increases agricultural mechanization by allowing farmers to overcome the fixed costs of renting machines. I use these insight to construct instrumental variable (IV) for agricultural mechanization at county level. My IV is constructed by combining data on fluctuations in the international prices of crops with data on the share of these crops in the agricultural revenues of counties. In turn, the share of a particular crop in a county's agricultural revenue is calculated using data on areas of agricultural land covered by the crop and the productivity of the county in that crop. I use FOA-GAEZ data on potential yields to measure productivity of each crop in each county.

Five main results emerge from my empirical analysis. First, agricultural mechanization led to sectoral reallocation of output. Counties that had faster agricultural mechanization rate experienced significant shrinks in the share of agricultural VA in GDP and an equivalent increase in the share of manufacturing VA in GDP, with no effect on the share of service sector. Second, agricultural mechanization led to reallocation of labor from agriculture to manufacturing, with no effect on employment share of services. Third, counties that experienced faster agricultural mechanization had a significant drop in the share of population aged 20-29. Fourth, in counties that had faster agricultural mechanization, both the number of industrial enterprises and their output grew significantly. That is, agricultural mechanization concurrently induced migration of productive age population out of the county and fostered significant development of the manufacturing sector within the county. Finally, I find that within-county reallocation of labor from agriculture to manufacturing due to agricultural mechanization led to significant contribution to aggregate productivity growth. This contribution widely varies across counties and in an average county the within-county reallocation of labor led to 7.5% growth in aggregate productivity between 2000 and 2010.

Both agricultural mechanization and manufacturing demand shocks lead to reallocation of labor from agriculture to manufacturing. However, the two can be distinguished in that the former suppresses real wages while the latter causes manufacturing wages to increase. I use this insight to verify that the effects of agricultural mechanization I find are not contaminated by manufacturing demand shock. Using average real wage and capital-labor ratio constructed from firm-level data, I show that in prefectures with faster agricultural mechanization rate real wages and capital-labor ratio decrease, after accounting for manufacturing demand shock due to trade liberalization.

My main results are robust to alternative identification strategies. First, one concern in my identification strategy relates to the validity of the exclusion restriction assumption. A shock to international crop prices and hence the farmers' agricultural income not only makes adoption of machines economically feasible but also may contribute directly to manufacturing development via: (i) increasing the demand for manufactured products by the farmers and (ii) increasing the farmers' savings which would lead to capital accumulation overall and increased supply of capital to the manufacturing sector. Though it is well documented that China's manufacturing success is largely driven by exports rather than domestic demand, I address this concern using two alternative variation to agricultural mechanization. The first variation is due to a national subsidy program for machinery purchase which commenced in 2004. Using the subsidy allocation across provinces, and counties share of agricultural machine stock in 1990, I construct a shift-share IV which is less susceptible to the above discussed concerns about exclusion restriction. The second alternative IV is based on spatial variation in farmland gradient and over time fluctuation in fuel costs. It is significantly more costly to mechanize steeper farm plots than flatter ones and this cost gap increases as fuel costs increase. I use the interaction of these two variables to construct an IV. I find that my results remain strongly robust to these alternative IVs.

Second, I explore robustness of my estimation results to identification assumptions in my baseline estimation. My identification strategy is related to the shift-share research design where, in my context, the "shifts" are fluctuations in the international crop prices and the "shares" are revenue shares of these crops in the counties' agricultural income. Identification in these kinds of research design requires exogeneity of either the "shifts" Borusyak et al. (2021) or the "shares" Goldsmith-Pinkham et al. (2020). While the exogeneity of the "shifts" (i.e., fluctuations in the international crop prices) is plausible from the way these measures are constructed, the exogeneity of the "shares" is not guaranteed.[3] My identification is threatened if revenue shares of crops in counties' agricultural income are related to change in sectoral composition of output and employment within the county via other mechanisms than capturing the counties exposure to international price shocks and its effect on agricultural mechanization. To allay this concern, I conduct a placebo test where I regress within county changes in sectoral employment shares between 1982 and 1990 on change in agricultural mechanization of the counties between 1990 and 2000 and I find null effect, suggesting that my estimates are capturing period-specific effects of agricultural mechanization and not some other spurious correlation between the revenue shares of crops and structural change.

This paper contributes to empirical studies on the role of agriculture in structural change and economic development. Influential papers by Bustos et al. (2016, 2020) study the effect of agricultural productivity growth due to adoption of genetically engineered soybean seeds (GE soy) in Brazil. GE soy is significantly labor saving because the seeds are herbicide-resistant, which allows for the use of no-tillage farm technique, obviating the need of labor for land preparation and weeding. In Bustos et al. (2016), they show that in regions that adopt the technology agricultural labor shrunk and industrial employment expanded. They contrast this with the adoption

[3]Note that, identification in the shift-share research design requires the exogeneity of either the "shifts" or the "shares". However, my claim is that both the "shifts" and the "shares" are plausibly exogenous in my setting.

of a different technology, a second harvest-season of maize, which is labor intensive (land-augmenting). Regions that adopted this latter technology experienced increase in agricultural labor and decrease in manufacturing employment. In Bustos et al. (2020), they show how the adoption of GE soy increased farm profit and capital accumulation. However, they find that the increased capital did not have significant effect on local structural transformation, instead, the capital was transferred to other regions via bank networks and contributed to structural change in the destination regions. My paper relates to these papers by studying how adoption of labor-saving agricultural machines facilitates structural transformation. The main difference in my paper is that I use a comprehensive measure of agricultural mechanization instead of a technological shock to the production of a single crop.

A closely related literature uses the agricultural productivity shock during the Green Revolution (GR) as a source of variation to study the link between agricultural productivity growth and structural change. Foster and Rosenzweig (2004) find that faster industrial development in India happened in areas with lowest rates of growth in crop yield. Similarly, using the same shock Moscona (2019) finds negative effect of agricultural productivity growth on urbanization and manufacturing development by comparing districts in India and in cross-country regressions.[4] Importantly, the negative effect on manufacturing development is stronger for districts and countries more open to trade. Other studies such as (Gollin et al., 2021; McArthur and McCord, 2017) conduct cross-country analysis using the GR shock and find a positive effect of GR on agricultural productivity, GDP percapita and a negative effects on agricultural share of employment. While cross-country analysis is well-suited to account for national level general equilibrium effects of the shock by accounting for spatial spillover of the shock across sub-national regions, it also has its own downsides. In particular, different countries have different economic institutions and economic environments which may fasten or thwart structural change. As long as these same factors may critically influence successful adoption of the GR, it could be difficult to isolate the effect of the GR from these factors. In other words, treatment subjects could be too disparate to compare.[5]

One potential reason why these studies that are based on the GR find inconsistent effect on structural change is that the GR is not necessarily labor saving. At its core GR includes the adoption of high-yield seed varieties and chemical fertilizers, which significantly increase yield without decreasing labor input. On the contrary,

[4]Another study that finds null effect of agricultural productivity shock on the development of non-agricultural sector is Hornbeck and Keskin (2015) who explore the long-term effect of a shock to irrigation technology in the post-war era which made the exploitation of ground water for irrigation easier in the regions of Ogallala Aquifer.

[5]See also (von der Goltz et al., 2020; Bharadwaj et al., 2020) for studies on the effects of GR on various development outcomes such as infant mortality.

the agricultural shock studied in the current paper (as well as the one in Bustos et al. (2016, 2020)) are labor saving. In particular, the agricultural mechanization studied in this paper is a canonical case of labor saving technology where labor inputs required for ploughing, weeding, harvesting and threshing are replaced by machine power.

Other studies on the long term effects of agricultural technology shock include (Schmidt et al., 2018; Nunn and Qian, 2011; Andersen et al., 2016; Chen and Kung, 2016; Marden, 2016; Carillo, 2018). These papers mostly study the transformative effects of adoption of some critical crops: potatoes in Nunn and Qian (2011), maize in Chen and Kung (2016), and clover in Schmidt et al. (2018). While adoption of these crops might have significant effect on agricultural production by increasing yields, they are unlikely to be labor saving. A more closely related paper to the current one is Andersen et al. (2016), who study the effect of adoption of heavy plow in medieval Europe.

A bulk of recent studies explore how manufacturing demand shock following China's accession to WTO led to migration, structural change and productivity growth (Tombe and Zhu, 2019; Erten and Leight, 2021; Imbert et al., 2022; Feng et al., 2017). My paper presents evidence that agricultural mechanization serves as a *push factor* which complements the *pull factor* induced by international trade. Much less attention is give to this push factor in comparison to the pull factor.

The rest of the paper is organized as follows. Section 2 describes the data and provides important definitions and cross-country descriptive results while Section 3 outlines the empirical methodology. Section 4 presents the main results, and is followed by analysis on potential mechanisms and heterogeneous effect of agricultural mechanization across counties in Section 5. Section 6 presents estimation of aggregate productivity growth from sectoral reallocation of labor due to agricultural mechanization. Section 7 concludes the paper.

2 Data

I combine data from several sources. My main dataset is county-level panel data covering the period 1997-2015 compiled from official government reports called Annual Provincial Yearbooks. This dataset includes information on sectoral value-added, GDP, and agricultural mechanization rate, among others. The Chinese government collects annual data on agricultural mechanization rate at county, prefecture, province and national levels. Because it is difficult to aggregate machines of different size (e.g., small and large tractors) or different function (e.g., threshers and water pumps), agricultural mechanization is measured in terms of total power of all agricultural machineries in 10,000KW. I use this as a measure of agricultural mechanization.

I combine this dataset with Global Agro-Ecological Zones (GAEZ) agricultural suitability data (Fischer et al., 2021), Global cropland data (Monfreda et al., 2008), Global irrigation areas data (Nagaraj et al., 2021), and Agricultural Producer Prices (AGPP) data (FAO 1991–2021) from the FAO. The GAEZ data provides potential yields for each crop at detailed geographical unit (9km × 9km at equator) under low, intermediate and high input scenarios and under rainfed and irrigation farm methods. I use the high-input potential yield scenarios, which is likely to reflect the actual agricultural practice in China during the past three decades. Because farming in China involves both rainfed and irrigation depending on geographic regions, seasons and crop types, I combine the yield estimates under rainfed and irrigation scenarios using the fraction of irrigated land in each county as a weight to obtain a county-level crop potential yield estimates z_{kc} for each county c and crop k combination. The fraction of irrigated areas is obtained from Nagaraj et al. (2021)[6] . I use the global cropland data to estimate harvested areas in hectare for each county-crop combination. The global cropland data is based on subnational crop production information for the period 1997-2003 and combines crop cover data over these multiple years to give an average harvested area for each crop at spatial resolution of 5 min by 5 min (9km × 9km at equator). I use this data to obtain harvested area in hectare for each county-crop combination h_{kc}.

3 Methodology

3.1 Baseline specification

My baseline estimation equation is written as follows:

$$y_{ct} = \beta_0 + \beta_1 m_{ct} + \mathbf{x}'_{ct}\delta + \gamma_c + \gamma_t + \varepsilon_{ct} \tag{1}$$

where y_{ct} is outcome variable for county c in year t which includes mainly the shares of value-added and employment in the primary, secondary, and tertiary sectors. m_{ct} measures agricultural mechanization rate in county c in year t measured in logs of total agricultural machinery measured in 10,000 kilowatts. This includes all agricultural machines including those used for ploughing, harvesting, threshing, pumping, etc aggregated together. Figure 3 shows trends in the total agricultural machinery where the stock of agricultural machinery increased nearly three folds over the sample period, with substantial heterogeneity across counties. Here, β_1 captures the effect of agricultural mechanization on the outcome variables. If agricultural mechanization significantly contributes to structural change, we should see the value-

[6]While Nagaraj et al. (2021) estimate irrigated areas for each year between 2001-2015, there is little variation over time for China. So I use the 2001 estimation in my calculations.

added and employment shares in the primary sector decrease and similar shares in the secondary and/or tertiary sectors increase in counties that adopted mechanization at faster rate than other counties.

3.2 The identification challenge

Identification of β_1 in equation 1 is complicated by measurement error in agricultural mechanization and potential endogeneity of mechanization rate. First, the measure of agricultural mechanization is obtained by aggregating different types of agricultural machineries which have different functions as well as features such as size, efficiency, age, etc. This measure is likely a noisy measure of the true level of agricultural mechanization which could lead to attenuation bias in β_1.

The second challenge is potential endogeneity problem illustrated in figures 1 and 2. Consider a two sector-economy with agriculture and manufacturing sectors (assume that services are included in manufacturing). For simplicity, suppose workers are freely mobile between sectors. Figure 1 illustrates the effect of agricultural mechanization. Agricultural mechanization decreases the demand for labor in agriculture, inducing outflow of workers to manufacturing. This increases labor supply to the manufacturing sector suppressing real wage in both sectors.[7] Now consider the case of a positive demand shock to the manufacturing sector (say, due to trade liberalization) illustrated in figure 2. The booming manufacturing sector would increase labor demand and wage in the manufacturing sectors, which induces flow of labor from the agricultural sector to the manufacturing sector. The increase in manufacturing labor demand and the decrease in agricultural labor supply would raise the cost of labor in both sectors.[8] The increase in labor costs induces farmers to mechanize their operation, thus creating an identification challenge.

To address the above identification challenge, I proceed as follows. First, to account for the demand shocks to the manufacturing sector, I include two variables $\mathbf{x}_{ct}$ capturing demand shock to manufacturing sector due to China's accession to the WTO in December 2001. The first is county-level share of employment in manufacturing interacted with post-2002 dummy variable. The second variable is prefecture-level measure of improved access to U.S. market due to U.S granting Permanent Normal Trade Relation (PNTR) status to China, which become effective in 2002 (see Pierce and Schott (2016)). This later measure captures heterogeneous impact of the granting of PNTR to China across 489 four-digit industries and

[7]Note that, this discussion ignores important costs of switching sectors (and rural-urban migration). Incorporating such costs does not affect the reasoning used above; it only creates wage gaps between the sectors by somewhat suppressing the the labor flows in response to agricultural mechanization or manufacturing demand shock.

[8]Other economic liberalization policies such as FDI liberalization will have similar consequences, by causing faster growth of and increasing labor demand in the manufacturing sector.

variation in the level of industrial specialization across Chinese prefectures in 1998.[9]

Second, I construct an exogenous variation to agricultural mechanization based on features of the agricultural sector in China. Agriculture in China is conducted predominantly by smallholder farmers with an average farm size of 0.6 hectares. Moreover, each farmer's farmland is divided into several smaller non-contiguous plots of about 0.1 hectare. Such plot fragmentation has its roots in 1978 reform where rural land was divided into plots taking into consideration the quality of the plots in terms of irrigation, soil quality, drainage condition etc. To ensure equitable access to land, farmers were contracted plots from each type, which caused plots operated by a single farmer to be highly fragmented Tan et al. (2006). Because buying machines is not efficient at small scale, farmers who seek to mechanize their operations rely on rental machines for ploughing, harvesting, threshing, etc. While rental machines from local agricultural mechanization services (AMS)(Qiu et al., 2021) reduce labor costs, they involve significant fixed costs including the costs of search and scheduling, commissions to brokers, costs of organization with neighboring farmers, costs of moving the machines across fragmented plots, etc. Most of these costs can be attributed to plot fragmentation, which limits scale economies from rental machines at the household (Wang et al., 2020; Qiu and Luo, 2021). AMS achieves scale economies by pooling several farm households.

In an environment with significant fixed costs of mechanization, a positive shock to prices of agricultural products increases agricultural mechanization because the increase in farmers' revenue allows them to overcome the fixed costs of renting machines. I present details of this mechanism in appendix A. Here, I present the main idea of the model, which is very similar to the model in Bustos (2011) where firms with heterogeneous productivity make choices on whether to upgrade their technologies and export. Only higher productive firms select into export and technology upgrade because the additional revenue from doing so is large enough to overcome the fixed costs involved in exporting and technology upgrade for the more productive firms. In the current context, farmers with heterogeneous productivity make choices between two technologies: labor-intensive (traditional) technology and mechanized (modern) technology. The former involves smaller fixed costs (the costs of traditional farm tools) and higher variable costs of labor while the latter involves significantly higher fixed costs (the costs of buying machineries or renting them from AMS) but lower variable costs due its labor-saving nature. Under monopolistic competition in crop market, only the most productive farmers find it profitable to mechanize their farms because less productive farmers will not be able to generate

[9] Prefecture-level exposure to China's accession to U.S market (granting of PNTR) is calculated as $\text{NTRgap}_r = \sum_j \alpha_{jc}\text{NTRgap}_j$, where NTRgap_j the difference between the non-NTR and NTR tariff rates for industry j; α_{jr} is industry j's employment share in prefecture r in 1998. I include the $Post_{2002} \times \text{NTRgap}_r$ to account for accession to U.S market.

enough revenue to overcome the higher fixed costs of mechanization. However, as the prices of crops grown by the farmer increases more farmers find it profitable to invest in mechanization, i.e., the productivity threshold to invest in mechanization decreases (see figure A.1). I use this insight to construct instrumental variable (IV) for agricultural mechanization at county level.

I use international shocks to crop prices and variation in the revenue shares of crops across counties at the base year to construct exogenous variation to different pace of agricultural mechanization across counties. I use the AGPP data to compute global shocks to crop prices. To do so, I drop China and calculate the weighted average price of each crop P_{kt} for each year over 1991-2015 using countries' share of global exports of each crop as a weight. Following Imbert et al. (2022), I use innovations from AR(1) regression of $\log P_{kt}$ on its lagged values and crop and year fixed effects as a measure of shocks to global crop prices, i.e, my measure of price shock is the residuals from the following regression:

$$\log P_{kt} = \alpha_0 + \alpha_1 \log P_{kt-1} + \gamma_k + \gamma_t + \varepsilon_{kt} \tag{2}$$

Suppose h_{kc} is harvest area in hectare for crop k in county c, z_{kc} is yield per hectare for crop k in county c, and P_{kt} is international price of crop k in year t which is the (export) trade weighted average price of crop k in year t across all countries in the world except China. County c's exposure to international crop price shock is measured as

$$S_{ct} = \frac{\sum_k P_{k0} h_{kc} z_{kc} \hat{\varepsilon}_{kt}}{\sum_k P_{k0} h_{kc} z_{kc}} \tag{3}$$

where $\hat{\varepsilon}_{kt}$ is the price innovations estimated from 2. P_{k0} is price in base year (1990). Note that because h_{kc} and z_{kc} are time invariant, within county variation in S_{ct} across years comes from international price fluctuations. Moreover, a county's exposure to international price shock of crop k is determined by the share of this crop in the county's revenue in the base year. For instance, when there is a positive shock to international wheat price, farmers in counties with higher wheat yield and/or larger areas of land allocated to the crop, relative to other crops, would experience larger revenue gain. I use S_{ct} as an instrument for m_{ct} in equation 1. To check sensitivity of my results to alternative approaches of constructing the IV, I also calculate versions of equation 3 where $\hat{\varepsilon}_{kt}$ is replaced by $\log P_{kt}$.

Recent econometric studies on the shift-share research designs (Borusyak et al., 2021; Goldsmith-Pinkham et al., 2020) suggest conditions under which S_{ct} can be considered as an exogenous shock to m_{ct}. Borusyak et al. (2021) suggest that S_{ct} can be considered as an exogenous variation to m_{ct} as long as the "shifts" (i.e.,

$\hat{\varepsilon}_{kt}$ or $\log P_{kt}$) are exogenous, regardless of whether or not the "shares" (i.e., the revenue shares of the crops in counties) are exogenous.[10] This condition is plausibly satisfied in my setting because international prices are constructed from export prices of all countries excluding China. On the other hand, Goldsmith-Pinkham et al. (2020) argue that S_{ct} can be considered as an exogenous shock to m_{ct} as long as the "shares" are exogenous, regardless of whether or not the "shifts" are exogenous. In my setup, counties' revenue shares from different crops are constructed based on the FAO-GAEZ data on potential yields and crop cover data, where the latter is itself strongly driven by the former. Hence, the shares in my setup are likely to satisfy the exogeneity assumption required in Goldsmith-Pinkham et al. (2020), unless counties that grow certain crops are predisposed to industrialize faster for some unknown reason. I rule out this possibility in my robustness exercise using a placebo test. To sum up, the exogeneity of my IV can be justified by appealing to both Borusyak et al. (2021) and Goldsmith-Pinkham et al. (2020).

A related study by Adão et al. (2019) discusses inference in shift-share research designs. They argue that standard inference methods, such as those based on robust or geographic clustered standard errors, may be inappropriate if locations that have similar exposures to the common shocks have correlated errors. In my setting, this happens if the error terms are correlated among counties that have similar share of agricultural revenues across crops. While clustering standard errors at province levels goes a long way to address this concern, it may not fully address the problem if counties in different provinces happen to have similar "shares". In my case, this is only relevant to my first-stage F-stat. I allay any concerns about the reliability of my first-stage F-stat by presenting plots of the first-stage scatter diagram to provide a clear visualization of the predictive power of the IV for the endogenous variable.[11]

A concern in causal interpretation of my estimated effects is that a shock to agricultural revenue, S_{ct}, may have a direct effect on manufacturing development by: (i) increasing agricultural households' demand for manufactured products,[12]

[10]The only other requirement for identification, other than exogenous shifts, is that shifts-level law of large numbers should hold, i.e., there should be *many* independent shift-level shocks where each shock contributes small to the average exposure. These *many* small and independent shocks could come either from large number of cross-sections or longer time series, which is likely guaranteed in my setup given that $\hat{\varepsilon}_{kt}$ includes shocks to prices of 20 crops across 25 years.

[11]I found that estimating standard errors using the procedures suggested in Adão et al. (2019) gives significantly smaller standard errors compared to the standard errors clustered at province level, in both the first- and second-stage estimations. This implies that correlation of the error terms across counties with similar revenue shares of crops is better taken care of by clustering standard errors at province levels. A potential reason for this is that geographically close counties are more likely to be suited to the same set of crops (and hence have similar revenue shares of crops) because they are more likely to share similar climatic and soil characteristics. Thus, clustering standard errors at larger geographic units such as provinces would effectively account for potential correlation of the error terms across counties within the same province.

[12]That is, as agricultural household revenues increase due to international shocks to crop prices, the share of income they spend on manufactured products increase because the income elasticity of

and (ii) increasing saving and capital accumulation which may then be invested in the manufacturing sector. Theoretically, these effects result in violation of the exclusion restriction assumption for the IV estimation. However, the above theoretical discussion emphasizes that the role of agricultural income growth as a source of demand to stimulate the manufacturing sector development is more relevant in a closed economy environment than export oriented economies such as the Chinese economy over the past three decades. A bulk of recent literature document that manufacturing development in China took off after the country's accession to WTO (Erten and Leight, 2021; Imbert et al., 2022; Feng et al., 2017, 2016; Dai et al., 2016; Liu and Ma, 2023; Bombardini and Li, 2020). China's global export share increased from 3% in 1995 to 20% in 2020.[13] Moreover, the fact that Chinese agricultural sector is dominated by smallholder farmers with average land size of just 0.6 hectare minimizes the concern that manufacturing development is driven by demand growth and surplus in the agricultural sector.

Nevertheless, as a robustness exercise, I exploit variation to agricultural mechanization driven by subsidy program for purchase of agricultural machineries commenced in 2004. The subsidy program is funded mainly by the central government but provincial governments also contributed a small fraction. The subsidy amount dramatically increased over years from 10 million USD in 2004 to 3.5 Billion USD in 2015. The subsidy allocation data is available at province level. I thus construct a "shift-share" type instrumental variable, where the "shifts" are province-level subsidies to agricultural mechanization (financed both by provincial and central governments) and the "shares" are each county's share of agricultural machine stock in the nation at the start of our sample (1990) which also precedes the commencement of the subsidy program.

$$\mathrm{IV}_{ct} = \mathrm{MachineryShare}_{c,1990} \times \mathrm{Log}\ (1{+}\mathrm{Subsidy})_{vt} \tag{4}$$

where $\mathrm{MachineryShare}_{c,1990}$ is the share of agricultural machine stock in China that existed in county c in 1990. $\mathrm{Subsidy}_{vt}$ the total amount of subsidy disbursement in province v and year t by both provincial and central governments. I find that this IV predicts growth in agricultural machine stock across counties. From discussion of the shift-share research design above, it is plausible that the shifts are exogenous because it is unlikely that variation in the machinery subsidy disbursements across provinces and years would affect manufacturing growth via other mechanism than reallocation of labor from agriculture to manufacturing. Hence, this IV is unlikely to violate the exclusion restriction.

demand for manufactured products is typically assumed to be larger than one.

[13] https://cepr.org/voxeu/columns/china-worlds-sole-manufacturing-superpower-line-sketch-rise

In Appendix B, I propose an alternative IV constructed from spatial variation in the gradient of farmland and over time variation in fuel costs. Mechanization of steeper farmland is significantly more costly than flat surfaces, and this cost gap significantly rises when fuel costs are high. I use this insight to construct alternative IV as robustness exercise.

3.3 Measurement

Structural transformation is defined as reallocation of economic activities across the broad economic sectors agriculture, manufacturing and services and can be measured using changes in employment share, value-added share and consumption shares of these sectors over time (Herrendorf et al., 2014).[14]

My main measures of structural transformation are changes in the value-added shares of primary sector (agriculture), secondary sector (manufacturing and mining) and tertiary sector (services), for which I have county-level annual panel data since 1997. Figure 4 presents trends in the shares of sectoral VA in GDP at national level. These sectoral VA shares in GDP are weighted averages across counties where the weights are county population size in 1990. The figure shows a persistent decrease in the share of VA in primary sector in GDP from 38% in 1997 to 18% in 2015. The figure also shows that the share of secondary sector VA in GDP increased from 35% in 1997 to its highest value of 47.4% in 2013 before slightly decreasing afterwards. The share of service sector VA in GDP increased relatively fast during the late 1990s and post 2010 period, but largely remained flat in between. Note that these average changes mask massive heterogeneity in the trends in the sectoral VA shares in GDP across counties. In some counties, the reallocation of VA production from the primary to the secondary and tertiary sectors is dramatic; in other counties it is relatively small.

I complement my analysis on sectoral reallocation of VA production with data on county level employment shares from the 1982, 1990, 2000, and 2010 censuses, which I use to trace changes in sectoral share of employment over longer period. Figure 5 presents changes in the sectoral share of employment across census rounds. The figure shows relatively stable employment shares of sectors over the two decades before 2000. However, between 2000 and 2010, the share of employment in the primary sector decreased by about 15 percentage points while the share of employment in the secondary and tertiary sectors increased by about 8 and 7 percentage points, respectively. Again, these average changes in the sectoral employment shares mask substantial heterogeneity across the counties, where some counties experienced

[14] Using data from multiple countries, Herrendorf et al. (2014) show that structural change measures based on employment and value-added shares shows very similar relationship with level of economic development.

significantly larger employment reallocation across sectors than others over the decades.

The two measures of structural change have their own up and downsides, and they complement each other when used together. There are two major advantages of using the value-added shares over the employment shares. The first is the availability of annual panel data on value-added shares but only decennial panel for employment shares. Second, workers, particularly in rural areas, may spend a significant fraction of their time working outside their primary sector of employment. Rural farmers are likely to work in non-agricultural sectors during slack seasons (seasonal migration). Besides seasonal migration, farmers may work in manufacturing or services to earn additional income. This is particularly the case for smallholder farmers with too small plot of land to support their full onfarm employment or generate enough living income. Rozelle et al. (1999) estimate that about 154 million farmers (34% of rural labor force) had some form of off-farm employment in 1995 in China.[15] Guang and Zheng (2005) discuss how Chinese government statistics fails to account for these facts and thus leads to overestimation of employment in the agricultural sector.[16] The value-added shares are not subject to these type of measurement problems. However, the value-added shares have some drawbacks that the employment shares may not. In particular, some scholars (Chen et al., 2019) have argued that statistical data compiled at local government levels are unreliable. This is because local government officials' performances are evaluated based on the economic performance of their jurisdictions and this motivates the officials to inflate certain statistics such as investment and industrial production. However, the employment shares are not subject to such criticism because they come from census data that is managed by the central government, and hence there is no above mentioned incentives to manipulate these measures.

4 Results

4.1 First-stage regressions

Before I discuss the main results, I assess the predictive power of the IVs. Table 1 presents the regression of log agricultural machinery on the suggested IVs. The regressions include county and year fixed effects. The first column reports the coefficient S_{ct} calculated using $\hat{\varepsilon}_{ct}$ obtained from an AR(1) process, which is my

[15] Su et al. (2016) find that the proportion of rural workers employed simultaneously in farm and non-farm sectors is rising due to increasing tendency of rural workers to work in agriculture on part-time basis.

[16] Similarly, while urbanization is used as a proxy measure for structural change in some contexts, it is likely to be misleading in the context of China where rural manufacturing (commonly known as Township and Village Enterprises (TVEs)) is extensive (see Cheng (1996) for instance).

preferred IV. This variable strongly predicts agricultural mechanization. In the second column, I assess the predictive power of S_{ct} calculated using $\log P_{ct}$ which shows a strong correlation with agricultural mechanization (see figure 6 for visual inspection of this relationship). Overall, both the suggested IVs possess strong predictive power of the endogenous variable and yield first-stage F-stats well above the critical points.

4.2 Sectoral reallocation in VA production

OLS results: I first present the OLS estimation result of sectoral reallocation of VA production in Panel A of table 2 to facilitate comparison with the IV results. The first column in Panel A shows negative effect of agricultural mechanization on the share of agricultural VA in GDP. One log point increase in agricultural machines power decreases the share of agricultural VA in GDP by 2.7 percentage points. Column 2 of Panel A reports a positive effect of agricultural mechanization on the share of manufacturing VA in GDP where one log point increase in agricultural machine power increases share of manufacturing VA in GDP by nearly 2.8 percentage points. The last column of Panel A reports no significant effect on share of services VA in GDP. Overall, the OLS results reported in table panel A of 2 suggest that agricultural mechanization could be a key driver of structural transformation in China. However, the magnitude of these OLS estimates is likely to be biased due to the endogeneity issues discussed in the previous section as well as measurement problems in agricultural mechanization. Next, I discuss how using the IV strategy affects the estimated effects.

IV results: Panel B of table 2 reports the IV results where the IV is constructed based on innovation from an AR(1) process in equation 2. Column 1 shows that one log point in crease in agricultural machine power resulted in 10.5 percentage points decrease in the share of agricultural VA in GDP. Column 2 shows that the share manufacturing VA in GDP increased by 12.2 percentage points following one log point increase in agricultural machine power. Column 3 shows a null effect on the share of services VA in GDP. The first-stage F-stat of 60 reported in the last row of Panel B suggests that IV constructed based on innovation from an AR(1) is strong predictor of agricultural mechanization.

To investigate sensitivity of the results to the construction of the IV, I redo the same estimation based on IV constructed using fluctuation in the international prices (instead of the innovations from the AR(1) process of the international prices). The results are reported in Panel C of table 2. Column 1 reports that share of agricultural VA in GDP decreased by 10.8 percentage points while Column 2 shows that the share of manufacturing VA in GDP increased by 11.9 percentage points following one log unit increase in agricultural machine power. Column 3 shows a null effect on share of services VA in GDP. The first-stage F-stat of 56 reported in the last row shows that this IV also has strong predictive power of the agricultural mechanization. Overall, comparing the results in Panel C against their counterparts in Panel B clearly shows that the point estimates are not sensitive to whether the IV is constructed based on log prices or innovations from an AR(1) process of log prices.

Discussion: It is worth commenting on the size of the estimated effects. My preferred estimation results are those in Panels B and C, in that order, mainly due to the performance of the IVs in predicting agricultural mechanization rate. Taking the point estimates in these two panels, one log point increase in agricultural machine power (measured in 10,000KW) resulted in about 10 percentage point decrease in the share of agricultural VA in GDP and almost an equivalent increase in the share of manufacturing VA in GDP. Over my sample period 1997-2015, the median (across counties) agricultural machine power increased by nearly 1 log points (from 2.48 in 1997 to 3.53 in 2015). Thus the above point estimates reflect the effect of agricultural mechanization in a median county. Over the same period, the share of agricultural VA in GDP decreased by 22 percentage points (from 40% in 1997 to 18% in 2015) while the share of manufacturing VA in GDP increased by 12 percentage points (from 32% in 1997 to 44% in 2015)[17], in a median county. Thus, my IV estimates in Panels B and C suggest that agricultural mechanization alone accounts for close to half of the observed decrease in the share of agricultural VA in GDP in a median county and the majority of the increase in share of manufacturing VA in GDP over the sample period.[18]

4.3 Sectoral employment reallocation

The theoretical literature discussed in section 1 suggests that agricultural mechanization leads to reallocation of labor force from the agriculture to the manufacturing and service sectors. This could mean either workers are switching sectors in their original location or they are migrating to a new location and a new sector. I explore the empirical content of this prediction by looking at how sectoral shares of employment change within a county and how county demographic characteristics change across decades between 1990-2010.

[17] The share of manufacturing VA in GDP peaked in 2012, reaching 47%, and continued to shrink afterwards.

[18] Note that these projections should be interpreted cautiously. First, the reported median changes in sectoral VA shares and agricultural machinery mask significant heterogeneity across counties. Second, the point estimates are strictly speaking valid for small changes in the independent variable and using them to project the effect of large change in the independent variable may lead to predictions that are outside the range of the dependent variable in the sample.

I first explore how employment shares of sectors change in response to agricultural mechanization. To do so, I combine county-level census data for the 1990, 2000, and 2010 census rounds. This data include sectoral employment shares calculated from the 10% sample of the census. I estimate the following stacked difference across decades:

$$\Delta Y_{cd} = \alpha_0 + \alpha_1 \Delta m_{cd} + \Delta \mathbf{x}'_{cd}\delta + \gamma_c + \gamma_d + \varepsilon_{cd} \tag{5}$$

where ΔY_{cd} change in percentage share of employment in a sector (for each of the primary, secondary and tertiary sectors), γ_c is county fixed-effects, and γ_d is decade (census-round) fixed effects. I instrument over-decade changes in agricultural mechanization $\alpha_1 \Delta m_{cd}$ by similar change in the instrumental variable discussed in the previous sections, ΔS_{cd}. $\mathbf{x}_{cd}$ is a vector of control variables accounting for the demand shock to manufacturing sector due to China's accession to WTO discussed under equation 1.

The results are reported in table 3. Panel A reports the OLS results. It shows that employment in primary sector decreases by 1.44 percentage points and employment in the secondary sector increases by 1.3 percentage points following one log point increase in agricultural machine power. The effect on employment share of services is null. Panel B presents the IV result based on the IV calculated using innovations from AR(1) process of $logP$. The IV estimation shows stronger effect. One log point increase in agricultural machine power decreases employment in the primary sector by 5.26 percentage points and increases employment in the secondary sector by 4.66 percentage points. Again, the effect on employment share in services is null, mimicking similar findings on the share of services VA in GDP in the previous subsection. Panel C replicates very similar results based on an IV calculated using $logP$. It is important to note that, even though much of the variations in the IV is lost due to long differencing, the IVs perform well in terms of the first-stage F-stats across all the specifications. Overall, the results in table 3 clearly show that reallocation of workers from the primary (agricultural) sector to the secondary (manufacturing) sector *within a county* is an important mechanism through which the agricultural mechanization contributed to structural transformation.

Robustness to alternative IVs: The main threat to my IV estimation comes from potential violation of the exclusion restriction assumption. As discussed in section 3, this could happen if shocks to agricultural income due to fluctuations in the international crop prices affect VA share of manufacturing in GDP directly beyond their effect via increasing adoption of agricultural machines and release of labor to the manufacturing sector. The candidate direct effects are increased demand for manufactured goods and/or increased saving (capital accumulation) that is then channeled to the manufacturing sector to fuel manufacturing development. I address this concern using alternative IV constructed from subsidy disbursements to households for purchase of agricultural machines. This IV is less susceptible to the exclusion restriction discussed above because the subsidy is directly spent on machinery purchases and the subsidy disbursement amount (3.5 billion USD in 2015) is insignificant relative to the size of the overall manufacturing sector.

Table 4 reports the estimated results. Panels A reports the results value-added in each sector. The point estimates in columns 1-3 are very close to their counterparts reported in Panel B of table 2. Though they are slightly smaller they are not statistically different from the estimates in Panel B of table 2. Panel B of table 4 reports the effects on sectoral employment shares. Again, the effects in Column 1 are very close to their counter parts in the main results reported in column 1 of Panel C of table 2. However, the effect on manufacturing sector employment share is significantly larger compared to its counterpart in column 2 of Panel C of table 2. Part of this larger effect is because we use sectoral employment shares from 2015 mini census (instead of 2010 census) to cover a period of significant increase in subsidy disbursements. I find insignificant effect on employment share of the tertiary sector based on the new IV, similar to my baseline result. Overall, my results remain largely robust to alternative identification strategy.

In Appendix B I provide results using alternative IV constructed based on land gradient and fluctuation in international fuel costs. The results are very similar to the baseline results for both VA reallocation and employment reallocation.

Placebo test: Exogeneity of my shift-share IV requires exogeneity of either the "shifts" (Goldsmith-Pinkham et al., 2020) or exogeneity of the "shares" (Borusyak et al., 2021). In my setting, the exogeneity of the shift is plausible given how the international crop price shocks are constructed. However, exogeneity of the shares (the revenue shares of crops across counties) is not guaranteed if counties that specialize in certain crops tend to industrialize faster. That is, the concern here is that the revenue shares of crops could drive structural change in a county in some other mechanism than capturing the counties exposure to international price shock and mechanization. To allay this concern, I conduct a placebo test where I regress changes in sectoral employment shares between 1982 and 1990 on change in agricultural mechanization between 1990 and 2000 instrumented by changes in S_{ct} between 1990 and 2000. The logic behind this placebo exercises is that if revenue shares of crops in counties are related to structural change within a county via other mechanisms than driving agricultural mechanization, we would identify this relationship by regressing changes in sectoral employment shares between 1982 and

1990 on change in agricultural mechanization between 1990 and 2000 instrumented by changes in S_{ct} between 1990 and 2000. I run the following regression for each sector:

$$\Delta Y_c = \alpha_0 + \alpha_1 \Delta m_c + \gamma_p + \varepsilon_c \tag{6}$$

where $\Delta Y_c = Y_{c,1990} - Y_{c,1982}$, $\Delta m_c = m_{c,2000} - m_{c,1990}$, γ_p is province fixed effects. I instrument Δm_c by $\Delta S_c = S_{c,2000} - S_{c,1990}$. Alternatively, I also regress $\Delta Y_c = Y_{c,1990} - Y_{c,1982}$ directly $\Delta S_c = S_{c,2000} - S_{c,1990}$.

The results are presented in table 5. Panels A-B regress $\Delta Y_c = Y_{c,1990} - Y_{c,1982}$ on $\Delta S_c = S_{c,2000} - S_{c,1990}$ where S_c is constructed based on AR(1) process or $\log P$. Panels C-D estimate equation 6 using changes in the outcome variables between 1982 and 1990 and changes in the agricultural mechanization and IVs between 1990 and 2000. Across all the panels, the estimation results show null effects on the changes in employment shares across all the three sectors. Not only are the estimated effects small in size but also that they are all statistically insignificant. This suggests that our estimation results in the previous section are not driven by some preexisting long-term trend in employment shares related to revenue shares of crops across counties, but rather capture period-specific effects of agricultural mechanization.

4.4 The effect on real wage

As discussed in section 3, both agricultural mechanization and manufacturing demand shocks lead to reallocation of labor from agriculture to manufacturing. However, agricultural mechanization can be distinguished from demand shock to manufacturing in that the former suppresses real wages while the latter causes manufacturing wages to increase.

Ideally, this can be tested using data on real wages of unskilled workers in the manufacturing sector. However, I do not have access to panel data on wages by worker skills. To make progress, I construct panel data of real wages and capital/labor ratio at prefecture-level for the period 1998-2007 using census of manufacturing firms. I calculate annual wages as average wages and salaries per worker across firms in the prefecture. I use national consumer price index to deflate the nominal wages. Similarly, capital/labor ratio is calculated as real value of capital[19] divided by number of workers. I use firms' sales value as a weight to aggregate the data across firms in a prefecture. Similarly, I calculate agricultural mechanization rate at prefecture level by taking average value across counties within the prefecture. Manufacturing demand shock is measured using NTRgap interacted with post 2002 dummy (see the discussion under equation 1).

[19] Real capital is defined as asset divided by CPI.

I then investigate the effect of agricultural mechanization and manufacturing demand shock on real wages and capital/labor ratio. The results are reported in Table 6. The results show that real wages and capital/labor ratio decreased in areas that experienced faster agricultural mechanization. In areas that experienced one-log point higher agricultural mechanization, real wages decreased by 9%. In a median prefecture, the stock of agricultural machine power increased by 0.5 log points between 1998-2007, implying that nearly 50% increase in agricultural machine power is associated with 4.5% lower relative manufacturing wage rate. Similarly, agricultural mechanization is associated with decrease in capital/labor ratio. That is, in prefectures with faster mechanization, firms' production technique became more labor intensive. This results are similar to Imbert et al. (2022) who find that manufacturing firms' labor costs and capital/labor ratio decreased in areas that received larger rural-urban migration, compared to firms in areas with lower rural-urban migration between 2000-2006. Turning to the manufacturing demand shock, it has a positive effect on real wages and a negative effect on capital/labor ratio though the point estimates are statistically insignificant.

To sum up, agricultural mechanization's effect on reallocation of output and labor from agriculture to modern sectors coupled with its wage suppression effect isolates it from the forces of manufacturing demand shock.

5 Evidence on mechanisms and heterogeneous effects

5.1 Migration and demographic change

The above discussions highlight how the agricultural mechanization led to reallocation of VA and employment from the primary to the secondary sector. The decrease in the employment share of agriculture and the increase in the employment share of manufacturing could be caused by: (i) reallocation of labor from agriculture to manufacturing *within* a county, and/or (ii) migration of surplus agricultural labor out of the county and hence change in the composition of labor. If migration of surplus labor is the mechanism, we should see a decrease in the population size (of migration age) in counties with faster rate of agricultural mechanization. To explore this potential mechanism, I explore how county populations for different age groups evolve over decades following the agricultural mechanization by combining county demographic data from 1990, 2000, and 2010 censuses. I run similar specifications as equation 5 on stacked first-differences. The results are presented in tables 7 and 8.

Table 7 shows that the agricultural mechanization does not significantly decrease county population size across census rounds. However, there is significant change in the age composition of county populations. In particular, agricultural mechanization leads to about 6.5 percentage point decrease in the share of population aged 20-29.

Previous studies document that migrant workers have an average and median ages between 20-29 (see for instance (Wang, 2008; UNFPA, 2019)). This decrease in the percentage of population aged 20-29 is reflected in increase in share of population aged 19 or below and those aged 30-39. In table 8, I redo similar estimations using (log of) population counts. The results confirm that counties with faster growth in agricultural mechanization indeed experience a relative decrease in number of population aged 20-29 and increase in the number of those aged 15-19, with no significant effect on the size of the other age groups. These results suggest that the decrease in the agricultural share of employment discussed above is partly driven by out migration of productive age population in counties that experienced faster agricultural mechanization.

5.2 Further evidence on mechanisms

The above analysis shows counties with faster expansion of agricultural mechanization experienced significant emigration of active age population (ages 20-29). Moreover agricultural mechanization led to decrease in the share of agriculture and increase in the share of manufacturing in both VA production and employment. If the emigration of active age population is heavily concentrated in the agricultural sector, this may contribute to shrink in the agricultural sector size and relative increase in the share of manufacturing sector without expansion of manufacturing production within the county. Here, I rule out this possibility by showing that the agricultural mechanization in a county not only increased agricultural output but also led to expansion of manufacturing within the county.

Table 9 presents the estimation results. I explore the effect of agricultural mechanization on grain production, the number of industrial enterprises and industrial production within the counties. These later two variables are based on survey of industrial enterprises with annual main business revenue of 5 million yuan or above (i.e., they do not include small enterprises, which are perhaps more likely to be located in rural areas). Panels A and B report results based on alternative IVs. Column 1, in both panels, shows that agricultural mechanization led to significant increase in grain production. The estimated elasticity of grain production with respect to agricultural machineries is 0.27. This implies that the emigration of active age workers did not hamper agricultural production, implying that the agricultural machineries replaced the workers. This is consistent with the hypothesis that the agricultural mechanization is labor-saving technology. Column 2 estimates the effect of agricultural mechanization on the number of industrial enterprises in the county. The estimates in both panels show the elasticity of number of industrial enterprises with respect to agricultural mechanization of 0.7-0.76. Column 3 repeats similar estimation now using output of industrial enterprises (deflated using CPI). The estimates across both panels show elasticity of industrial output with respect to agricultural mechanization of about 2.3. The results in columns 2 and 3 imply that agricultural mechanization in a county led to expansion of manufacturing production within the county.

Overall, the results in table 9, together with the above analyses, clearly show that agricultural mechanization resulted in not only expansion of both agricultural and manufacturing production within the counties but also caused emigration of surplus labor out of the counties. Thus, the increase in manufacturing share of VA is not caused by shrink in the agricultural VA but by a faster expansion of the manufacturing sector within the counties. Moreover, the increase in manufacturing share of employment is not totally caused by change in worker composition due to emigration of agricultural workers. Instead, it is partially driven by workers switching sectors within their own counties. Further evidence of this is given in table 10 which shows agricultural mechanization caused significant growth in the log of real value-added in all of the primary, secondary and tertiary sectors.

5.3 Heterogeneous effects: proximity to ports

In this section, I explore potential heterogeneous effect of agricultural mechanization across locations in China depending on their proximity to ports (access to international trade). In areas with more access to international trade, agricultural mechanization may lead to larger reallocation of workers towards manufacturing. In hinterland regions where the potential for manufacturing development is weaker, worker reallocation from agriculture is likely to be either muted or is more directed towards non-tradable (service) sectors. To test this hypothesis, I calculate each county centroid's distance from the nearest port (sea/river). I interact a dummy variable indicating above/below median distance to port with log of agricultural mechanization.

The results are reported in table 11. Panel A is based on IV from AR(1) process while Panel B use IV based on LogP. In both panels, there is no statistically significant difference in the effect on employment share in the primary sector across counties with varying proximity to port. Column 2 shows that the increase in employment share of secondary sector is weaker in counties above median distance from ports (though the statistical significance of the interaction term is sensitive to the IV chosen). However, column 3 shows that employment share of tertiary sector significantly increases in counties with above median distance from ports but does not change in counties with less than median distance from ports. This result is true regardless of the IV used. Note that the average effect on employment in the tertiary sector is null (see table 3). Taken together, these results suggest that agricultural mechanization led to larger reallocation of labor from agriculture to manufacturing in counties closer to ports while it led to reallocation of labor from agriculture to services in counties further from the ports where the potential for manufacturing is weaker due to higher transport costs to international markets. That is, while the effect of agricultural mechanization on agricultural employment is more or less similar across counties, the sector absorbing surplus labor released from agriculture varies depending on the counties' proximity to ports.[20]

6 The productivity gain from sectoral reallocation

My goal in this section is to quantify the productivity gain from reallocation of workers from the primary to the secondary sector within a county in response to agricultural mechanization.[21] This analysis is related to McCaig and Pavcnik (2018) who estimate the aggregate productivity gain due to worker reallocation from the informal to formal sector in response to export demand shock in Vietnam. The key difference in my analysis is that I estimate the productivity gains at subnational (county) level. This empirical exercise requires county-level estimates of labor reallocation from the primary to the secondary sectors (which can be inferred from the regression results in table 3) and county-level measure of sectoral productivity gap between the primary and secondary sectors. I use value-added per worker in each sector as a measure of sectoral productivity and I combine county-level sectoral employment data from rounds of census data with county-level sectoral value-added data from provincial yearbook, to calculate VA per worker in each sector for the years 2000 and 2010 to measure productivity in the sector.

Let county-level aggregate productivity T (in logs) is given by the weighted average of sectoral VA per worker where the weights are employment shares of the sectors:

$$T = S^p T^p + S^s T^s + (1 - S^p - S^s) T^t \tag{7}$$

where T^p, T^s and T^t are log of value-added per worker in the primary, secondary and tertiary sectors, and S^p, S^s and $(1 - S^p - S^s)$ are employment shares of these

[20] I conduct similar analysis using proximity the nearest city using location of largest 350 cities. However, the result is omitted because I find no strong heterogeneity in the effects of agricultural mechanization based on proximity to cities.

[21] Because I do not have data on migration between counties, I cannot quantify the effect of sectoral reallocation of workers that happens due to workers migrating out of their counties. However, this channel is likely to be very important since migration in response to agricultural mechanization is quantitatively significant (particularly for people aged 20-29) and it is highly likely that migrant workers switch sectors more often than those who stay in their counties.

sectors. Productivity growth between years 2000 and 2010 is given by:

$$\begin{aligned} \Delta T = & \ S_0^p \Delta T^a + S_0^s \Delta T^s + (1 - S_0^p - S_0^s) \Delta T^t \\ & + \Delta S^p T_0^p + \Delta S^s T_0^s - (\Delta S^p + \Delta S^s) T_0^t \end{aligned} \tag{8}$$

where ΔX denotes changes in variable X between years 2000 and 2010, and 0 subscript denotes value of the variable in 2000.

From equation 8, the total productivity gain from reallocation of workers across sectors is given by $\Delta s^p T_0^p + \Delta s^s T_0^s - (\Delta s^p + \Delta s^s) T_0^t$. My goal is to quantify the productivity gain from within-county reallocation of labor from the primary to secondary sector due to agricultural mechanization because I find no evidence of labor reallocation from the primary to tertiary sector. From the IV estimation results (Panel B) in table 3, the share of labor in primary sector decreased by 0.05261, the share of employment in manufacturing increased by 0.04664, and the share of employment in tertiary sector did not change. Also, note that the employment share in secondary sector increased by closely the same amount as the decrease in employment share in the primary sector (the two point estimates are not statistically different). Imposing this equality and the fact that the employment share in service sector did not change in response to agricultural mechanization, the productivity gain from reallocation of labor from primary to secondary sector is given as $\Delta S^s (T_0^s - T_0^p)$. Here ΔS^s is change in share of employment in secondary sector due to agricultural mechanization and this can be calculated for each county as $\widehat{\Delta S_c^s} = 0.04664 \times \Delta \text{logAg.Mech}_c$. I first estimate the productivity gain from this labor reallocation in each county $\widehat{\Delta S_c^s}(T_{0,c}^s - T_{0,c}^p)$. I then calculate average productivity gain from the reallocation by taking weighted average across the counties where the weights are counties' share of national employment $\sum_c \widehat{\Delta S_c^s}(T_{0,c}^s - T_{0,c}^p) \times \theta_c$ where θ_c is the employment share in county c.

The results are reported in table 12. Column 1 reports the sectoral productivity gap (log of ratio of value-added productivity in the secondary sector to the value-added productivity in the primary sector) in the year 2000. It shows that the sectoral productivity gap varies widely across counties. In few counties, it is negative implying that the agricultural sector is more productive than the manufacturing sector. Columns 2-5 report: overall productivity growth ΔT, productivity growth due to sectoral productivity growth $S_0^p \Delta T^a + S_0^s \Delta T^s + (1 - S_0^p - S_0^s) \Delta T^t$, productivity growth due to between sector reallocation $\Delta s^p T_0^p + \Delta s^s T_0^s - (\Delta s^p + \Delta s^s) T_0^t$, and productivity growth due to reallocation from the primary to secondary sector due to agricultural mechanization $\widehat{\Delta S_c^s}(T_{0,c}^s - T_{0,c}^p)$. Panel A reports summary statistics of these estimates across counties while Panel B reports the national aggregate using the employment shares of counties as weights.

Column 2 shows significant difference in aggregate productivity growth between 2000 and 2010 across counties, ranging from 51% in the lowest percentile counties to 286% in the highest percentile counties while the average county experienced a growth of 148.5%. Column 3 shows that most of the aggregate productivity growth is attributed to productivity growth in each of the sectors. In an average county, out of 148.5% growth in aggregate productivity, 120% of it was caused by growth in productivity in the primary, secondary and tertiary sectors. The remaining growth in aggregate productivity of counties is caused by reallocation of labor across sectors within the counties. This is reported in column 4 and it is evident that within-county reallocation of labor accounts for smaller share (roughly 20%) of aggregate productivity growth in the counties. In some counties, reallocation of labor actually decreased aggregate productivity. Column 5 reports how much of the productivity growth due to sectoral reallocation of labor within counties is attributed to agricultural mechanization. Aggregate productivity growth due to reallocation of labor induced by agricultural mechanization significantly varies across counties, ranging from -5% in the lowest percentile counties to 28% in the highest percentile counties. In an average county, agricultural mechanization led to 7.5% aggregate productivity growth between 2000 and 2010.

Panel B reports the country level aggregate. Aggregate productivity increased by 145% at national level, of which 119% is attributed to growth within each sector and the remaining 26% caused by reallocation across sectors. Sectoral reallocation due to agricultural mechanization accounts for 7% of the overall growth.

Note that this exercises do not take into account sectoral reallocation of workers due to migration. Because migrant workers are more likely to switch sectors (since most of the migrations are rural-urban) it is likely that such migrations could contribute significantly to aggregate productivity growth. In this sense, my above estimates about the aggregate productivity growth due to sectoral reallocation induced by agricultural mechanization could be interpreted as a lower bound of the gain from overall reallocation induced by mechanization. Another limitation of this exercises is that it does not account for worker heterogeneity. It is likely that workers who move from agriculture to manufacturing or services are relatively more educated or have some attributes that makes them different from those who stay in agriculture. However, this is unlikely to have significant effect on the above analysis (see, for instance, Herrendorf and Schoellman (2018); Vollrath (2014)).

7 Conclusion

In this paper I study how agricultural mechanization could be a driving factor for structural transformation using a uniquely comprehensive measure of the stock of agricultural machinery at detailed geographic unit, spanning over two decades, from China.

The main empirical challenge in studying the effect of agricultural mechanization on structural transformation is endogeneity problem: agricultural mechanization could be both the driver of structural change by releasing surplus labor from agriculture and be driven by manufacturing/service sector development which may increase rural wages by drawing rural labor to these sectors. I address this endogeneity problem by constructing alternative exogenous source of variation to adoption of mechanization based on (i) international fluctuation in crop prices and the share of these crops in counties' agricultural revenues, and (ii) subsidy disbursement for purchase of agricultural machines.

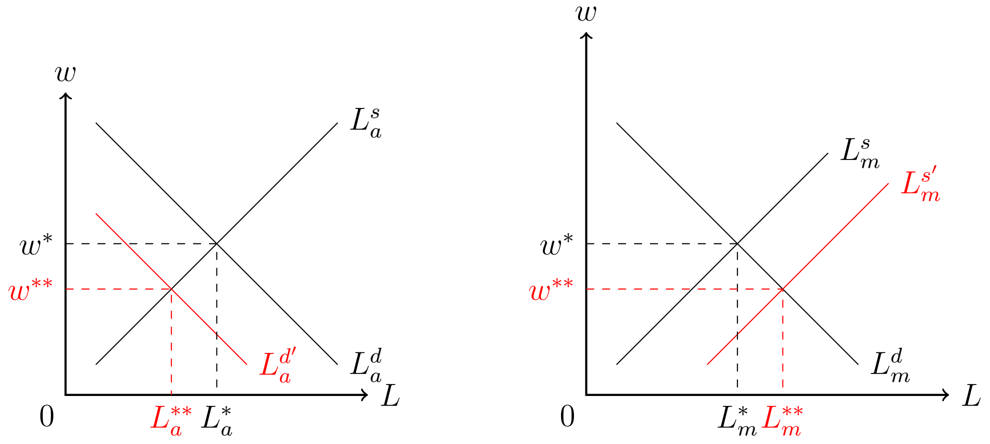
I find that agricultural mechanization led to significant reallocation of labor and value-added production from the agricultural sector towards the manufacturing sector. It led to not only out-migration of productive-age population but also caused significant expansion of manufacturing sector within the counties. In counties closer to ports, agricultural mechanization led to reallocation of labor from agricultural sector to manufacturing sector whereas in areas farther from ports agricultural labor was reallocated towards service sectors.

I quantify the aggregate productivity gain from within-county reallocation of labor at both county level and national level. I find that the effect on aggregate productivity significantly varies across counties. On average, the within-county reallocation of labor from agriculture to manufacturing led to about 7% growth in aggregate productivity at national level between 2000 and 2010. An important avenue for future studies is to quantify the aggregate productivity gain from inter-regional reallocation of labor and economic activities due to agricultural mechanization.

Figures

Figure 1: The effect of agricultural mechanization

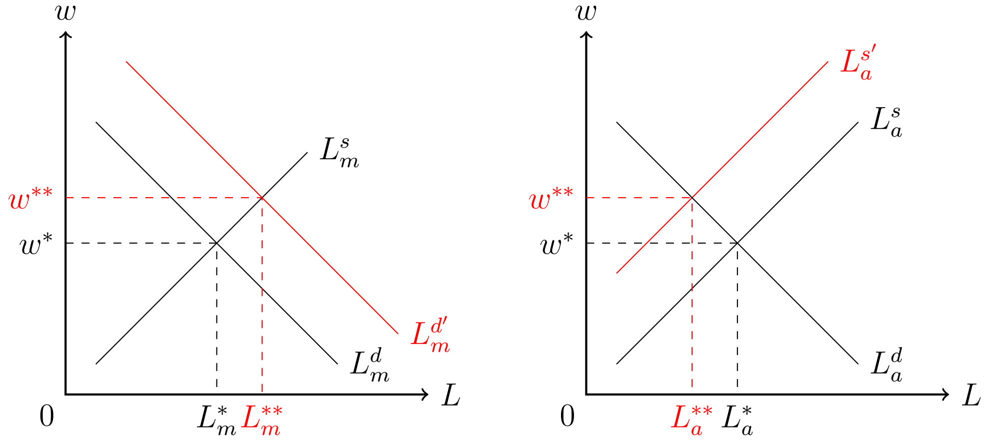
(a) Agriculture (b) Manufacturing



Notes: Mechanization decreases the demand for labor in agriculture, and induces increase in labor supply to the manufacturing sector. Wage decreases in both sectors.

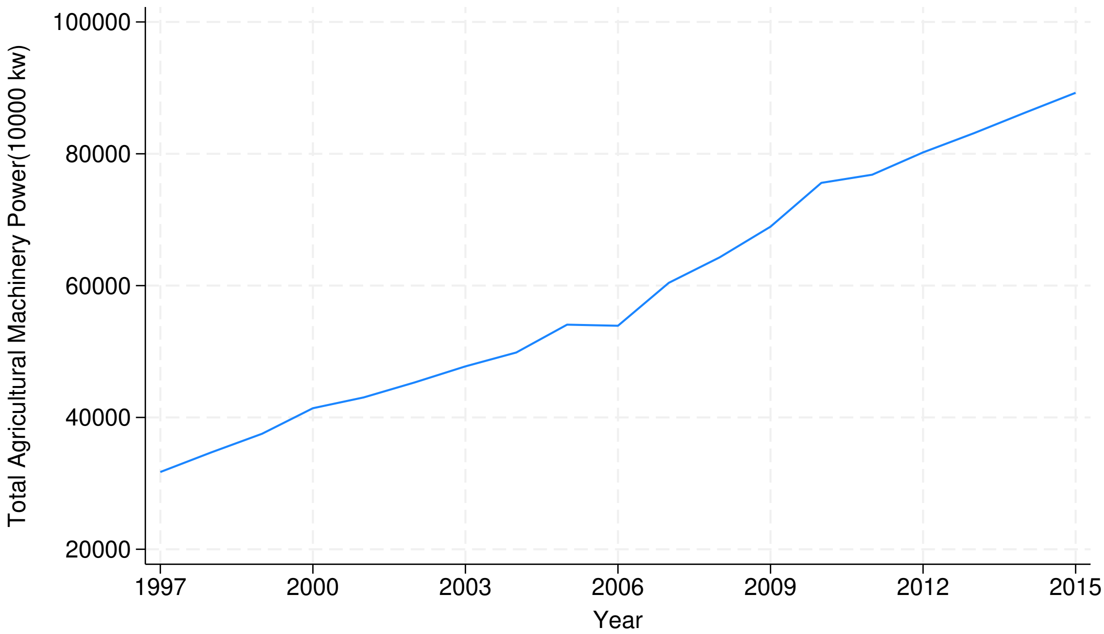
Figure 2: The effect of manufacturing demand shock

(a) Manufacturing (b) Agriculture



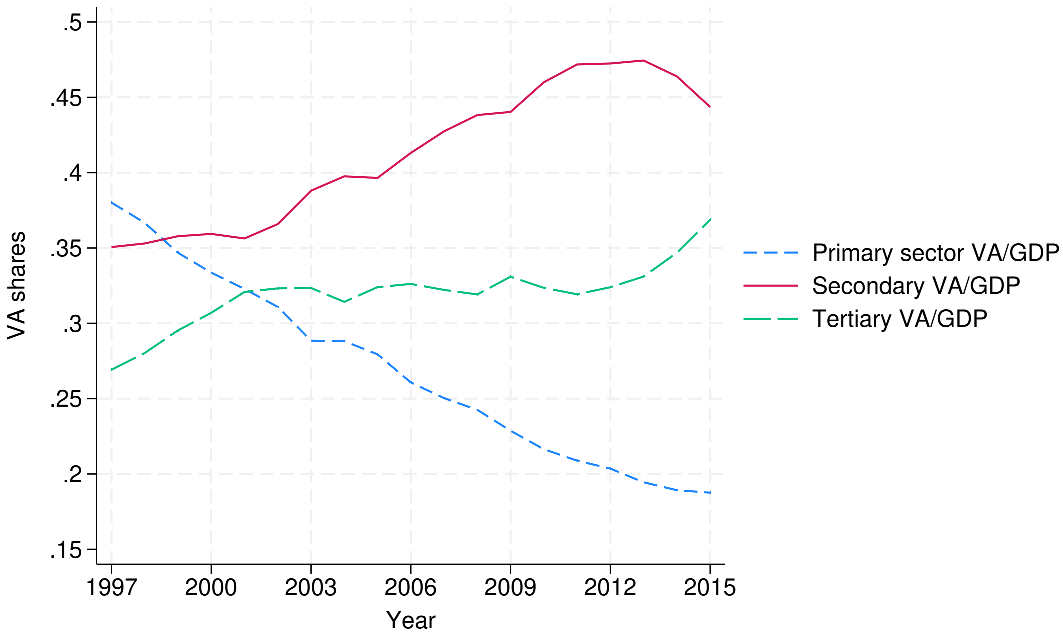
Notes: A positive demand shock to the manufacturing sector increases manufacturing labor demand and increases manufacturing wage. This induces worker migration from agriculture to manufacturing, which decreases labor supply in agriculture. Wage increases in both sectors.

Figure 3: Total agricultural machine stocks over time



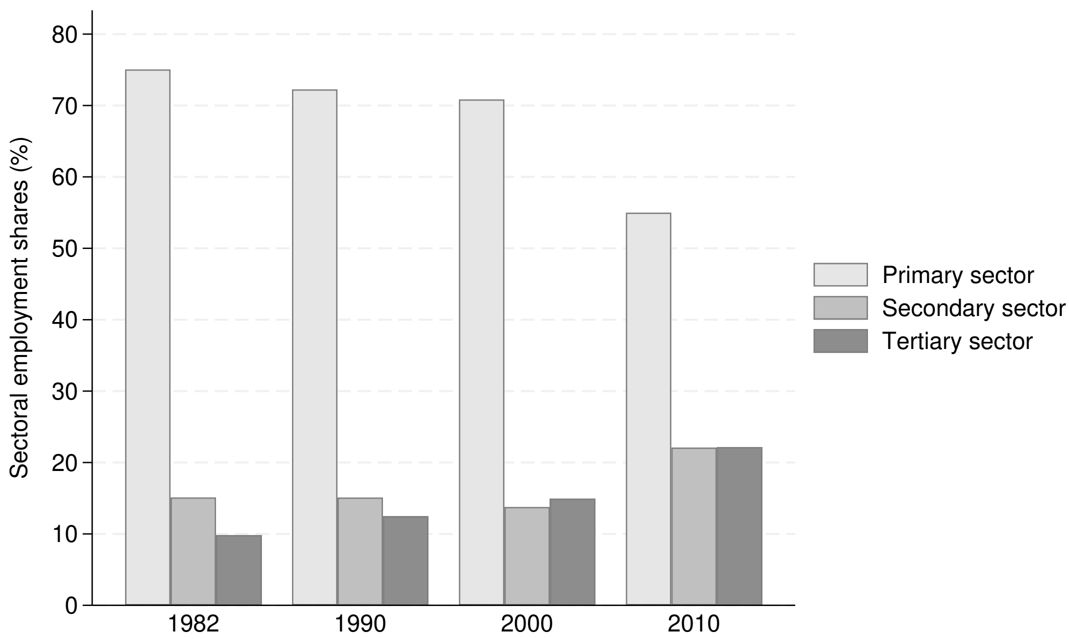
Notes: This figure presents total agricultural machine power (in 10,000KW) over time.

Figure 4: Trends in sectoral VA shares in GDP



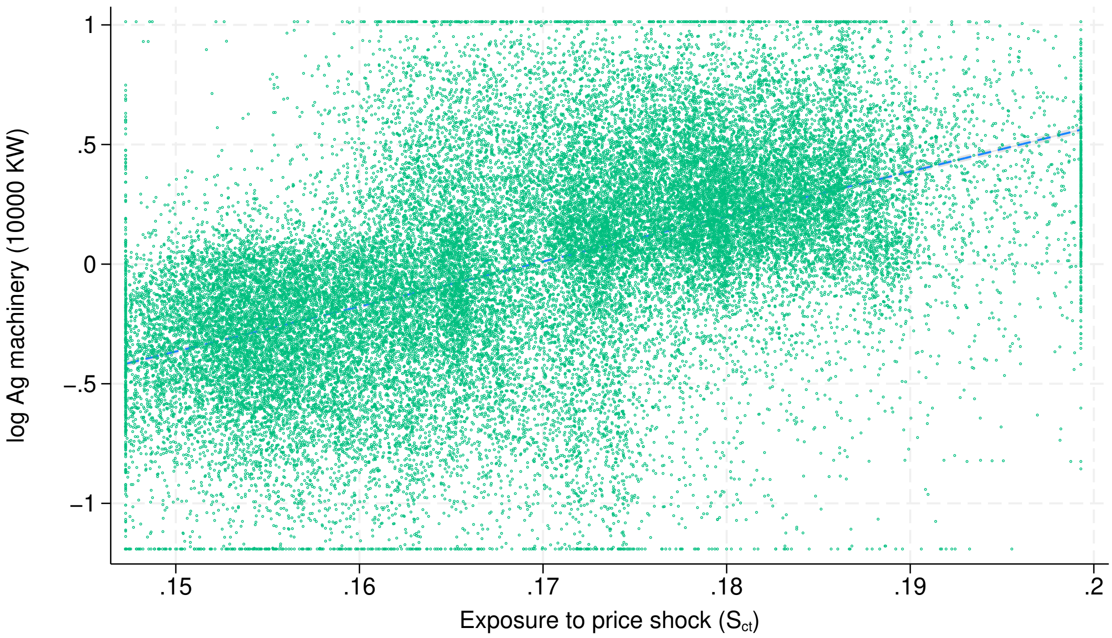
Notes: The vertical axis measures the share in GDP of value added in primary, secondary and tertiary sectors. These shares are weighted average across counties where the weights are counties' population in 1990 census. The primary sector includes agriculture, forestry, animal husbandry and fishery. The secondary sector includes manufacturing, mining and quarrying, construction, and production and supply of electricity and gas. The tertiary sector includes all the other industries (services).

Figure 5: Sectoral shares of employment over decades



Notes: This figure presents the employment shares of the primary, secondary and tertiary sectors across census years. These shares are weighted average across counties where the weights are counties' total employments in each census. The primary sector includes agriculture, forestry, animal husbandry and fishery. The secondary sector includes manufacturing, mining and quarrying, construction, and production and supply of electricity and gas. The tertiary sector includes all the other industries (services).

Figure 6: First-stage correlation



Notes: The vertical axis measures agricultural mechanization rate (measured as log of machine power in 10,000KW). I remove province fixed effects from this variable. The horizontal axis measures counties' exposure to international price shock of crops calculated using equation 3. Each dot represents a county-year observation pooled from 1997-2021. Both the dependent and independent variable are winsorized at the 1% and 99% for better visualization.

Tables

Table 1: First-stage regressions

	(1) Log Ag. Mach	(2) Log Ag. Mach
S_{ct} (based on AR(1))	4.561***	
	(0.887)	
S_{ct} (based on $\ln P_{ct}$)		4.555***
		(0.883)
N	34186	34186
R^2	0.935	0.935

Notes: All regressions include county and year fixed effects. Standard errors are clustered at county level. The estimation includes county level data from 1997-2015. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: The effect of agricultural mechanization on sectoral share of VA: main result

	(1) VA prim/GDP	(2) VA sec/GDP	(3) VA ter/GDP
	Panel A: OLS		
Log Ag.machinery	-2.694***	2.795***	-0.101
	(0.313)	(0.400)	(0.310)
N	29947	29947	29947
R^2	0.855	0.807	0.655
	Panel B: IV - based on AR(1)		
Log Ag.machinery	-10.495***	12.201**	-1.706
	(3.794)	(4.875)	(3.981)
N	29947	29947	29947
First-stage F-stat	39	39	39
	Panel C: IV - based on log$P$		
Log Ag.machinery	-10.764***	11.889**	-1.125
	(3.851)	(4.926)	(4.039)
N	29947	29947	29947
First-stage F-stat	37	37	37

Notes: All regressions include county and year fixed effects and Post 2002 dummy interacted with Prefecture-level NTR gap. Standard errors are clustered at county level. The estimation includes data from 1997-2015. The IV is constructed as MachineryShare$_{c,1990}$ $\times$ Subsidy$_{pt}$ * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: The effect of agricultural mechanization on sectoral share of employment

	(1) Δ % Primary	(2) Δ % Secondary	(3) Δ % Tertiary
		Panel A: OLS	
Δ Log Ag.machinery	-1.437***	1.303***	0.245
	(0.368)	(0.243)	(0.191)
N	3972	3972	3972
R^2	0.638	0.688	0.515
		Panel B: IV- based on AR(1)	
Δ Log Ag.machinery	-5.261***	4.664***	0.156
	(1.579)	(1.317)	(0.867)
N	3640	3640	3640
First-stage F-stat	111	111	111
		Panel C: IV - based on log$P$	
Δ Log Ag.machinery	-5.187***	4.528***	0.238
	(1.567)	(1.319)	(0.864)
N	3640	3640	3640
First-stage F-stat	112	112	112

Notes: The estimation is based on county-level data of stacked first difference across the 1990, 2000, and 2010 census rounds. All regressions include county and decade fixed effects. Standard errors are clustered at county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: The effect of agricultural mechanization: IV based on subsidy for purchase of agricultural machinery

	(1) Primary	(2) Secondary	(3) Tertiary
	Panel A: Value-added reallocation		
Log Ag.machinery	-7.557***	7.181***	0.376
	(1.328)	(1.511)	(1.236)
N	33508	33508	33508
First-stage F-stat	185	185	185
	Panel B: Employment reallocation		
Log Ag.machinery	-7.226***	11.402***	-4.331*
	(2.641)	(2.268)	(2.479)
N	5464	5249	5457
First-stage F-stat	30	30	30

Notes: All regressions include county and year fixed effects and Post 2002 dummy interacted with Prefecture-level NTR gap and a host of other county characteristics including poor county dummy and presence of specialized economic zones in the county. Standard errors are clustered at county level. Panel A is based on annual data from 1997-2015. Panel B based on census data in 1990, 2000, and 2015. All standard errors are clustered at county level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: The effect of agricultural mechanization on sectoral share of employment: placebo

	(1) Δ % Primary	(2) Δ % Secondary	(3) Δ % Tertiary
	Panel A: OLS		
Δ S_{ct} AR(1)	-0.102	0.175	-0.098
	(0.308)	(0.222)	(0.147)
N	1921	1921	1921
	Panel B: OLS		
Δ S_{ct} (based on $\ln P_{ct}$)	-0.380	0.660	-0.366
	(1.070)	(0.770)	(0.520)
N	1921	1921	1921
	Panel C: IV- based on AR(1)		
Δ Log Ag.machinery	-0.322	1.081	-1.005
	(3.060)	(2.161)	(1.508)
N	1828	1828	1828
First-stage F-stat	17	17	17
	Panel D: IV - based on log$P$		
Δ Log Ag.machinery	-0.508	1.318	-1.056
	(3.048)	(2.153)	(1.526)
N	1828	1828	1828
First-stage F-stat	17	17	17

Notes: The estimation is based on county-level data on changes in sectoral share of employment between the 1982 and 1990 census rounds and change in agricultural mechanization between the 1990 and 2000 census rounds. All regressions include province fixed effects. Standard errors are clustered at county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Prefecture-level real wage and K-L ratio

	(1) log Labor cost	(2) log K-L ratio	(3) log Labor cost	(4) log K-L ratio	(5) log Labor cost	(6) log K-L ratio
Log Ag.machinery	-0.090***	-0.086**			-0.089***	-0.087**
	(0.033)	(0.043)			(0.033)	(0.043)
NTRgap*Post2002			0.102	-0.097	0.064	-0.040
			(0.191)	(0.317)	(0.190)	(0.332)
N	2089	2089	2089	2089	2089	2089
R^2	0.868	0.834	0.867	0.833	0.868	0.834

Notes: All regressions include Prefecture and year fixed effects. Standard errors are clustered at county level. The estimation includes prefecture-level data from 1998-2007. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: The effect of agricultural mechanization on county demographic features: % composition

	(1) ΔPop	(2) Δ%Below 14	(3) Δ%15-19	(4) Δ%20-29	(5) Δ%30-39	(6) Δ%40-64	(7) Δ%above 64
Δ Log Ag.machinery	-0.007	1.618***	3.980***	-6.484***	1.256***	0.057	-0.187
	(0.055)	(0.608)	(0.596)	(0.810)	(0.434)	(0.404)	(0.141)
N	3638	3638	3638	3638	3638	3638	3638
First-stage F-stat	98	98	98	98	98	98	98
	Panel B: IV - based on log$P$						
Δ Log Ag.machinery	-0.014	1.539**	4.062***	-6.662***	1.354***	0.072	0.072
	(0.054)	(0.603)	(0.599)	(0.817)	(0.437)	(0.404)	(0.404)
N	3638	3638	3638	3638	3638	3638	3638
First-stage F-stat	98	98	98	98	98	98	98

Notes: The estimation is based on county-level data of stacked first difference across the 1990, 2000, and 2010 census ro regressions include county and decade fixed effects.Standard errors are clustered at county level. * $p < 0.10$, ** $p <$ $p < 0.01$

Table 8: The effect of agricultural mechanization on county demographic features: ΔLog numbers

	(1) Age below 14	(2) Age 15-19	(3) Age 20-29	(4) Age 30-39	(5) Age 40-64	(6) Age above 64
	Panel A: IV- based on AR(1)					
Δ Log Ag.machinery	0.051	0.439***	-0.420***	0.068	-0.013	-0.103*
	(0.060)	(0.098)	(0.066)	(0.064)	(0.059)	(0.054)
N	3640	3640	3640	3640	3640	3640
First-stage F-stat	98	98	98	98	98	98
	Panel B: IV - based on log$P$					
Δ Log Ag.machinery	0.040	0.440***	-0.438***	0.065	-0.019	-0.100*
	(0.059)	(0.098)	(0.066)	(0.063)	(0.058)	(0.054)
N	3640	3640	3640	3640	3640	3640
First-stage F-stat	99	99	99	99	99	99

Notes: The estimation is based on county-level data of stacked first difference across the 1990, 2000, and 2010 census rounds. All regressions include county and decade fixed effects.Standard errors are clustered at county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Mechanisms

	(1) Log Grain production (10000 tons)	(2) Log Number of Industrial Enterprises	(3) Log Output of Industrial Enterprises
	Panel A: IV- based on AR(1)		
Log Ag. Machinery	0.273***	0.764**	2.298***
	(0.105)	(0.311)	(0.689)
N	35499	19336	24514
First-stage F-stat	66	40	26
	Panel B: IV- based on LogP		
Log Ag. Machinery	0.270**	0.697**	2.255***
	(0.106)	(0.305)	(0.681)
N	35499	19336	24514
First-stage F-stat	65	40	27

Notes: All regressions include county and year fixed effects and Post 2002 dummy interacted with manufacturing share of county employment in 1990 and Prefecture-level NTR gap. Standard errors are clustered at county level. Data for grain production covers the period 1997-2016. Data for Industrial enterprises covers 1996-2010 for output and 2000-2010 for number of enterprises. The data on number and output of industrial enterprises is based on survey of industrial enterprises with annual main business revenue of 5 million yuan or above. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: The effect of agricultural mechanization on sectoral VA: dependent variables are in logs

	(1) VA prim	(2) VA sec	(3) GDP
	Panel A: OLS		
Log Ag.machinery	0.101***	0.210***	0.109***
	(0.012)	(0.021)	(0.014)
N	33857	33853	33857
R^2	0.967	0.946	0.971
	Panel B: IV- based on AR(1)		
Log Ag.machinery	0.808***	2.415***	1.356***
	(0.198)	(0.491)	(0.281)
N	33857	33853	33857
First-stage F-stat	31	31	31
	Panel C: IV - based on log$P$		
Log Ag.machinery	0.832***	2.449***	1.385***
	(0.204)	(0.503)	(0.289)
N	33857	33853	33857
First-stage F-stat	30	30	30

Notes: All regressions include county and year fixed effects and Post 2002 dummy interacted with manufacturing share of county employment in 1990 and Prefecture-level NTR gap. Standard errors are clustered at county level. The estimation includes county level data from 1997-2015. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Proximity to ports and heterogeneous effect of mechanization

	(1) Δ % Primary	(2) Δ % Secondary	(3) Δ % Tertiary
	Panel A: IV- based on AR(1)		
Δ Log Ag.machinery	-4.547**	5.307***	-1.389
	(1.883)	(1.450)	(0.938)
Δ Log Ag.machinery × (> median dist. from port)	-1.809	-1.632	3.919***
	(1.758)	(1.363)	(0.810)
N	3640	3640	3640
First-stage F-stat	58	58	58
	Panel B: IV - based on log$P$		
Δ Log Ag.machinery	-5.992***	6.571***	-1.043
	(2.056)	(1.618)	(1.096)
Δ Log Ag.machinery × (> median dist. from port)	1.983	-5.035***	3.158***
	(2.245)	(1.562)	(1.100)
N	3640	3640	3640
First-stage F-stat	66	66	66

Notes: The estimation is based on county-level data of stacked first difference across the 1990, 2000, and 2010 census rounds. All regressions include county and decade fixed effects. Standard errors are clustered at county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Decomposing productivity growth between 2000 and 2010

	(1) Log productivity gap (2000)	(2) Total productivity growth	(3) Growth due to growth in each sector	(4) Growth due to overall sectoral reallocation	(5) Growth due to sectoral reallocation induced by agri. mechan.
		Panel A: County-level productivity gains			
P1	-0.35	51.20	40.97	-35.59	-5.02
P5	0.88	83.67	66.42	-13.38	-0.28
P10	1.26	98.10	80.71	-4.31	0.93
P25	1.77	120.32	99.16	8.81	3.37
P50	2.33	144.93	115.61	24.52	6.56
Mean	2.31	148.53	119.99	27.98	7.50
P75	2.87	170.42	135.85	44.53	10.69
P90	3.44	203.74	162.81	66.14	15.07
P95	3.79	227.17	185.95	84.56	18.13
P99	4.50	286.27	236.13	124.81	27.95
		Panel B: Aggregate productivity gains			
National aggregate	2.37	145.05	118.89	26.12	6.94

Notes: This table decomposes growth in productivity (value-added per worker) between 2000 and 2010. Panel A reports the results for county-level exercises. Panel B reports the national aggregate result, which is obtained by aggregating counties using the national employment shares as weights. Log productivity gap in the first column is calculated as log of the ratio of value-added per worker in the secondary sector to value-added per worker in the primary sector, for the year 2000.

Appendix A Model of demand shock and investment in machines

Suppose there are K different crops indexed as $k = \{1, 2, ..., K\}$. Each crop comes with a continuum of varieties in $\nu \in [0, 1]$. Consumer preferences over different crops is given by

$$U = \sum_{k=1}^{K} \beta_j \log Q_k, \quad \text{where} \quad \sum_{k=1}^{K} \beta_k = 1, \quad \beta_k \geq 0 \tag{9}$$

Preferences of a representative consumer over a continuum of varieties of each crop k is given by:

$$Q_k = \left(\int_0^1 q_k(\nu)^{\frac{\sigma_k - 1}{\sigma_k}} d\nu \right)^{\frac{\sigma_j}{\sigma_k - 1}}, \quad \sigma_k > 1 \quad \forall k \tag{10}$$

where $q_k(\nu)$ is the quantity of variety ν of crop k consumed and σ_k is the elasticity of substitution between varieties of crop k. Let Y denotes aggregate income and $X_k = \beta_k Y$ denotes spending on crop k. The demand for each variety ν of crop k:

$$q_k(\nu) = A_k p_k(\nu)^{-\sigma_k}, \quad \text{where} \quad A_k = X_k P_k^{\sigma_k - 1} \tag{11}$$

where P_k is given by $P_k = \left(\int_\nu p_k(\nu)^{1-\sigma_k} d\nu \right)^{\frac{1}{1-\sigma_k}}$.

A positive shock to international prices of different varieties of crop k increases demand for the crop by increasing the crop price index P_k and hence the demand parameter A_k.

I Simple case

Suppose a farmer makes a choice between two technologies: labor-intensive (traditional) technology l and mechanized (modern) technology m. The total costs of producing q output of any variety of crop k on a unit of land using traditional farm technology (i.e., labor-intensive technique) is given by

$$TC_k^l = f + \frac{q}{\varphi_k} \tag{12}$$

where f is the fixed costs of traditional farm tools such as ploughs, rakes, hooks, etc, and φ_k is productivity drawn from some distribution $G_k(.)$, and $1/\varphi_k$ is a constant marginal costs of production.[22] In counties that are more productive in crop k, φ_k is likely to be higher, thus reducing the marginal costs of production.

[22] Alternatively, q can be interpreted as area of land farmed. If output per area of land is constant, this alternative interpretation gives equivalent conclusions.

Mechanized farming involves higher fixed costs but reduces variable costs by decreasing labor demand:

$$TC_k^m = f\eta + \frac{q_k}{\gamma\varphi_k} \tag{13}$$

where $f\eta$ (with $\eta > 1$) is the fixed costs of mechanized farming which includes the purchase prices of machines or rental costs of machines including the search costs, costs of transporting the machines to farm sites, costs of organizing with neighboring farmers, etc. Note that mechanization does not completely replace labor because human labor is required to operate the machines and to do activities that machines cannot effectively do. Mechanization increases labor productivity by $\gamma > 1$, thus decreasing marginal costs to $1/\gamma\varphi_k$. From now on, I drop the crop and county subscripts for convenience, unless when they are required.

Profit maximization: The CES demand and monopolistic competition imply that the profit maximizing price is given as a markup over marginal costs:

$$p = \begin{cases} \frac{\sigma}{\sigma-1}\frac{1}{\varphi}, & \text{if l} \\ \frac{\sigma}{\sigma-1}\frac{1}{\gamma\varphi}, & \text{if m} \end{cases} \tag{14}$$

Farm profit under alternative technologies is given as:

$$\Pi^l = \Gamma A\varphi^{\sigma-1} - f, \quad \textbf{and} \tag{15}$$

$$\Pi^m = \Gamma A(\gamma\varphi)^{\sigma-1} - \eta f \tag{16}$$

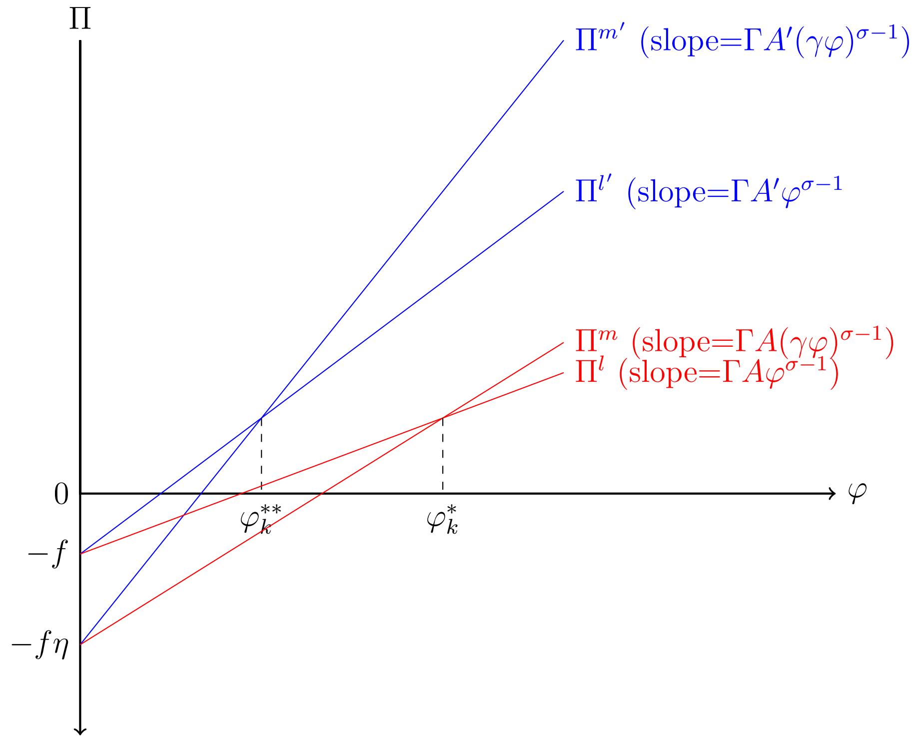
where $\Gamma = (\frac{\sigma}{\sigma-1})^{-\sigma}\frac{1}{\sigma-1}$. These profit functions are drawn in figure A.1.

The effect of international price shock on mechanization: Prior to international price shock, farmers with productivity higher than φ_k^*, where $\varphi_k^* = (\frac{(\eta-1)f}{\Gamma A_k(\gamma^{\sigma-1}-1)})^{1/(\sigma-1)}$, choose to mechanize because profit from mechanization is larger than profit from labor-intensive technique ($\Pi^m > \Pi^l$).

With a positive international price shock to crop k, the demand parameter A_k increases and the new profit functions $\Pi^{l'}$ and $\Pi^{m'}$ become steeper (see Figure A.1). This decreases the productivity cutoff for mechanization to φ_k^{**}. That is, farmers in the productivity range between φ_k^{**} and φ_k^* will not choose to mechanize before a price increase but would mechanize their farms after the price increase.

Figure A.1 shows how the fraction of farmers who choose to mechanize their farms increases in response to increase in the price of a given crop. This insight can be used

Figure A.1: The effect of crop price shock on agricultural mechanization



Notes: Π^l and Π^m farm profits corresponding to labor-intensive and mechanized technology, respectively. Farmers with productivity above φ_k^* find it profitable to mechanize while those with lower productivity choose to use labor-intensive technique instead. Positive shock to international crop prices rotate Π^l and Π^m to $\Pi^{l'}$ and $\Pi^{m'}$ and decreases the cutoff productivity for mechanization to φ_k^{**}, implying that an increase in prices lead to more farms being mechanized.

to construct exogenous variation to agricultural mechanization across counties and over years because different counties specialize in different crops according to their natural climatic suitability and crop prices vary widely across years. For instance, a positive shock to international prices of rice in a given year increases revenues of farmers in rice producing regions, thus increasing mechanization in these regions.

Appendix B Alternative identification technique

In my main identification strategy, I use exogenous variation in the profitability of mechanization across counties and years driven by fluctuations in the international prices of crops over time and variations in the crop portfolio across counties. In this section I explore the robustness of my conclusions to alternative identification technique. Specifically, I exploit plausibly exogenous variation in agricultural mechanization rates across counties and years driven by the interaction of variation in the steepness of farmlands across counties and fluctuations in the international oil prices.

The steepness of agricultural land is a crucial factor that affects the costs of mechanization, with the cost of mechanization sharply increasing on farmlands steeper than 15^o. Moreover, fuel costs account for significant fraction of the costs of agricultural mechanization. Zhang et al. (2017) find that fuel costs account for nearly 30% of the operation costs of agricultural mechanization services in China. In particular, higher fuel prices raise the costs of mechanization relatively more on steeper farmlands than flat ones. There are a number of reasons for this.

First, farming a steep terrain requires specialized equipments with additional features for stability and traction which increases the prices of the machines. Moreover, because of these additional features and the technical requirement that these equipments need to be heavier to be used on steep terrain (to avoid rollovers), they are also relatively less fuel efficient. Second, tractors and harvesters need increased engine load to climb and work on steeper slopes due to the added gravitational resistance. This increased workload leads to higher fuel consumption as the engine works harder to maintain speed and perform harvesting operations. Third, ploughing/harvesting on hills often requires frequent adjustments and maneuvers to maintain stability and prevent slippage or rollover. These additional movements contribute to higher fuel consumption as the vehicle accelerates and decelerates more frequently. Fourth, farming on steep hills may require more time to complete compared to flat terrain, especially if the machine needs to navigate complex terrain or work at reduced speeds for safety reasons. Extended operating time translates to increased fuel consumption over the duration of the ploughing/harvesting process. Fifth, operating machinery on steep slopes can accelerate wear and tear on components such as brakes, transmissions, and tires. Increased maintenance requirements, including more frequent replacements and repairs, contribute to higher overall costs.

To construct an instrumental variable that utilizes the above technical details, I use Digital Elevation Model (DEM) to calculate the average land gradient in each Chinese county. Consistent with the technical prediction, I find that, crossectionally, counties with high average slope experienced significantly lower agricultural mechanization. I combine the slope data with time series data on international prices of crude oil, which I obtain the World Bank Commodity Price Data (The Pink Sheet).[23] Because oil prices tend to be persistent, I contract shocks to oil prices from residuals of an AR(1) regression of log oil prices (following similar procedure as in my main identification strategy).[24] I then construct my IV as the interaction of the slope and the oil price shocks.

Compared to the IV in my main identification strategy, the current IV is less susceptible to violation of the exclusion restriction because it is unlikely that shocks to oil prices would affect economic activity in non-agricultural sectors differently across counties with different average land gradient. The downside of the current IV

[23] Available at this link: https://www.worldbank.org/en/research/commodity-markets.

[24] I show that my results are very similar if I use a residual from AR(2), instead of AR(1) process.

compared to my main IV is that its variation across years comes from only variation in international oil prices and its variation across counties comes only from variation in the average slope of farmlands across villages. In contrast, the cross-county variation in my main IV comes from the variation in the crop portfolio of counties combined with the counties' productivities in these crops whereas the variation over time comes from fluctuation in the international prices of over 20 crops. In sum, the two IVs have their own merits and it is crucial to examine the robustness of the main results to these alternative identification strategies.

Because the analysis on changes in sectoral employment shares utilizes changes between three censuses (1990, 2000, and 2010), I cannot rely on fluctuations in oil prices. Instead, I instrument between decade changes in agricultural mechanization within a county by the counties' average land slope. I consider both changes between 1990 and 2000 and between 1990 and 2010.

Results: The results are reported in tables A.1 and A.2. First, it is worth mentioning that the IV is strongly negatively correlated with log of agricultural mechanization, in both analysis on changes in sectoral VA shares (table A.1) and long differences in employment shares (table A.2). The first-stage F-stats reported in the two tables show these strong predictive power of the IVs.

The results in table A.1 show that agricultural mechanization increases the share of manufacturing VA by 8.5% and decreases the VA shares in agriculture and service sectors, respectively, by 3.4% and 5.1%. These results are similar across Panel A (that uses the residual from an AR(1) process of log oil prices for construction of the IV) and Panel B (that uses the residual from an AR(2) process of log oil prices). Compared to the results based on the main IV in tables 2 and 4, the current results show similar magnitude of increase in VA share of manufacturing. However, the results for changes in VA shares of agriculture and service sectors are different. In the main analysis, the VA shares in the agricultural sector declined by the same magnitude as the increase in the VA shares of the manufacturing sector, leaving the VA shares of the service sector unaffected by agricultural mechanization. The results based the current identification strategy show that agricultural mechanization leads to decreases in VA shares of both agriculture and service sectors. A potential explanation for this result is the heterogeneous effect of agricultural mechanization across counties with different geographic characteristics such as distance from ports or cities discussed above.

Turning to the changes in employment shares of sectors, Panel A of table A.2 regresses within-county changes in the employment shares of sectors between the 1990 and 2000 censuses on within-county changes in agricultural mechanization, where the latter is instrumented by the slope of land gradient in the county. The result shows that agricultural mechanization led to about 6 percentage point decrease in the share of employment in the agriculture, which led to 3 percentage points increase in share of employment in the manufacturing sector and 2.2 percentage points increase in the share of employment in the service sector. Panel B of table A.2 repeats similar analysis using between 1990 and 2010 changes sectoral employment shares. The result shows that agricultural mechanization led to about 10.6 percentage points decrease in the employment share of agriculture, the vast majority of which is absorbed into the manufacturing sector (7.3 percentage point increase) while the effect on employment share of services is small and statistically insignificant. Overall, the results in table A.2 are broadly consistent with their counterparts in the main analysis (table 5).

Table A.1: The effect of agricultural mechanization on sectoral share of VA: Alternative IV based on land gradient and fuel price fluctuation

	(1) VA prim/GDP	(2) VA sec/GDP	(3) VA ter/GDP
	Panel A: IV- based on AR(1)		
Log Ag.machinery	-0.047***	0.081***	-0.035**
	(0.014)	(0.021)	(0.016)
N	35375	35375	35375
First-stage F-stat	141	141	141
	Panel B: IV - based on AR(2)		
Log Ag.machinery	-0.044***	0.078***	-0.034**
	(0.014)	(0.022)	(0.017)
N	35375	35375	35375
First-stage F-stat	122	122	122

Notes: All regressions include county and year fixed effects and Post 2002 dummy interacted with manufacturing share of county employment in 1990 and Prefecture-level NTR gap. Standard errors are clustered at county level. The estimation includes county level data from 1997-2017. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: The effect of agricultural mechanization on sectoral share of employment: Using slope as an IV

	(1)	(2)	(3)
	Δ % Primary	Δ % Secondary	Δ % Tertiary
	Panel A: Change between 1990 and 2000		
Δ Log Ag.machinery	-6.099***	3.025**	2.188**
	(2.047)	(1.396)	(1.032)
N	1918	1918	1918
First-stage F-stat	51	51	51
	Panel B: Change between 1990 and 2010		
Δ Log Ag.machinery	-10.574**	7.300**	1.996
	(4.293)	(2.951)	(1.959)
N	1901	1901	1901
First-stage F-stat	29	29	29

Notes: The estimation is based on county-level data of difference between 1990 and 2000 in Panel A and difference between 1990 and 2010 in Panel B. Within-county change in agricultural mechanization is instrumented for by the average slope of land in the county. All regressions include province fixed effects control variables for proximity to port and distance to city. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

Adamopoulos, T. and Restuccia, D. (2021). Geography and Agricultural Productivity: Cross-Country Evidence from Micro Plot-Level Data. *The Review of Economic Studies*, 89(4):1629–1653.

Adão, R., Kolesár, M., and Morales, E. (2019). Shift-Share Designs: Theory and Inference*. *The Quarterly Journal of Economics*, 134(4):1949–2010.

Andersen, T. B., Jensen, P. S., and Skovsgaard, C. V. (2016). The heavy plow and the agricultural revolution in medieval europe. *Journal of Development Economics*, 118:133–149.

Bharadwaj, P., Fenske, J., Kala, N., and Mirza, R. A. (2020). The green revolution and infant mortality in india. *Journal of Health Economics*, 71:102314.

Bombardini, M. and Li, B. (2020). Trade, pollution and mortality in china. *Journal of International Economics*, 125:103321.

Borusyak, K., Hull, P., and Jaravel, X. (2021). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies*, 89(1):181–213.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms. *American Economic Review*, 101(1):304–40.

Bustos, P., Caprettini, B., and Ponticelli, J. (2016). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review*, 106(6):1320–65.

Bustos, P., Garber, G., and Ponticelli, J. (2020). Capital Accumulation and Structural Transformation. *The Quarterly Journal of Economics*, 135(2):1037–1094.

Carillo, M. (2018). Agricultural policy and long-run development: evidence from mussolini's battle for grain. Mpra paper, University Library of Munich, Germany.

Chen, S. and Kung, J. (2016). Of maize and men: the effect of a new world crop on population and economic growth in china. *Journal of Economic Growth*, 21(1):71–99.

Chen, W., Chen, X., Hsieh, C.-T., and Song, Z. (2019). A forensic examination of china's national accounts. Working Paper 25754, National Bureau of Economic Research.

Cheng, H. (1996). Chapter 11 promoting township and village enterprises as a growth strategy in china. In *Inflation and Growth in China*, page ch013. International Monetary Fund, USA.

Dai, M., Maitra, M., and Yu, M. (2016). Unexceptional exporter performance in china? the role of processing trade. *Journal of Development Economics*, 121:177–189.

Erten, B. and Leight, J. (2021). Exporting Out of Agriculture: The Impact of WTO Accession on Structural Transformation in China. *The Review of Economics and Statistics*, 103(2):364–380.

Feng, L., Li, Z., and Swenson, D. L. (2016). The connection between imported intermediate inputs and exports: Evidence from chinese firms. *Journal of International Economics*, 101:86–101.

Feng, L., Li, Z., and Swenson, D. L. (2017). Trade policy uncertainty and exports: Evidence from china's wto accession. *Journal of International Economics*, 106:20–36.

Fischer, G., Nachtergaele, F., van Velthuizen, H., Chiozza, F., Francheschini, G., a. H. M., Muchoney, D., and Tramberend, S. (2021). Global agro-ecological zone v4 – model documentation. Technical report, FAO IIASA.

Foster, A. D. and Rosenzweig, M. R. (2004). Agricultural productivity growth, rural economic diversity, and economic reforms: India, 1970–2000. *Economic Development and Cultural Change*, 52(3):509–542.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Gollin, D., Hansen, C. W., and Wingender, A. M. (2021). Two blades of grass: The impact of the green revolution. *Journal of Political Economy*, 129(8):2344–2384.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review*, 92(2):160–164.

Gollin, D., Parente, S. L., and Rogerson, R. (2004). Farm work, home work and international productivity differences. *Review of Economic Dynamics*, 7(4):827–850.

Guang, L. and Zheng, L. (2005). Migration as the second-best option: Local power and off-farm employment. *The China Quarterly*, (181):22–45.

Herrendorf, B., Rogerson, R., and Ákos Valentinyi (2014). Chapter 6 - growth and structural transformation. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 2 of *Handbook of Economic Growth*, pages 855–941. Elsevier.

Herrendorf, B. and Schoellman, T. (2018). Wages, human capital, and barriers to structural transformation. *American Economic Journal: Macroeconomics*, 10(2):1–23.

Hornbeck, R. and Keskin, P. (2015). Does agriculture generate local economic spillovers? short-run and long-run evidence from the ogallala aquifer. *American Economic Journal: Economic Policy*, 7(2):192–213.

Imbert, C., Seror, M., Zhang, Y., and Zylberberg, Y. (2022). Migrants and firms: Evidence from china. *American Economic Review*, 112(6):1885–1914.

Kongsamut, P., Rebelo, S., and Xie, D. (2001). Beyond balanced growth. *The Review of Economic Studies*, 68(4):869–882.

Lewis, A. (1951). The principles of economic planning. *American Political Science Review*, 45(2).

Liu, C. and Ma, X. (2023). Migration, tariffs, and china's export surge. *Journal of International Economics*, 140:103696.

Marden, S. (2016). The agricultural roots of industrial development: 'forward linkages' in reform era china. Working paper series, Department of Economics, University of Sussex Business School.

Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory*, 58(2):317–334.

McArthur, J. W. and McCord, G. C. (2017). Fertilizing growth: Agricultural inputs and their effects in economic development. *Journal of Development Economics*, 127:133–152.

McCaig, B. and Pavcnik, N. (2018). Export markets and labor allocation in a low-income country. *American Economic Review*, 108(7):1899–1941.

Monfreda, C., Ramankutty, N., and Foley, J. A. (2008). Farming the planet: 2. geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, 22(1).

Moscona, J. (2019). Agricultural development and structural change, within and across countries. Technical report.

Nagaraj, D., Proust, E., Todeschini, A., Rulli, M. C., and D'Odorico, P. (2021). A new dataset of global irrigation areas from 2001 to 2015. *Advances in Water Resources*, 152:103910.

Ngai, L. R. and Pissarides, C. A. (2007). Structural change in a multisector model of growth. *American Economic Review*, 97(1):429–443.

Nunn, N. and Qian, N. (2011). The Potato's Contribution to Population and Urbanization: Evidence From A Historical Experiment*. *The Quarterly Journal of Economics*, 126(2):593–650.

Nurkse, R. (1953). *Problems of Capital Formation in Underdeveloped Countries*. Oxford University Press.

Pierce, J. R. and Schott, P. K. (2016). The surprisingly swift decline of us manufacturing employment. *American Economic Review*, 106(7):1632–62.

Qiu, T. and Luo, B. (2021). Do small farms prefer agricultural mechanization services? evidence from wheat production in china. *Applied Economics*, 53(26):2962–2973.

Qiu, T., Shi, X., He, Q., and Luo, B. (2021). The paradox of developing agricultural mechanization services in china: Supporting or kicking out smallholder farmers? *China Economic Review*, 69:101680.

Ranis, G. and Fei, J. C. H. (1961). A theory of economic development. *The American Economic Review*, 51(4):533–565.

Restuccia, D., Yang, D. T., and Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics*, 55(2):234–250.

Rostow, W. W. (1956). The take-off into self-sustained growth. *The Economic Journal*, 66(261):25–48.

Rozelle, S., Guo, L., Shen, M., Hughart, A., and Giles, J. (1999). Leaving china's farms: Survey results of new paths and remaining hurdles to rural migration. *The China Quarterly*, (158):367–393.

Schmidt, T. D., Jensen, P. S., and Naz, A. (2018). Agricultural productivity and economic development: the contribution of clover to structural transformation in denmark. *Journal of Economic Growth*, 23(4):pp. 387–426.

Schultz, T. (1953). *The Economic Organization of Agriculture*. McGraw-Hill.

Su, W., Eriksson, T., Zhang, L., and Bai, Y. (2016). Off-farm employment and time allocation in on-farm work in rural china from gender perspective. *China Economic Review*, 41:34–45.

Tan, S., Heerink, N., and Qu, F. (2006). Land fragmentation and its driving forces in china. *Land Use Policy*, 23(3):272–285.

Tombe, T. and Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–72.

UNFPA (2019). Domestic migrants in China (2018): trends, challenges and recommendations. Technical report, UNFPA.

Vollrath, D. (2014). The efficiency of human capital allocations in developing countries. *Journal of Development Economics*, 108:106–118.

von der Goltz, J., Dar, A., Fishman, R., Mueller, N. D., Barnwal, P., and McCord, G. C. (2020). Health impacts of the green revolution: Evidence from 600,000 births across the developing world. *Journal of Health Economics*, 74:102373.

Wang, D. (2008). Rural-urban migration and policy responses in China : challenges and options. Technical report.

Wang, X., Yamauchi, F., Huang, J., and Rozelle, S. (2020). What constrains mechanization in chinese agriculture? role of farm size and fragmentation. *China Economic Review*, 62:101221.

Zhang, X., Yang, J., and Thomas, R. (2017). Mechanization outsourcing clusters and division of labor in chinese agriculture. *China Economic Review*, 43:184–195.