Financial Frictions and Agricultural Productivity Differences

Junmin Liao and Wei Wang∗

December 28, 2014

Abstract

We explore the role of financial frictions in accounting for labor productivity differences across provinces in China using a quantitative sectoral model featuring non-homothetic preference, intermediate input and financial frictions. We explore household-level data and find financial frictions exist in rural area and are more severe in poorer regions. Limited credit in poor area depresses the use of intermediate inputs and hence encourages the use of labor inputs. Quantitatively, financial frictions alone explain more than 1/4 of the observed employment share and productivity differences. Moreover, financial frictions amplify the effect of TFP differences on agricultural productivity differences by 30 percent.

JEL: O11, O13, O4, E00, Q14

Keywords: Rural financial frictions, agricultural productivity, agriculture, China.

∗Liao: Department of Economics, Washington University in St. Louis, Campus Box 1208, 1 Brookings Drive, St. Louis, MO 63130 (e-mail: liaojunmin@wustl.edu); Wang: Department of Economics, Washington University in St. Louis, Campus Box 1208, 1 Brookings Drive, St. Louis, MO 63130 (e-mail: wwangc@wustl.edu). We are indebted to Rody Manuelli, Costas Azariadis, Yongs Shin for their guidance and support. We thank Diego Restuccia, Mark Rosenzweig, Mike Waugh and Xiaodong Zhu for very helpful comments. We also thank Jiahua Che, Douglas Gollin, B. Ravikumar, Roberto Samaniego, Todd Schoellman, Robert Tamura, and Ping Wang for helpful suggestions and discussions. We have benefited from constructive comments by seminar participants at Washington University in St. Louis, 2013 Spring Midwest Macro Conference, 2013 China Meeting of the Econometric Society, 2013 Annual Meeting of the Society for Economic Dynamics, 2013 Tsinghua Workshop for Macroeconomics, 2014 North American Winter Meeting of the Econometric Society, 2014 Chinese Economist Society North America Annual Meeting, and 2014 North America Summer Meeting of the Econometric Society. We are grateful to Tsinghua University and SWUFE for data sharing and Washington University in St. Louis for financial support. All remaining errors are our own.
1 Introduction

A growing literature documents that agricultural productivity differences can account for a substantial fraction of the large aggregate productivity differences and hence the income differences across countries (see Caselli (2005); Restuccia, Yang and Zhu (2008); Gollin, Lagakos and Waugh (2014), among others). Since to understand the large income differences across countries is one of the most important issues in development economics, theoretical and quantitative frameworks that could rationalize huge agricultural productivity differences are needed. To gain some insight to this issue, this paper explores the role of financial frictions in explaining the agricultural productivity differences.

The main economic mechanism is as follows. Farmers in poor area do not have sufficient funds to purchase intermediate inputs such as fertilizers and chemicals. They have to borrow from financial intermediaries. Severe financial frictions keep the purchase of intermediate inputs for agricultural production at a low level. On the one hand, farmers have to rely more on labor inputs when they do not have adequate intermediate inputs. On the other hand, demand for non-agricultural good and hence non-agricultural labor are low, meaning employment share in agriculture is large. Therefore, more severe financial frictions depress the use of intermediate inputs and encourage the use of labor in agriculture. As a consequence, in poor area, labor productivity in the agricultural sector is lower and moreover, due to a large employment share in agriculture, aggregate labor productivity is lower as well. Financial friction differences thus are able to account for the productivity differences in the agricultural sector and the aggregate economy.

Instead of studying cross-country agricultural productivity differences, we focus on cross-province productivity differences in China. To this end, we construct a novel data set to document a set of empirical regularities for Chinese economy. The challenge is that official provincial data are not comparable. The cross-province sector-level data we construct features the following: 1) PPP-adjusted provincial sectoral GDPs; 2) adjusted provincial agricultural employment. We then document that, in 2007, aggregate and agricultural labor productivity differences between the richest and the poorest 10 percent of the provinces are 5.9-fold and 5.3-fold. Both differences are very large within a country. However, labor productivity differences in the non-agricultural sector is merely 3.8-fold, which is much smaller.

1 This is what Schultz (1953) called “food problem”.
2 Lack of sector-level cross-country data which are needed to measure financial frictions is the main reason that we focus on a cross-region study. Moreover, Acemoglu and Dell (2010) argues that the sources of within-country and across-country are related.
3 The official provincial data of nominal variables from the National Bureau of Statistics of China (NBS) is not comparable among provinces since they are not adjusted by the Purchasing Power Parity (PPP). Moreover, official data of employment in each sector in each province are problematic, as described in Brandt, Tombe and Zhu (2013). Therefore, we construct our own data set to overcome these issues. See Appendix A.1 for the details.
Moreover, poorest provinces allocate 44 percent of employment to agriculture while rich provinces allocate only 6 percent. Therefore, aggregate productivity difference is mainly due to the fact that poor provinces allocate a large fraction of their employment to the low-productivity agricultural sector.

We conduct empirical and development accounting exercises and find that intermediate input is an important factor to account for productivity differences. First, we empirically document that there is a significant positive correlation between agricultural intermediate input-output ratio and agricultural labor productivity. This correlation is still significant after controlling for other factors such as physical capital, human capital, lands, etc. Second, development accounting exercises show that physical capital, human capital and lands are not quite successful in accounting for agricultural productivity differences, though physical capital is more important in non-agriculture. The evidence together shows that intermediate input is important to account for agricultural productivity differences.

To study the role of financial frictions, we mainly explore the data from the 2006 Farmer’s Credit Situation Survey (FCSS), 2005-2007 Rural Finance Development Survey (RFDS), and 2011 China Household Finance Survey (CHFS). All of them are household-level surveys on financial situation of households and hence provide the information about financial frictions that restrict borrowing ability of rural households (farmers) in China. One of the empirical contributions of this paper is that we document several empirical facts about the financial situations of rural households in China by exploring the aforementioned surveys. First, the fraction of rural households that need to borrow is from 53 to 57 percent based on different surveys, which implies that a large friction of rural households need external finance. Second, the expenditure share of intermediate inputs, which is the largest, takes more than 50 percent in total rural household expenditure. Third, financial frictions differ a lot across regions at both county and province level. To this end, we construct our measure for the degree of financial frictions (DFF), which is defined as the fraction of households who applied bank loans but did not get full amount out of all households that need external credits. DFF varies from a number slightly larger than 0 to almost 100 percent at county level and from 50 percent to almost 90 percent at province level. This implies that financial frictions are severe in some places while does not exist at all in some other places. These facts together mean that a large fraction of households need external funds to purchase intermediate inputs for agricultural production, but most of them are financially constrained in the sense that they are financially constrained in the sense that they are

---

4 The main reason that we focus on 2007 is the data availability.
5 See Appendix A.2 for detailed descriptions about these datasets.
6 For example, among 12257 rural households in 10 provinces, 57 percent report that they need to borrow, according to the data from FCSS. RFDS data implies that in the year 2007, 53 percent of rural households in the sample report the same.
7 The second and third largest expenditure are children education and health expenditure, according to the data from RDFS.
not able to get sufficient funds that they need to purchase intermediate inputs. The variation of financial frictions across different regions has great potential to account for productivity differences.

We then study the relationship between financial frictions and macro variables that related to productivity: intermediate input ratio, agricultural productivity and aggregate productivity. Combining the above empirical facts on rural financial frictions and macro data of corresponding regions, we document a couple of cross-province empirical facts regarding intermediate inputs used for agricultural production and financial frictions in the agricultural sector. First, there is a significant negative correlation between DFF and agricultural intermediate input use. Second, a significant negative correlation between DFF and agricultural labor productivity is also reflected by the data. Finally, we document a significant negative correlation between DFF and aggregate productivity. Moreover, these correlations are quite robust. Since poor farmers have to buy intermediate inputs through external finance, these empirical facts suggest that financial frictions are potentially quite important to account for agricultural labor productivity differences.

A two-sector general equilibrium model with financial frictions is constructed to explain and quantify the importance of financial frictions in agricultural labor productivity differences. There are three main ingredients in the model. The first one is intermediate inputs for agricultural production. The second one is that there is a subsistence consumption requirement of agricultural goods due to non-homothetic preferences. The last ingredient is financial frictions: people can not borrow as much as they need to purchase intermediate goods. Our model is then calibrated to match the data in Beijing which is our benchmark economy. In various experiments, we keep the value of most parameters unchanged and vary one or more of the following parameters: economy-wide efficiency, agriculture-specific efficiency, land-to-employment ratio, labor market friction parameter and financial friction parameter. These five parameters are calibrated to match the data in corresponding provinces when they are varied.

In our baseline experiment, we vary economy-wide efficiency (TFP) and financial frictions, keeping other parameters unchanged. Quantitative exercises show that economy-wide ef-

---

8To the best of our knowledge, cross-country agricultural financial data are rare if not unavailable. We use cross-province data from China instead to investigate the relationship between financial frictions and intermediate input usage. Note that cross-county investigation might be more satisfactory. We choose not to do that simply because reliable country-level aggregate data are not available.

9These correlations still hold significantly after adding control variables such as lands and human capital. We also use loan rejection rate and loan-output ratio in agriculture as other measures for the degree of financial frictions for robustness check. Results are quite similar.

10Economy-wide TFP difference will be able to generate the same patterns. Admittedly, TFP difference is able to account for a large fraction of productivity differences in China, mainly because provinces within a country have less heterogeneity than different countries. However, there is still a large fraction of productivity differences that can not be explained by TFP difference.
efficiency and financial frictions are able to account for almost all the differences in agricultural employment share, agricultural labor productivity, and aggregate labor productivity. To show how important are financial frictions, we conduct an experiment in which only the financial frictions parameter is varied. Quantitative results show that financial frictions alone are able to produce 27 percent (agricultural employment share increases from 4 percent to 7 percent), 32 percent and 1.6 percent of the observed elasticities of agricultural employment share, agricultural labor productivity and aggregate labor productivity, with respect to agricultural intermediate input-output ratio. The reason that we can only explain a small fraction of aggregate labor productivity is the income effect is small when we start from a rich benchmark economy (high TFP). If we start from the poorest province (low TFP) and conduct the same experiment, financial frictions alone can generate 30 percent (agricultural employment share decreases from 62 percent to 34 percent), 32 percent and 25 percent of the observed elasticities of agricultural employment share, agricultural labor productivity and aggregate labor productivity, with respect to intermediate input ratio.\textsuperscript{11}

We then vary TFP and financial frictions together. For each province, we calibrate TFP and financial frictions to match the data of that province. Our model is able to account for more than 80 percent of the observed differences in agricultural and aggregate productivity. Comparing the results from this experiment and the experiment in which only TFP is calibrated to match the data of corresponding province, we find that financial frictions amplify the effect of TFP differences substantially when accounting for agricultural productivity: financial frictions are able to explain 30 percent more. Relative prices of agricultural good are not targeted in our quantitative exercises. However, they match data quite well.

The contribution of this paper is three-fold. First, we construct a novel data set featuring PPP-adjusted provincial GDPs in China using standard Geary-Khamis method. Second, based on several household-level surveys, we document certain important empirical facts of financial frictions in rural area across provinces in China. We then derive correlations of labor productivity differences, financial friction differences, and agricultural intermediate input differences across provinces in China. Finally, we quantify the importance of financial frictions in explaining the large agricultural labor productivity differences across provinces in China.

**Related literature** This paper joins a growing literature trying to account for large per capita income difference across different regions (Klenow and Rodriguez-Clare (1997); Hall and Jones (1999); Parente and Prescott (2002); Caselli (2005)). Accounting exercises by Caselli (2005), Restuccia, Yang and Zhu (2008), and Gollin, Lagakos and Waugh (2014) show that

\textsuperscript{11}We have data from all provinces including poorest provinces so that we are able to conduct experiments where the poorest province is the benchmark economy.
low per capita income in poor regions is driven by low agricultural productivity and large agricultural employment share in poor area. Several important works have been completed. Restuccia, Yang and Zhu (2008) explores the role of intermediate inputs and emphasizes the substitution effect of indirect friction on the labor market. Lagakos and Waugh (2013) considers occupational choice as an amplification mechanism of economy-wide efficiency difference. Adamopoulos and Restuccia (2014) emphasizes policy distortion on size distribution of farms. Gollin, Parente and Rogerson (2004) and Herrendorf and Schoellman (Forthcoming) attribute the difference to mismeasurement due to home production or under-estimation of hours worked. Donovan (2013) investigates the interaction of non-homothetic preferences and incomplete markets. Frictions on infrastructure is studied in Adamopoulos (2011).\footnote{Other work on this issue includes Priyo (2012) which introduces physical capital and human capital into Restuccia, Yang and Zhu (2008)'s model as exogenous parameters.}

To the best of our knowledge, our paper is the first one that studies financial frictions in agriculture to understand the large productivity differences across regions.\footnote{There is a large literature investigating the effect of financial frictions on non-agriculture. See Greenwood, Sanchez and Wang (2010); Buera, Kaboski and Shin (2011); Buera and Shin (2011); Moll (2014); Buera and Shin (2013); Allub and Erosa (2013); Itskhoki and Moll (2013) among others.} The most related paper is Restuccia, Yang and Zhu (2008). Our paper differs from theirs by emphasizing financial frictions, instead of frictions on labor market, as an important factor to account for low labor productivity.

This paper is organized as follows. Section 2 provides motivating empirical facts on financial frictions, agricultural intermediate inputs, and productivity. The model is described and characterized in Section 3. Calibration and quantitative findings are presented in Section 4. Finally, Section 5 concludes. Data descriptions, proofs and model extensions are included in the appendix.

\section{Empirical Underpinnings}

This section documents several key empirical facts as our motivation. We first show some empirical facts about cross-province differences in agricultural and aggregate productivity and employment share in agriculture. We then show a positive correlation between agricultural labor productivity and intermediate input. A development accounting exercise is conducted showing intermediate input is important to account for productivity differences. We also provide evidence that rural households face severe credit constraints and the variation of rural financial frictions is large across provinces. We finally show, using both micro and macro data, that rural financial frictions are the key to explain different agricultural intermediate input shares and agricultural labor productivity.
2.1 Agricultural and Aggregate Labor Productivity in China

We construct PPP-adjusted GDPs of each province in both agriculture and non-agriculture. Moreover, as Brandt, Tombe and Zhu (2013) points out, official employment data from NBS are problematic. Hence, we modify employment of each province in both sectors by official census data.\textsuperscript{14}

Table 1: Labor Productivity Gap across Provinces in China

<table>
<thead>
<tr>
<th>Decile</th>
<th>GDP/N</th>
<th>GDP\textsubscript{a}/L\textsubscript{a}</th>
<th>GDP\textsubscript{n}/L\textsubscript{n}</th>
<th>L\textsubscript{a}/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1\textsubscript{st}</td>
<td>102027.9</td>
<td>23619.1</td>
<td>106584.9</td>
<td>0.06</td>
</tr>
<tr>
<td>10\textsuperscript{th}</td>
<td>17333.3</td>
<td>4490.6</td>
<td>27818.4</td>
<td>0.44</td>
</tr>
<tr>
<td>1\textsuperscript{st} decile</td>
<td>5.9</td>
<td>5.3</td>
<td>3.8</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: Provinces are ranked based on observed aggregate GDP per worker descendingly and divide them into 10 deciles, each with 3 provinces. The decile ranking is preserved for the comparisons of agricultural GDP per worker, non-agricultural GDP per worker and agricultural employment share. Hainan and Tibet are excluded due to lack of price data hence the sample consists of 29 provinces.

Source: Various issues of China Statistical Yearbook. PPP-adjusted GDP and output data are from authors’ own calculation based on the Geary-Khamis method. Employment data is adjusted based on the method in Brandt and Zhu (2010). See Appendix A.1 for details.

Table 1 shows the aggregate, agricultural and non-agricultural GDP per worker (measured in RMB, PPP adjusted) and agricultural employment share across provinces in China in 2007. The third row shows that the average GDP per worker of the richest decile is about 6 times that of the poorest decile. The log-variance of GDP per worker for the whole sample is 0.25. By decomposing the total GDP per worker, we find that the agricultural labor productivity gap is larger than that of the non-agricultural sector. The gap in the agricultural sector is a factor of 5.3 while the gap in the non-agricultural sector is a factor of 3.8. Besides labor productivity, there is also a large gap in agricultural employment share. Average agricultural employment share of the poorest decile is almost 50 percent while that number of the richest decile is only 6 percent.\textsuperscript{15}

Simple counterfactual analyses would show the importance of labor productivity and em-

---

\textsuperscript{14} A detailed description on how we construct these data is in Appendix A.1.

\textsuperscript{15} Results from the official GDP and employment data is as follows.

<table>
<thead>
<tr>
<th>Decile</th>
<th>GDP/N</th>
<th>GDP\textsubscript{a}/L\textsubscript{a}</th>
<th>GDP\textsubscript{n}/L\textsubscript{n}</th>
<th>L\textsubscript{a}/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1\textsubscript{st}</td>
<td>6.81</td>
<td>3.63</td>
<td>3.71</td>
<td>10%/57%</td>
</tr>
</tbody>
</table>

Agricultural productivity difference is much smaller than what we obtain from PPP-adjusted data. Aggregate productivity difference is larger. Productivity differences in agriculture and non-agriculture are almost the same. This is in contrast to the cross-country evidence documented by Gollin, Lagakos and Waugh (2014). As shown in Table 1, the gap between factor differences of agricultural and non-agricultural productivity obtained from the PPP-adjust data is 1.4, which is comparable to the results in their paper. Official agricultural employment share in the 10\textsuperscript{th} decile is much larger than what we derive from the census data. To sum up, it is necessary to fix the two problems in the official data, otherwise the differences we are trying to explain is misleading.
ployment share in the agricultural sector. Suppose all the provinces have the same highest agricultural productivity, then the $1_{st}/10_{th}$ decile ratio of aggregate GDP per worker would be 3.9 and log-variance would be driven down from 0.25 to 0.15. If we assume all the provinces has the same smallest agricultural employment share, then the $1_{st}/10_{th}$ decile ratio would be driven down to 4 and log-variance would be 0.16. Therefore agricultural labor productivity and employment share play important roles in explaining aggregate labor productivity. In the next section we will discuss the possible reasons that could result in the agricultural labor productivity differences.

2.2 Agricultural Intermediate Input Shares across Provinces

One important reason to explain the agricultural productivity differences is the disparate usage of intermediate non-agricultural inputs in the agricultural sector, such as fertilizer, pesticide, etc., which indicates the modernization level of agricultural production. Restuccia, Yang and Zhu (2008) shows that there is a positive correlation between final output per worker and intermediate input to output ratio in the agricultural sector across countries. This is also true according to provincial data in China.

Figure 1 shows the significant positive correlation between final output per worker and intermediate input output ratio in the agricultural sector across provinces in China. The horizontal axis is agricultural intermediate input output ratio ($X/Y_a$) and the vertical axis is relative agricultural final output per worker ($Y_a/L_a$). The slope is 0.72 ($R^2 = 0.52$). The poorest province, Guizhou, has a comparable arable land area and agricultural employment with one of the richest provinces, Jiangsu, but they differ largely in the use of intermediate non-agricultural input, as well as the agricultural output per worker. In Appendix A.3.1 we test this relationship by controlling other variables such as land quality and human capital. The result is still significance.

2.3 Development Accounting

The differences in the use of factor inputs such as physical capital, human capital, or land across provinces could result in different labor productivity. In this section we try to deter-

---

16 Through out this paper by agricultural intermediate inputs we refer to the intermediate non-agricultural inputs in the agricultural sector, which is denoted by $X$.

17 See Appendix A.1 for the descriptive statistics of intermediate input data across provinces. We do not have price data on agricultural intermediate input, hence the agricultural intermediate input-output ratio ($X/Y_a$) actually contains the price information of intermediate input. Since $X$ is produced by the non-agricultural sector, with this relative price we can measure $X$ in terms of agricultural goods (see Lagakos and Waugh (2013) for a similar treatment when comparing productivity across sectors). It is also necessary to keep this price information to investigate the role of financial friction because farmers have to consider the price of intermediate input when making production decision.
Figure 1: Intermediate Input Output Ratio vs Relative Agricultural Output per Worker

Note: The largest agricultural final output per worker is normalized to be 1.
Source: 2007 China Regional Input-Output Tables and authors’ own calculation based on various issues of China Statistical Yearbook.

We follow Caselli (2005) to mainly examine two measures of the “success” of factor inputs in explaining agricultural labor productivity differences across provinces. $success_1$ is defined as the ratio of log variance in output per worker generated by the factor-only model divided by the real log variance in the data. $success_2$ is defined as the 90-10 ratio of output per worker generated by the factor-only model with the observed 90-10 ratio. Table 2 reports the results.

Our results suggest that none of the factor inputs are crucial to explain agricultural labor productivity differences across provinces and therefore other factors, such as intermediate inputs, which are abstracted away from the simple model are important. In Table 2 we can see even if we include all the main factor inputs, we find a ratio of 0.38 for $success_1$ and 0.49 for $success_2$, both leave large unexplained fractions. For comparison, we also conduct similar analysis for the non-agricultural sector. The result shows that although physical capital is more important in the non-agricultural sector than in the agricultural sector, it still leaves a large room for TFP, which is consistent with the results in Caselli (2005) and Lagakos and Waugh (2013).
Table 2: Development Accounting

<table>
<thead>
<tr>
<th>Sector</th>
<th>Factor</th>
<th>success&lt;sub&gt;1&lt;/sub&gt;</th>
<th>success&lt;sub&gt;2&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>$K_a/L_a$</td>
<td>0.14</td>
<td>0.34</td>
</tr>
<tr>
<td>Agriculture</td>
<td>$Z/L_a$</td>
<td>0.10</td>
<td>0.31</td>
</tr>
<tr>
<td>Agriculture</td>
<td>$h$</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>Agriculture</td>
<td>$K_a/L_a$, $Z/L_a$, and $h$</td>
<td>0.38</td>
<td>0.49</td>
</tr>
<tr>
<td>Non-agriculture</td>
<td>$K_n/L_n$</td>
<td>0.39</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Note: $K_a/L_a$ means capital per worker in agriculture; $Z/L_a$ means land per worker in agriculture; $h$ means human capital per worker in agriculture; and $K_n/L_n$ is capital per worker in non-agriculture.

Source: Authors’ own calculation. Provincial sectoral capital is from Wu (2009). Agricultural human capital data is calculated from the year of schooling data in 2005 The Second National Agricultural Census.

2.4 Rural Financial Frictions

In this section, we describe how we measure financial frictions. Essentially, we take the fraction of households who are financially constrained as our major measure for financial friction within a certain area. The main challenge here is to identify households that are financially constrained. We explore data from several surveys to provide some empirical facts about the financial situations in rural China. This will guide us on how we construct our measure of the degree of financial frictions (DFF) in the next section.

There are two major channels to borrow: from the Rural Credit Cooperative (RCC) and from family and friends. Almost all the borrowings in each province in the survey are from these two channels. Each of them consists of about a half of the total borrowing. The former one is formal and we call the latter one informal channel. Formal finance is more expensive and restrictive than informal finance. The difference of borrowing costs between the two channels is very large.

Due to this feature, if possible, rational rural households tend to borrow from their family members or friends at the first place. If they can not borrow enough funds from the informal channel, they have to come to the RCC. If they could not get enough funds, they are financially constrained. Therefore, it is reasonable to assume that households...

---

<sup>18</sup> According to the data from Rural Financial Development Survey (RDFS) data in 2007, in 8 provinces where the survey was conducted, 90 percent to 95 percent (in different provinces, with a mean of 93.4 percent) of the households that get external funds borrow from the RCC and family and friends. Therefore, we could essentially ignore the other way of financing.

<sup>19</sup> Based on Survey on China’s Rural Finance (SCRF) conducted in 2005, Han (2007) documents that the average loan rates from the RCC and the informal channel are 6.42 and 1.11 percent, respectively. While the informal channel involves no time delay, the formal channel has a long time delay due to background verification. Moreover, the RCC loans usually requires collateral that rural households may not have. 67 percent of the RCC loans involves collateral in the sample of the survey (Han, 2007). Another point is that every RCC loan has a clear due date, in contrast to the informal case. All these differences between the two channels are largely due to that family and friends have connections and perfect information while the RCC does not. The main disadvantage of the informal channel for rural households is the volume of can not be large. Hence, borrowing from the informal channel may not be enough.
that do not apply loans are financially unconstrained. Therefore, we only focus on financial frictions result from formal channel.

In this paper, we do not consider financial frictions in the non-agricultural sector. Recently, Brandt, Hsieh and Zhu (2008) estimates that in 2004 the average TFP in the non-state sector is 80 percent higher than that of the state sector for non-agriculture. Hsieh and Klenow (2009) estimate that, within the manufacturing sector, the average revenue TFP (TFPR) of state-owned firms is about 40 percent lower than that of private firms during the time period from 1998 and 2005. Song, Storesletten and Zilibotti (2011) argue that the TFP gap between state-owned firms and private firms results from the fact that state-owned firms are well-connected to the government and hence have better access to external credit while private firms do not and hence severely financially constrained. We argue that even though the non-agricultural sector in China is subject to financial frictions, in particular private firms, financial frictions may not vary across provinces. Hsieh and Klenow (2009) find that average TFPR levels in manufacturing differ modestly (within 10 percent) across Chinese province. Since TFPR reflects the underlying distortions including financial frictions, financial frictions in manufacturing is not likely to be an important factor accounting for cross-province income differences.

2.5 Rural Financial Frictions and Agricultural Labor Productivity

We explore the data from several novel household-level surveys to study the relationship between rural financial frictions and agricultural labor productivity. One is Farmer’s Credit Situation Survey (FCSS) which was conducted by the People’s Bank of China in year 2006.

Table 3: Descriptive Statistics of Degree of Financial Frictions

<table>
<thead>
<tr>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Coefficient of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.505</td>
<td>0.872</td>
<td>0.747</td>
<td>0.792</td>
<td>0.111</td>
<td>0.148</td>
</tr>
</tbody>
</table>

Source: Authors’ own calculation based on 2006 Farmer’s Credit Situation Survey.

Households were asked whether they need external credits, whether they can get credits from the bank and whether they can get enough credits. The degree of financial frictions (DFF) is defined as the number of households who apply for loans but are not able to get enough credits divided by the number of households who need loans.\(^{20}\) We use DFF as

\(^{20}\)See Appendix A.2 for the details.
an index of rural borrowing constraint.\textsuperscript{21} DFF ranges from zero to one and larger number means tighter borrowing constraint (more severe financial frictions). The sample has ten provinces, which are representative in terms of geography and income distribution. The descriptive statistics of the ten-province DFF are reported in Table 3. From the table we can see the differences in DFF are not small and have the potential to explain the differences in agricultural input output ratio and agricultural labor productivity.

![Figure 2: Degree of Financial Frictions vs Agricultural Sales Income per Worker](image)

*Figure 2: Degree of Financial Frictions vs Agricultural Sales Income per Worker*

*Note: The sales income per worker of provincial aggregation is PPP-adjusted based on authors’ own calculation. Source: Authors’ own calculation based on 2006 Farmer’s Credit Situation Survey.*

Figure 2 shows a significant negative correlation between agricultural sales income per worker with the DFF in the micro data. At the county level (left panel), the coefficient for DFF is significant at 0.1 percent level with a sample size of 151. From the right panel of Figure 2 we can see that rural households in a rich province, such as Jiangsu, is less likely to be credit constrained, with a DFF of 0.698. While 78.6 percent of rural households who need loan are credit constrained in the poorest province, Guizhou. The DFF also shows a negative correlation with the agricultural intermediate input-output ratio in the aggregate data (see Figure 3), which means provinces with a high degree of financial friction are more likely to have a low intermediate input-output ratio.\textsuperscript{22}

\textsuperscript{21}Instead of using the total sample size as the denominator of the DFF, in this paper we use the number of farmers who need external funds. Feder et al. (1990) used a similar measure.

\textsuperscript{22}In Appendix A.3.2 we test this relationship by controlling human capital and land size. The result is still significant.
We also check for 2011 China Household Finance Survey (CHFS) which covers more provinces (but with much less households in each province and less detailed questions on agricultural loan, see Appendix A.2). Unfortunately, this survey does not ask questions about whether farmers need loans or whether they can get enough loans. So we can only use the loan rejection rate, which is defined as the number of farmers who have been rejected by the bank over the number of farmers who have applied for loans, to measure borrowing constraint. Figure 4 shows the correlation between the loan rejection rate and three macro variables that we are interested in. We can see the significant negative correlation quite clear.

In addition, the correlation between agricultural financial frictions and the use of agricultural intermediate inputs is also preserved in aggregate data. We use macro data, in particular provincial total short-term agricultural loan over agricultural final output as another measure of borrowing constraint. This agricultural loan refers to the short-term (less than or equal to one year) loan issued by all financial intermediaries only for the production of agriculture, forestry, animal husbandry and fishery. The loan might be used to purchase agricultural intermediate inputs, non-agricultural intermediate inputs, agricultural capital, or employment expenditure. Although we cannot distinguish the detailed purposes, this agricultural loan-output ratio can still measure financial frictions because it is reasonable to assume all specific activities within agricultural production would suffer similar borrowing constraints. Figure 5 shows that significant positive correlation between this measure and

---

23 In Appendix B.3, we use this measure as an indirect target to calibrate the financial friction parameter. See Figure 18.
In sum, we document that rural financial friction potentially has a strong impact on the intermediate input use in agriculture. Therefore, variation of rural financial frictions could lead to differences in agricultural employment share and productivity in different areas.

### 2.6 Other Micro Evidence on Financial Frictions and Productivity

Recently, many experimentalists conduct randomized controlled experiments in developing countries to study how important is credit constraint on investment decisions in agriculture. The evidence from these studies is mixed in the sense that some support that financial frictions are important while the others do not. Notice that these experiments are all conducted outside of China. No such experiments have been conducted in China. Nevertheless, based on a 10-province household-level panel data, Shen and Wang (2014) show that the

---

24Since credits are usually needed at the beginning of each production cycle, we use agricultural short-term loan reported at the end of 2006 (which should be paid back in 2007) and output data in 2007.

25Banerjee and Duflo (2014) shows that medium-size Indian firms had been severely credit constrained and a relax of credit constraint leads to a huge expansion of production. Crepon et al. (2011) concludes that more credit results in a decrease of consumption and an increase in agricultural production of rural households in Morocco. Kaboski and Townsend (2012) finds that an improvement of credit constraint increases agricultural investment in Thai villages. Duflo, Kremer and Robinson (2011) argues that the improvement of factors that facilitate the purchase of fertilizer and are orthogonal to financial frictions, leads to a large increase of fertilizer use in rural Kenya. Karlan et al. (2013) conducted experiments in rural Ghana and show that credit constraint is not likely to cause low expenditure in agriculture production.

---

Figure 4: Loan Rejection Rate (CHFS)

*Source: Authors’ own calculation from 2011 China Household Finance Survey.*
elimination of the lump-sum agricultural tax increases intermediate inputs use and labor productivity while decreases the use of labor, which is consistent with the mechanism in our paper.  

3 The Model

Our model economy has two sectors: an agricultural sector and a non-agricultural sector. The two sectors produce two final goods respectively: an agricultural good and a non-agricultural good. The output of the agricultural sector is used for consumption only and the output of the non-agricultural sector has to be used for agricultural good production in addition to consumption. The model economy is populated by a continuum of mass $N$ homogeneous households. Assume that labor market is competitive but there are labor market distortions that in effect incur a cost of reallocating labor from the agricultural sector to the non-agricultural sector. Both final goods markets are competitive. There exists a representative financial intermediary which behaves competitively and has a deep pocket. The

---

26 Starting from 2003, the Chinese government reduced agricultural taxes gradually and completely eliminated agricultural taxes on the first day of 2006. One important characteristic of agricultural tax is that they are essentially lump-sum. Agricultural tax is a non-trivial burden for rural farmers in China: 7 percent of per capita agricultural income. The abolition of agricultural tax should immediately increase rural households’ income and may have impact on agricultural production.

---

Figure 5: Agricultural Debt Output Ratio (log-scale)

Source: Authors’ own calculation from China Compendium of Statistics 1949-2008.
economy coincides with a province in China.\textsuperscript{27}

### 3.1 Technologies

The non-agricultural good is produced by labor only:

\[ Y_n = AL_n, \]

where \( Y_n \) and \( L_n \) are output of non-agricultural good and labor input, respectively.\textsuperscript{28} Parameter \( A \) is the economy-wide productivity. We adopt linear production function due to the lack of accurate capital data.\textsuperscript{29} Denote the wage rate in non-agricultural sector by \( w_n \). Profit maximization of the representative firm in non-agricultural sector implies \( w_n = A \). The non-agricultural good is used as numeraire.

The agricultural production uses both intermediate non-agricultural inputs and labor. The production function is

\[ Y_a = X^\alpha \left( Z^{1-\sigma} (\kappa AL_a)^\sigma \right)^{1-\alpha}, \quad 0 < \sigma < 1, 0 < \alpha < 1, \kappa > 0, \]

where \( Y_a \) is agricultural output and \( L_a \) is labor input.\textsuperscript{30} \( Z \) is land and \( X \) is the intermediate inputs supplied by the non-agricultural sector. As Restuccia, Yang and Zhu (2008) argues, this intermediate inputs may consist of chemical fertilizers, pesticides, hybrid seeds, fuel, energy and other purchased factors produced by the non-agricultural sector. The economy-wide productivity \( A \) and agricultural specific productivity parameter \( \kappa \) are both labor augmenting.

\textsuperscript{27}We treat each province as a closed economy and there is no trade among provinces for simplicity and tractability. We adopt this setting for several reasons. First, although inter-province trade is common, there are still large differences in producer’s price across different provinces, meaning there are trade barriers. For example, the highest producer’s price of rice is more than twice as large as the lowest one. Second, there are labor barriers (such as Hukou system) across provinces making labor difficult to move. This is reflect by provincial real wage differences (see Candelaria, Daly and Hale (2009)). What we emphasize is the role of financial frictions in accounting for cross-province productivity differences given that there are trade barriers and labor mobility frictions across provinces. Quantitatively, in our model, trade barrier and labor mobility frictions as well as provincial TFP differences together can only account for 50 percent of the observed agricultural productivity difference across provinces. This leaves a large fraction of difference in the data not explained.

\textsuperscript{28}The subscript \( n \) denotes non-agriculture and the subscript \( a \) denotes agriculture.

\textsuperscript{29}This is consistent with the literature, such as Restuccia, Yang and Zhu (2008) and Lagakos and Waugh (2013). In Appendix D, we introduce physical capital into the baseline model to take into account the role of physical capital and see how quantitative results change when the marginal product of labor is diminishing. Quantitative results show that a large fraction of productivity differences could be accounted for by financial frictions even after physical capital is taken into account. Similar results are also obtained in Caselli (2005), Lagakos and Waugh (2013). Priyo (2012) arrives at the conclusion that labor market distortion outweights the role of both human and physical capital in accounting for productivity differences, though he treats physical capital as a fixed factor. See also Chanda and Dalgaard (2008) and Vollrath (2009) for accounting for cross-country income differences through physical capital.

\textsuperscript{30}We use Cobb-Douglas production function which is broadly used in agriculture production and consistent with empirical findings. See Restuccia, Yang and Zhu (2008).
A general technology progress will have different effects on agriculture and non-agriculture. Agricultural specific productivity links efficiency of agriculture production and economy-wide productivity, which captures the specific effect of TFP progress in agriculture.\footnote{See Restuccia, Yang and Zhu (2008) for detailed interpretations of agricultural specific productivity.} Assume that land is a fixed factor in the production function so that labor and intermediate input both exhibit decreasing returns to scale.\footnote{According to Deininger and Jin (2005), the rental market for land in China is actually very small. Based on a survey data conducted in three provinces in China in 2000-2001, they document that the average arable land area per household rented out is 0.0055 mu (one fifteenth of a hectare), while in one of our survey data (FCSS) the average arable land area per household is about 11.2 mu (in 2005 The Second National Agricultural Census this number is 9.1 mu).} \(\alpha\) is the income share of intermediate input and \(\sigma\) captures the income share of labor. For simplicity, we assume that there is a representative farmer behaving competitively.

We assume that there is a wage gap between agriculture and non-agriculture. The gap may result from distortions in the labor market such as migration costs from agriculture to non-agriculture. Migration costs include transportation costs, more expensive living costs in cities, among others.\footnote{Agricultural productions take place in rural area while non-agriculture productions are mainly operated in urban or suburban area. Hence, the transportation cost and more expensive living costs (at least additional expenditure on housing rental) are obvious.} These obstacles usually suppress agricultural wages and we model the mobility cost as a fraction \(\theta\) of the wage rate in the non-agricultural sector, so the following non-arbitrage condition holds:

\[
    w_a = (1 - \theta) w_n, \quad 0 \leq \theta < 1. \tag{3}
\]

While labor market friction leads to a lower wage in agriculture, the wage gap between the two sectors is larger in poorer regions, as is reflected in the data (see Appendix B.1.1). Lower wages provide a strong incentive to use more labor and less intermediate inputs in agriculture production.

### 3.2 Preferences

There is a representative household in the economy. There are several family members in the household. There is a household planner who makes decisions for other family members. There are a farmer and a firm-manager who manage production and behave competitively in agriculture and non-agriculture, respectively. There are a continuum of mass \(N\) workers with unity labor endowment. At the beginning of each period, to conduct production activities, the firm-manager hire workers, and the farmer hire workers as well as use the funds borrowed from the financial intermediary to buy intermediate inputs. The productions then take place. After the two goods are produced, workers get paid, the loans as well as the in-
terest rate are repaid. Profits, if any, go to the farmer and the firm-owner. Finally, the planner collects all incomes from the farmer, the firm-owner, and the workers all together to make consumption decision.

The representative household gains utility from consuming the agricultural good $c_a$ and non-agricultural good $c_n$. Assume that the representative household owns the profits of agriculture production. Since labor endowment is supplied inelastically, the total labor supply (or aggregate employment) is equal to $N$. The preference of each household is represented by a Stone-Geary utility function, which implies Engel’s law:

$$U = a \log(c_a - \bar{a}) + (1 - a) \log c_n, \ 0 \leq a < 1,$$

where $\bar{a}$ is the subsistence level of consumption of agricultural good and $a$ is a utility weight over the two goods.

### 3.3 Financial Frictions

Financial frictions are considered to be barriers that keep farmers from substantively using intermediate inputs: labor market friction and financial friction. As motivated in Section 2, financial friction in agriculture is embedded in the model as follows. Before production, farmers (producers of agricultural goods) have to buy intermediate goods in advance. However, the farmer is assumed to be too poor to purchase any intermediate goods upfront. The only way to finance the purchase of intermediate goods is to borrow from the financial intermediary through financial contracts. The farmer have limited commitment and he could default. In particular, after the production, the farmer could renege on the financial contracts. In that case, only a fraction $1 - \lambda$ of the output net of wage payment could be kept by the farmer. Denote the exogenous interest rate, wage rate in the agricultural sector and the relative price of agricultural goods by $r^*, w_a$ and $p_a$. For any $X$, an incentive-compatible financial contract requires that the following condition hold:

$$\max_{L_a, X} \{ p_a Y_a - w_a L_a - (1 + r^*) X \} \geq (1 - \lambda) \max_{L_a} \{ p_a Y_a - w_a L_a \}.$$

The above condition states that a farmer must end up with no less economic resources when he fulfills his credit (left-hand side) than when he defaults (right-hand side). The degree of financial friction is captured by the parameter $\lambda$. Larger $\lambda$ means smaller default value and hence less severe financial friction in agriculture.\textsuperscript{34}

\textsuperscript{34}An alternative specification of financial constraint is provided in the Appendix C.1. In that specification, the first-order condition with respect to labor has a wedge between the marginal labor productivity and the wage rate for agricultural labor. This wedge could potentially generate labor productivity differences even if there is no distortion in labor market (no wage differences).
It is easy to check whether the financial constraint is binding or not. In particular, we have the following result.

**Proposition 1.** The farmer is financially constrained if and only if

\[ \lambda \leq \frac{\alpha}{1 - \sigma (1 - \alpha)}. \]  

(4)

Intuitively, as argued above, larger \( \lambda \) implies the degree of financial friction is smaller. If degree of financial friction is smaller enough (\( \lambda \) is larger than the above cutoff value), the farmer will not be financial constrained (optimal level of intermediate inputs will be attained).

### 3.4 Theoretical results

The following proposition states the effect of financial frictions on the several key variables that we are interested in when financial constraint is binding.

**Proposition 2.** When financial constraint is binding, a larger \( \lambda \) decreases the use of labor inputs \( L_a \) and increases intermediate input-labor ratio \( X / L_a \). More importantly, agricultural labor productivity \( Y_a / L_a \) increases as \( \lambda \) increases.

**Proof.** See Appendix C.3.

Consider an economy where the degree of financial frictions is very high. Intermediate input demand is depressed. If the financial constraint is relaxed, intermediate input must increase. This in turn leads to a decrease of labor input through general equilibrium effects: more labor is needed to produce more non-agricultural good to meet intermediate input demand and consumption demand (due to a higher income). Agricultural productivity increases because of a reduction of agricultural labor input and an increase of output due to an increase of intermediate input. Employment share in agriculture is smaller since more workers work in non-agriculture now.

Notice that relationships between other key variables, such as intermediate input use and hence total input, and financial frictions can not be determined analytically. We examine these correlations numerically in Appendix B.4. We also consider a special case where \( a = 0 \), meaning only the subsistence agricultural goods consumption is needed, i.e. \( c_a = \bar{a} \). Hence, income does not affect agricultural good demand, allowing us to derive analytical solutions and more comparative statics. The following proposition summarizes comparative statics results. Analytical solutions can be found in the Appendix C.2.1.
**Proposition 3.** When financial constraint is binding, a larger $\lambda$ implies a smaller use of agricultural labor inputs $L_a$, a larger use of intermediate inputs $X$, a smaller total input, a larger intermediate input-labor ratio $X/L_a$ and a higher agricultural labor productivity $Y_a/L_a$.

**Proof.** See Appendix C.5.

Total inputs have two components: labor and intermediate inputs. When financial frictions are less severe, the use of labor is smaller and the use of intermediate inputs is larger. A smaller total inputs in better financial situation implies that labor cost reduction is more than the increase of the cost due to larger intermediate inputs use. Intermediate input use will increase whenever there is an improvement of financial situation. The magnitude of increase depends on the demand of agricultural good. Non-agricultural good market clearing condition dictates that agricultural labor input must decrease. The magnitude of reduction depends on the demand of non-agricultural good. Under non-homothetic preference, when financial frictions are very severe, households spend most of their income in agricultural good to survive. When financial frictions become less severe and hence income becomes higher, relative expenditure on agricultural good decreases, meaning that non-agricultural consumption will increase more than agricultural consumption. In this special case, agricultural goods demand does not change at all. Consequently, the reduction of labor input cost is larger than the increase of intermediate input cost.

### 4 Quantitative Analysis

This section quantitatively analyzes the role of financial friction (captured by parameter $\lambda$) in accounting for labor productivity differences. We first calibrate most of parameters to match data in our benchmark province: Beijing. Then we conduct several quantitative analyses to show how our model is able to account for productivity differences in the data. Specifically, we vary $\lambda$ alone to match intermediate input share of each province and compute how much the key variables (such as $L_a/N$, $Y_a/L_a$, and $Y/N$) would change as $\lambda$ changes. The results show that financial frictions alone are able to explain substantial variation of agricultural employment share and agricultural labor productivity differences.\(^{35}\)

\(^{35}\)In Appendix B.1, we also vary other province-specific parameters to see to what extent financial frictions can explain the labor productivity differences across provinces.
4.1 Calibration

We use the data (PPP-adjusted) described in Section 2 to calibrate the model. We need to determine the value of the following 10 parameters: financial friction parameter $\lambda$, labor market friction parameter $\theta$, land over employment ratio $Z/N$, province-wide productivity $A$, agricultural specific productivity parameter $\kappa$, interest rate $r^*$, and the rest 4 parameters which capture the production technology and utility $\sigma, \alpha, a, \bar{a}$. The first 5 parameters may be different across provinces when we conduct quantitative experiments (variable parameters) while the other 5 parameters are fixed to be the same throughout this paper (fixed parameters). If an experiment requires the value of a variable parameter to be the same across provinces, it is set to be equal to the calibrated value in our benchmark economy. Otherwise, the value of that variable parameter in a particular province is calibrated to match the data of that province.

4.1.1 Fixed Parameters

We first calibrate parameters that are kept fixed in our quantitative analysis. Interest rate parameter $r^*$ is set to match the annual real interest rate in 2007, which is 0.0156. $\alpha$ is selected to match the highest intermediate input ratio (the income share of intermediate inputs) among the provinces in 2007. We get $\alpha = 0.5252$. By doing so, we can make sure all provinces are binding (the one with the highest intermediate input income share is indifferent between binding and not binding). The NBS input-output tables do not report land income share because there is no market for land. Chow (1993) and Cao and Birchenall (2013) estimate factor income shares based on production function regression and they achieved similar estimation on land income share. Here we choose the average of their estimation which is $1 - \sigma = 0.36$, hence $\sigma = 0.64$. Restuccia, Yang and Zhu (2008) use $\sigma = 0.7$ for cross-country study which is similar as ours.

To calibrate $a$ and $\bar{a}$, we follow Restuccia, Yang and Zhu (2008) and Lagakos and Waugh (2013) setting the target of long-run agricultural employment share to be 0.5 percent. First,

---

36Hainan and Tibet are excluded due to lack of price data.
37Our benchmark economy is Beijing in all but one experiment. In that exception, Guizhou is taken as our benchmark to investigate our model performance when income effect is large.
38See Appendix A.1.8. We assume the same interest rate across province because in China interest rate is highly regulated by the central bank. This value of $r^*$ might be conservative because farmers would be charged a higher interest rate by other informal financial intermediaries if they were rejected by the formal banks. However it will not affect our quantitative results because given our calibration strategy, $\lambda$ is proportional to $1 + r^*$ and the $X/Y_a$ generated by the model will not be affected.
39In 2007, Shanghai has the highest intermediate input-output ratio.
40Chow (1993) and Cao and Birchenall (2013) use valued added data to estimate the agricultural production function with only factor inputs of land, capital and labor, hence the land input share in their papers is consistent with $1 - \sigma$ in this paper.
41In Appendix B we test the sensitivity of $\alpha$ and $\sigma$ and in an acceptable interval, the results are quite similar.
we assume in the long run $L_a/N$ will approach 0.005 (the target) to back out $a$ based on equation (15). The implied value of $a$ is 0.002. Then we calibrate $\bar{a}$ to match the agricultural employment share and agricultural labor productivity, which leads to $\bar{a} = 1801.5$. The long-run agricultural employment share target is a conservative choice. Lower target will result in lower relative importance of agricultural goods (lower $a$) and higher subsistence level (higher $\bar{a}$). Due to the increase of subsistence requirement, the model can explain more of the agricultural employment share and aggregate labor productivity differences.

### 4.1.2 Variable Parameters

In this section, we calibrate the 5 variable parameters in Beijing to match the data of Beijing. Calibration of these 5 parameters of other provinces follows the same strategy but uses the data of the corresponding provinces. The economy-wide TFP parameter $A$ is calibrated to match non-agricultural GDP per worker of each province in 2007. $A = 108158$ in our benchmark economy. We choose $\kappa = 540.2$ to match the agricultural output per worker. $Z/N$, the land-to-employment ratio, is calculated from quality adjusted land data and total employment as in Appendix A.1. In 2007, the arable land per worker (in hectares) in our benchmark economy is 0.0687. We calibrate the wage gap $\theta$ indirectly by using average labor productivity ratio between agriculture and non-agriculture. $\theta = 0.8798$ in our benchmark economy. Since there is no direct empirical counterpart for the financial friction parameter $\lambda$, we use first-order condition in the model to back it out. In the model we assume that the only way to finance the purchase of intermediate non-agricultural input is borrowing. Hence we use intermediate non-agricultural input data in 2007 as our direct target to calibrate financial friction parameter $\lambda$. It gives $\lambda = 0.6352$. See Table 4 for all the calibrated parameters in 2007 (only variable parameter values in Beijing are displayed).

### 4.2 Quantitative Results

In this section, we conduct several quantitative analyses to show the explanatory power of our model. We mainly examine two measures: elasticity among variables that we are interested in, and factor differences of these variables. Through these measures we can quantify

---

42Remember that $N$ is total employment and $Z/N$ means how large arable land we have to feed one labor.

43Wage data are not available. See Restuccia, Yang and Zhu (2008) for the same method.

44We also use agricultural debt as our indirect target to calibrate $\lambda$. Results of quantitative exercises based on this calibration are similar and they are reported in Appendix B.3.

45Notice that the cutoff value of financial frictions is $\frac{\delta}{1-e(1-\alpha)} = 0.7544$ (see Proposition 1), which means borrowing constraint of Beijing is binding (and so are other provinces).
Table 4: Parameter Values

<table>
<thead>
<tr>
<th>Fixed</th>
<th>Value</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^*$</td>
<td>0.0156</td>
<td>Annual real interest rate</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.64</td>
<td>Chow (1993); Cao and Birchenall (2013)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.5252</td>
<td>Highest intermediate input share</td>
</tr>
<tr>
<td>$a$</td>
<td>0.0020</td>
<td>Long-run share of agricultural employment</td>
</tr>
<tr>
<td>$\bar{a}$</td>
<td>1801.5</td>
<td>Agricultural employment share in Beijing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Target (in Beijing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>108158</td>
<td>Non-agricultural labor productivity</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>540.2</td>
<td>Agricultural labor productivity</td>
</tr>
<tr>
<td>$Z/N$</td>
<td>0.0687</td>
<td>Land-to-employment ratio</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.8798</td>
<td>Wage gap</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.6352</td>
<td>Intermediate input</td>
</tr>
</tbody>
</table>

the role of financial frictions in explaining productivity differences. We take Beijing to be our benchmark economy and if it is not mentioned, parameters are set to be identical to the benchmark economy.\(^{46}\)

4.2.1 Aggregate Impact of Financial Frictions

Elasticity We compare model-generating elasticities of variables that we are interested in with respect to financial frictions with those generated from the data to show the explanatory power of financial frictions alone. Financial frictions are province-specific and other parameters are kept unchanged. In particular, we vary $\lambda$ to span a range of intermediate input ratios $X/Y_a$ observed in the data. Values of $\lambda$ range from 0.29 to 0.75, an interval below the cutoff value of $\lambda$. This results in the value range of $X/Y_a$ is from 0.20 to 0.52. When we shutdown variations of productivity parameters, land-to-employment ratio, and labor market friction, across provinces other than the financial friction parameter $\lambda$, we are able to assess the capacity of our model to explain productivity differences through the financial friction channel. The calibrated values of parameters $A$, $\kappa$, $Z/N$, and $\theta$ for each province are the same as in the benchmark economy. Our theory predicts that higher $\lambda$ leads to lower $L_a/N$, higher $X/Y_a$, higher $Y_a/L_a$, and higher $Y/N$.

The results are drawn in Figure 6. Solid dots (red) are simulation results generated by our model and hollow dots (blue) are data observations. Dash line (red) and solid line (blue) are

\(^{46}\)To compare aggregate output per worker across provinces, the relative price of agricultural good is set to be the same as in our benchmark economy, which is unity. This setup is also consistent with the data we constructed. The model predicts higher relative price of agricultural goods in poorer areas. By setting the same relative price across provinces, we can eliminate the impact of regional relative price differences and compare provincial aggregate output in real term. Also to be consistent with the data, the intermediate input ratio $(X/Y_a)$ we use in the following analysis is the ratio of value.
fitted lines for the model and data respectively. They show the correlations among the key variables implied by $\lambda$ in log scale. Therefore the three panels in the figure show the elasticity of $L_a/N$, $Y_a/L_a$ and $Y/N$ (from the left to the right) with respect to the intermediate input share $X/Y_a$: the slope means the percentage change of $L_a/N$, $Y_a/L_a$ and $Y/N$ when there is 1 percent change of $X/Y_a$ implied by $\lambda$.

![Graph](image)

**Figure 6: Model vs Data: Different $\lambda$ Only (log-scale)**

From Figure 6 we can see the model can generate an elasticity of 0.62 between $X/Y_a$ and $Y_a/L_a$ (the middle panel), which means 1 percent decrease of $X/Y_a$ decreased by financial frictions could result in 0.62 percent decrease of $Y_a/L_a$, and it is 32 percent of the elasticity observed in the data, which is 1.96. In the left panel, the slope of the fitted line based on the model is $-0.56$ while that from the data is $-2.08$. The agricultural employment share increases from 4 percent to 7 percent. Hence, one may conclude that differences in $\lambda$ can predict 27 percent of the elasticity between $X/Y_a$ and $L_a/N$ observed from the data. In the right panel one can see that the elasticity of $Y/N$ with respect to $X/Y_a$ predicted by the model is only 0.03 and hence our model predicts merely 1.5 percent of what is observed in the data.

Land is another important factor in agriculture production. If poor provinces have less land, then poor provinces have more incentive to use labor and hence labor productivity in poor provinces will be depressed. However, we document from the data that land-to-employment ratio is actually higher in poor provinces. Hence, poor provinces have less incentive to use
labor and this will dampen the effect of financial frictions. To see this quantitatively, we also vary land-to-employment ratio $Z/N$ in addition to financial frictions to match the data in each province. As is clear from Figure 7, our model could explain about 20 percent of the elasticities of agricultural employment share and productivity observed in the data.

The above experiments can not generate aggregate productivity difference that is comparable to the explanatory power of our model in accounting for agricultural employment share and agricultural labor productivity. The main reason is as follows. Aggregate productivity could be decomposed into weighted average of agricultural and non-agricultural productivity:

$$
\frac{Y}{N} = \frac{Y_a}{L_a} \cdot \frac{L_a}{N} + A \left(1 - \frac{L_a}{N}\right). 
$$

First of all, in this simulation exercise, non-agricultural productivity parameters are the same across provinces. Second, agricultural employment share is 29 percent of what is observed in the data and hence in our simulation poor provinces have much smaller agricultural employment share. This means non-agricultural productivity is the dominant factor determining aggregate productivity. Non-agricultural productivity in each province is the same also implies aggregate labor productivity difference is quite small. Figure 8 confirms our intuition. In Figure 8, we conduct another simulation in which we only vary non-agricultural productivity $A$ and financial friction $\lambda$. It is clear from Figure 8 that aggregate productivity differences in the data are explained quite well.

As before, our model generate less differences if we also vary land-to-employment ratio to match the data. Quantitatively, this experiment explains about 10 percent less of the elasticities of agricultural employment share and productivity observed in the data than the experiment that varies TFP and financial frictions only.
Since $A$ in the benchmark economy is very high, the income effect is too small to generate enough agricultural employment share, limiting the power of $\lambda$ to explain $Y/N$. When $A$ decreases, the income effect increases, and financial frictions would push more labor out of the agricultural sector. This intuition is confirmed by our simulation exercise starting from the poorest economy. In this exercise, we change the benchmark economy to the province with the lowest $Y/N$ and simulate $\lambda$ again. Figure 9 shows the results. We can see that comparing with Figure 6 (the richest province as benchmark economy), the elasticity of $Y_a/L_a$ with respect to $X/Y_a$ is similar (0.62 versus 0.63). The elasticity of $L_a/N$ with respect to $X/Y_a$ is also comparable as before (0.56 versus 0.63), but the income effect is much larger (agricultural employment share decreases from 62 percent to 34 percent). As a result, this simulation can generate larger differences of $Y/N$, and can explain 25 percent of aggregate productivity difference in terms of elasticity.\footnote{In 2007, Guizhou has the lowest GDP per worker.\footnote{We also conduct the same experiment except that the benchmark economy is chosen so that the TFP is the median among all provinces. Comparing with the experiment where the richest province is chose as the benchmark economy, elasticities of $Y_a/L_a$ and $L_a/N$ with respect to $X/Y_a$ are almost the same. The income effect is also large (agricultural employment share decreases from 35 percent to 15 percent) but smaller than that in the experiment which takes the poorest province as the benchmark economy. Hence, this experiment generates only 8 percent of aggregate productivity difference in terms of elasticity. However, this number is much better than 1.5 percent.}}

Figure 8: Model vs Data: Different $\lambda$ and $A$ (log-scale)
Factor differences  Above experiments only allow us to compare two trends (regression slopes). Here, we directly compare the factor differences of important endogenous variables generate by the model with the data. We allow provinces to have their specific economy-wide productivity $A$ and check how the calibrated model can explain the factor differences in $L_a/N$, $Y_a/L_a$, and $Y/N$ across provinces. In order to isolate the contribution of financial frictions, we consider 3 different versions of the model:

1. All provinces share the same parameter except $\lambda$.
2. All provinces share the same parameter except $A$.
3. All provinces share the same parameter except $\lambda$ and $A$.

The quantitative results are presented in Table 5. We report the ratio of the equilibrium outcomes between the average value of richest decile provinces and that of the poorest decile provinces. Since we choose $\lambda$ to match $X/Y_a$ in the data and they match perfectly, so there is no need to report the factor differences in $X/Y_a$. The model with no frictions can only explain limited factor differences of $Y_a/L_a$ (53 percent), and $Y/N$ (76 percent). These differences only come from the variation of $A$. After we introduce the differences in borrowing constraint parameter (version 2), the model can explain additional 29 percent and 7 percent factor differences in $Y_a/L_a$ (82 percent), and $Y/N$ (83 percent), respectively. In other words,
the factor difference in agricultural labor productivity of the model with financial frictions is approximate 1.55 times of that difference in the model with no frictions. The factor difference of $Y/N$ is amplified by financial friction for only 8 percent, confirming the importance of TFP in accounting for aggregate labor productivity through the income effect.

The factor differences generated by the model with variation of both $A$ and $\lambda$ (version 2) are comparable to the data and hence our calibrated model has a strong explanatory power in agricultural input and output differences. To see this more clearly, we also present the average values by quartile based on observed aggregate GDP per worker (Table 6). The model is able to generate quantitative results that are quite close to the data except for the agricultural employment share. $X/Y_a$ matches perfectly in each quartile. Comparing with the data, as province becomes poorer, productivity gap between model and data becomes larger because of the small income effect. Hence the model is not good at generating large agricultural employment share in poor provinces.

### 4.3 Assessment of Model’s Implications

In our calibration, we do not directly target relative price of agricultural goods and agricultural employment. Hence, comparing the values of these two variables generated from our
model with the data is a good way to evaluate our model. We choose to compare the results generated by the model with variations only in $\lambda$ and $A$ with the data.\textsuperscript{50} It turns out our model fits the data quite well.

**Agricultural employment share** Due to subsistence requirement of the agricultural good, the model predicts that agricultural employment share should be higher in poorer area. In the left panel of Figure 10 we compare the results of model and data. The quantitative results of the model show a strong trend that agricultural employment share declines in aggregate GDP per worker, while the data shows an even stronger pattern.

**Agricultural relative price** The model predicts that agricultural relative price should be higher in poor provinces than rich provinces, which is confirmed by the quantitative results in the right panel of Figure 10. In this figure, we compare the relative prices $p_a$ generated by the model, with the PPP price ratio $PPP_a/PPP_n$. We can see that the model result $p_a$ (red solid dots) declines in aggregate GDP per worker and show a similar slope as what we observed in the data (blue hollow circles).

![Figure 10: Agricultural Employment Share](image)

*Note:* PPP GDP per worker and relative price $p_a$ are in log scale.

\textsuperscript{50}The results generated by the model with variations in $\lambda$, $\theta$, $\kappa$, $Z/N$, and $A$ are quite similar.
5 Conclusion

This paper explores the role of financial frictions in accounting for labor productivity differences across provinces in China. A novel dataset featuring PPP-adjusted sector-level GDPs in each province is constructed to compare labor productivity across provinces. In particular, we find that cross-province agricultural and aggregate output per worker differences are 5.3-fold and 5.9-fold respectively. Moreover, employment share of agriculture is 6 percent on average in the richest 10 percent of all provinces while the poorest 10 percent of provinces have almost half of their employment working in agriculture. We then explore household-level data from various surveys to find financial frictions exist in rural China with the feature that financial frictions in poorer regions are more severe. A two-sector general equilibrium model with a subsistence consumption requirement, modern intermediate input and financial frictions is constructed. Limited credit depresses the use of intermediate inputs while it encourages the use of labor inputs. Consequently, workers are trapped in the agricultural sector and agricultural labor productivity is low. Because of a large weight of agricultural employment, aggregate labor productivity is also low. Our theoretical predictions are broadly consistent with empirical evidence in the literature. Our quantitative exercises show that a substantial part of observed agricultural employment share and labor productivity differences can be accounted for by financial frictions.

References


A Data Description

A.1 Macro Data

The aggregate and disaggregate data, including GDP and total output mainly come from various series of *China Statistical Yearbook* which are published by National Bureau of Statistics of China (NBS). All goods and services are measured in nominal value.

The sample consists of 31 provinces in China. They are Beijing, Tianjin, Hebei, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Guangxi, Hainan, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang.

The agricultural sector consists of farming, forestry, animal husbandry, fishery and agricultural services.\(^{51}\) Detailed agriculture sector data come from *China Agricultural Statistical Yearbook*.

A.1.1 Agricultural Employment

The official employment data is criticized by the literature for several inconsistencies.\(^{52}\) We follow the literature to construct a set of alternative data which can fix the problems in the official data.

We calculate provincial employment shares from the *National Population Census* in the year 2000, 2005, and 2010. For the shares between these census years, we use cubic interpolation. Then we apply the results to the national employment data from *China Statistical Yearbook* in various years to compute provincial total employment in each year.

We follow Brandt and Zhu (2010) to use a national wide adjustment factor to correct provincial agricultural employment data in each year. We divide the national agricultural employment share from the *Rural Household Survey* by national agricultural employment share from the *China Statistical Yearbook* to calculate the adjustment factor. Finally, the provincial agricultural employment data from the *China Statistical Yearbook* is multiplied by the adjustment factor to generate the adjusted provincial agricultural employment.

\(^{51}\)There are no significant differences if we only focused on farming and animal husbandry. According to 2005 *National Agricultural Census*, at least 97.7 percent of agricultural employment mainly engaged in farming and animal husbandry.

\(^{52}\)See Brandt and Zhu (2010); Brandt, Tombe and Zhu (2013).
A.1.2 Purchasing Power Parity

Although this paper focuses on productivity differences within a country, and all values are measured using the same currency, regional price differences still exist. To eliminate the impact of price differences, we use Geary-Khamis method to calculate purchasing power parity (PPP) for both agricultural output and non-agricultural output. PPP of Beijing is normalized to be unity.

**Agricultural PPP** Provincial agricultural output quantities are collected from *China Agricultural Statistical Yearbook* and *China Statistical Yearbook* in 2007. We do not have the producer price data of agricultural goods in 2007 so we collect the producer price data in 1990 from *China Price Statistical Yearbook* in 1991 and use detailed producer price index to derive the data we need for the year after 1990. After 1994, only price index data can be found in the statistical yearbook. The detailed producer price index we use are from *China Market Statistical Yearbook*, *China Price and Urban Household Income and Expenditure Statistical Yearbook*, *China Rural Statistical Yearbook*, and *China Agricultural Commodity Price Statistical Yearbook*.

We only include farming and husbandry in the calculation because we do not have accurate price data for forestry or fishery. Hainan and Tibet are excluded because of the lack of price data. From Figure 11 we can see agricultural PPP is negatively correlated with agricultural productivity per worker, which is consistent with the literature. This means the price of agricultural goods is lower in richer areas.

**Non-agricultural PPP** We calculate non-agricultural PPP based on urban household expenditure for the year 2010 since we only have urban consumer price data in 2010. We collect price and expenditure data for food, tobacco, beverage, clothing, housing, utility, household equipment and services, health, transportation, communication, and education for 31 provinces. The expenditure data come from the *China Statistical Yearbook* in 2011. We calculate provincial composite consumption price for each category based on the price data from the price monitoring center of National Development and Reform Commission and the price monitoring center of each province. Hainan and Tibet are excluded to be consistent with agricultural PPP. From Figure 12 we can see non-agricultural PPP is positively correlated with non-agricultural productivity per worker. This means the price of non-agricultural goods in richer areas is higher. PPP-adjusted aggregate value added is obtained by adding PPP-adjusted agricultural value added and PPP-adjusted non-agricultural value added.

A.1.3 Intermediate Non-agricultural Inputs

We use similar definitions as in Restuccia, Yang and Zhu (2008). Total agricultural output is the sum of intermediate agricultural input, intermediate non-agricultural input ($X$), and
agricultural value added. Final agricultural output \((Y_a)\) is equal to total agricultural output minus intermediate agricultural input. Intermediate non-agricultural input is obtained by subtracting PPP-adjusted agricultural value added from PPP-adjusted final agricultural output. We collected the data from 2007 China Regional Input-Output Tables published by NBS. Tibet is excluded due to the lack of data.

Since we do not have price data on agricultural intermediate input, the agricultural intermediate input-output ratio \((X/Y_a)\) actually contains the relative price information of intermediate input and agricultural output. It is also necessary to keep this price information to investigate the role of financial friction because farmers have to consider the price of intermediate input when making production decision. The descriptive statistics of \(X/Y_a\) for the year 2007 is reported in Table 7.

### A.1.4 Land Quality

Department of Land and Natural Resources of China published Investigation and Evaluation of Cultivated Land Quality in 2009. It uses standard food production per unit land to measure land quality based on several characteristics of land (such as inherent soil, climate, type of...
Figure 12: Non-agricultural PPP

Source: Authors’ own calculation based on 2011 China Statistical Yearbook and price data from price monitoring center of National Development and Reform Commission and the price monitoring center of each province.

Table 7: Descriptive Statistics of Agricultural Intermediate Input Output Ratio

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Coefficient of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.20</td>
<td>0.52</td>
<td>0.32</td>
<td>0.30</td>
<td>0.07</td>
<td>0.225</td>
</tr>
</tbody>
</table>

Source: Authors’ own calculation based on 2007 China Regional Input-Output Tables.

crops and other characteristics). Based on this data, we linearly divided land quality into 4 categories: excellent (16,500 ~ 22,500 kg/hm²), high (10,500 ~ 16,500 kg/hm²), medium (4,500 ~ 10,500 kg/hm²) and low (0 ~ 4500 kg/hm²), with weights of 4, 3, 2 and 1 respectively (the standard food production per unit land in the excellent group is 4 times as high as that in the low group). Then we adjusted the land size based on the land quality and weights.

A.1.5 Physical Capital

The physical capital data we use is from Wu (2009). It provides estimated physical capital stock data of China by sector and province from 1978 to 2006. The data is in constant price
in 1978. First we use GDP deflators to adjust the price of capital to 2006. Then we use agricultural PPP to adjust agricultural capital and use non-agricultural PPP to adjust non-agricultural capital. The aggregate capital is the sum of the PPP-adjusted agricultural and non-agricultural capital. We use provincial capital data in 2006 in our analysis.

A.1.6 Year of Schooling

Our year of schooling data is from 2005 *The Second National Agricultural Census*. Average year of schooling in richer areas is higher than that in poorer areas. See Table 8 for a summary by region.

Table 8: Average Year of Schooling

<table>
<thead>
<tr>
<th>Region</th>
<th>0</th>
<th>6</th>
<th>9</th>
<th>12 and more</th>
</tr>
</thead>
<tbody>
<tr>
<td>East</td>
<td>8%</td>
<td>39%</td>
<td>49%</td>
<td>5%</td>
</tr>
<tr>
<td>Central</td>
<td>9%</td>
<td>37%</td>
<td>49%</td>
<td>5%</td>
</tr>
<tr>
<td>West</td>
<td>13%</td>
<td>47%</td>
<td>37%</td>
<td>4%</td>
</tr>
<tr>
<td>Northeast</td>
<td>3%</td>
<td>39%</td>
<td>55%</td>
<td>3%</td>
</tr>
</tbody>
</table>

*Source: Authors’ own calculation based on the year of schooling data in 2005 *The Second National Agricultural Census*.*

A.1.7 Agricultural Zone

We divided provinces into several agricultural zones according to *Comprehensive Agricultural Zoning of China*. There are nine areas: Dong Bei, Gan Xin, Huang Huai Hai, Huang Tu Gao Yuan, Qing Zang, Xi Nan, Chang Jiang Zhong Xia You, Hua Nan, and Inner Mongolia and the Great Wall. Some provinces belong to multiple areas.

A.1.8 Agricultural Loan

We use the provincial agricultural loan data from *China Compendium of Statistics 1949-2008*. This agricultural loan refers to the short-term (less than or equal to one year) loan issued by all financial intermediaries only for the production of agriculture, forestry, animal husbandry and fishery. The loan might be used to purchase agricultural intermediate inputs, non-agricultural intermediate inputs, agricultural capital, or employment expenditure (we cannot distinguish the purpose).

Since the purchase of intermediate inputs happened at the beginning of each production cycle, we use the average nominal loan rate (base rate set by the central bank) from Aug 2006
to May 2007, which is 6.36 percent. Chinese government does not report official inflation rate, so we use official CPI (Consumer Price Index) in 2007 instead, which is 4.8 percent. Then the annual real loan rate in 2007 we use is $r^* = 6.36\% - 4.8\% = 1.56\%$. Our model results are not sensitive at all to the value of $r^*$. Because given our calibration strategy, $\lambda$ is proportional to $1 + r^*$ and the $X/Y_a$ generated by the model will not be affected.

A.2 Micro Data

2006 Farmer's Credit Situation Survey (FCSS) Farmer's Credit Situation Survey was conducted by the People's Bank of China (the central bank of China) in year 2006. It covers 10 provinces, 263 counties, 20040 households. The 10 provinces are Anhui, Fujian, Guizhou, Henan, Hubei, Inner Mongolia, Jiangsu, Jilin, Ningxia, Sichuan, which are nationally representative. First, there are 3-4 provinces in each income group. Fujian, Inner Mongolia and Jiangsu are among top 10 provinces in terms of per capita GDP. Henan, Hubei, and Jilin are among medium income provinces. Anhui, Guizhou, Ningxia and Sichuan are in the group of low income provinces. Second, geographically, Fujian, Jiangsu and Jilin are in the east; Anhui, Henan, Hubei and Inner Mongolia are in the central area; Guizhou, Ningxia and Sichuan are in the west. The richest province in the sample is Jiangsu which is the fifth richest province in China in 2006. The poorest province in the sample is Guizhou which is the poorest province in China in 2006.

We focus on the households who mainly engaged in agricultural production in year 2006 and eliminate those who do not have the habit of borrowing money (then we have 12260 cases). Among these households, 6934 cases (56.57 percent) answered that they needed to borrow from the bank. Borrowing households were asked whether they can get credits from the bank and whether they can get enough credits. The household who is credit rationed is defined as who can not get the loan from the bank or the loan is less than demand. The degree of financial frictions is defined as the number of farmers who are credit rationed divided by the number of households who need to borrow. The number of workers of each household is defined as the number of labor over the age of 16 minus the number of migrant workers.

For the county level data, we exclude the counties which contain less than 20 households to improve the accuracy of binding ratio. From Figure 13 we can see that the negative correlation does increase as we raise the criteria level (below which we need to exclude) from 0 to 20. As we increase the cutoff sample size, the significance starts to drop because of the severe reduction of total sample size.

In any case, we stick to this survey if the data is available to obtain empirical findings. However, some important questions are not included in this survey. To this end, we use the
other surveys which have questions we are interested in. The other surveys are also used as robustness checks if necessary.

2005-2007 Rural Financial Development Survey (RFDS)  
Rural Financial Development Survey is a household-level survey and was conducted by Tsinghua University during 2006-2008. Although it has data in 3 years, it is not a panel data. Data in 2005 covers 4 provinces: Gansu, Hebei, Qinghai and Xinjiang; 29 counties and 1523 households. Data in 2006 covers 8 provinces: Anhui, Henan, Heilongjiang, Hubei, Hunan, Jiangxi, Jilin and Shanxi; 31 counties and 1990 households. Data in 2007 covers 4 provinces: Inner Mongolia, Liaoning, Shaanxi and Shandong; 11 counties and 1263 households.

Sampling strategy is not mentioned in the data. Although the survey in each year is slightly different from each other, all of them cover basic demographic information, education attainment, information about household expenditure, information on agricultural production and financial situation. There are more detailed questions on input costs, outputs and profits of agricultural production. Household income and expenditure information are divided into several categories. More importantly, there are very detailed questions on household borrowing including borrowing record in that year, the way of borrowing (from banks or family and friends or other sources), the reason for borrowing.
We use the dataset to obtain several useful descriptive statistics in the introduction section.

**2011 China Household Finance Survey (CHFS)**  *China Household Finance Survey* is the only nationally representative survey in China that has detailed information about household assets and liabilities. In addition, the survey also has information on income and expenditure. The survey was conducted by Southwestern University of Finance and Economics (SWUFE) in 2011. The sampling design for the China Household Finance Survey (CHFS) consists of two major components, an overall sampling scheme and an onsite sampling scheme based on mapping. The survey employs a stratified three-stage probability proportion to size (PPS) random sample design. The primary sampling units (PSU) include 2,585 counties from all provinces in China except Tibet, Xinjiang, Inner Mongolia, Hong Kong, Macau, and Taiwan. The second stage of sampling involves selecting residential committees/villages from the counties/cities selected at the earlier stage. The last stage is to select households from the residential committees/villages chosen at the previous stage (onsite sampling scheme based on mapping is adopted in this stage). Then at least the member who knows the financial situation best in selected households are interviewed. Every stage of sampling is carried with PPS method and weighted by its population size. The 2011 CHFS collects information from 8,438 households consisting of 29,463 individuals.

Households employed in agricultural production of any kind will be asked detailed questions on agricultural productions and corresponding financial situation. For agricultural productions, there are questions on revenue from agricultural production and inputs, such as fertilizer use or the cost of hired labor, in terms of money. For financial situation, there are questions on borrowings, from either banks or other channels (very specific), only for the purpose of agricultural productions and related factors.

For this survey, we focus on the households who only engaged in agricultural production. Beijing, Tianjin, Shanghai and Qinghai are excluded because of zero observations. Liaoning and Guizhou are excluded because they are two clear outliers after calculation. At last we end up with 2587 observations.

### A.3 Additional Evidence

#### A.3.1 More Evidence on Intermediate Inputs

Because of different climate and land quality, various regions may produce different agricultural product and have distinct intermediate input level. Physical and human capital may

---

54 Usually, more or all household members are interviewed.
55 There are no question on outputs such as the quantity or value of each crop growth, etc.
also affect the use of intermediate inputs. To test the correlation between agricultural output per worker and intermediate inputs, we control for land size (quality adjusted), human capital and physical capital in the agricultural sector:

\[
\ln \left( \frac{Y_a}{L_a} \right)_i = \alpha_0 + \alpha_Z \ln \left( \frac{Z}{L_a} \right)_i + \alpha_K \ln \left( \frac{K_a}{Y_a} \right)_i + \alpha_h \ln (h_i) + \alpha_X \ln \left( \frac{X}{Y_a} \right)_i + \varepsilon_i,
\]

(5)

where \(Y_a\) is agricultural final output, \(Z\) is quality adjusted land size (controlling for inherent soil, climate, type of crops and other characteristics, see Appendix A.1.4), \(K_a\) is agricultural physical capital, \(h\) is human capital, \(L_a\) is agricultural employment, \(X\) is intermediate input, and \(i\) is the index for provinces.\(^{56}\) Since we do not take into account provincial specific TFP level, to eliminate the possible endogeneity problem, we use \(X/Y_a\) and \(K_a/Y_a\) instead of \(X/L_a\) and \(K_a/L\) on the right hand side of (5).

If \(X/Y_a\) is highly correlated with other factor inputs such as quality adjusted land, physical or human capital, then when we control for these variables, they will affect the significance of \(X/Y_a\). The regression results for (5) are reported in Table 9. We can see that land quality, physical capital and human capital are not significant in determining output and the coefficient of \(X/Y_a\) is very significant.

Table 9: Regression on intermediate inputs

<table>
<thead>
<tr>
<th>(\ln Z/L_a)</th>
<th>(\ln K_a/Y_a)</th>
<th>(\ln (h))</th>
<th>(\ln X/Y_a)</th>
<th>Adjusted (R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2246</td>
<td>-0.2586</td>
<td>0.8582</td>
<td>1.7947***</td>
<td>0.5615</td>
</tr>
<tr>
<td>(0.2093)</td>
<td>(0.1652)</td>
<td>(1.0333)</td>
<td>(0.4166)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard errors are reported in parentheses. \^{****}: 0.1% significance.

Source: Authors’ own calculation. Agricultural intermediate input and output are based on 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010). Land size is calculated from 2009 Investigation and Evaluation of Cultivated Land Quality. The agricultural human capital data is calculated from the year of schooling data in 2005 The Second National Agricultural Census.

For a more specific study on the regional effect, we also divided provinces into several agricultural zones based on their latitude, altitude, type of production and climate.\(^{57}\) None of the regional dummies are significant and \(X/Y_a\) is still very significant.

Transportation of intermediate non-agricultural input could also affect agricultural production. We use total length of highway, length of highway per capita, or length of highway over surface area for each province as various proxies for transportation and add these proxies in (5). The results show that none of them are significant and \(X/Y_a\) is still very significant.

\(^{56}\)The measure of human capital is \(h_i = \exp(0.1 \times s_i)\). \(s_i\) is average year of schooling of province \(i\). We simply assume that each year of schooling increases income by around 10 percent. See Appendix A.1.6.

\(^{57}\)See Appendix A.1.7.
Therefore agricultural intermediate input is an important factor to determine agricultural output.

A.3.2 More Evidence on DFF and Intermediates

Borrowing constraint usually correlates with land size and human capital. To test the correlation between financial frictions and income per worker, we also control for land size and average year of schooling from the survey data using the following regression:

\[ y = \gamma_0 + \gamma_1 \times DFF + \gamma_Z Z + \gamma_S \exp(10\% \times S) + \varepsilon, \]  

(6)

where \( y \) is agricultural sales income per worker, \( Z \) is land size and \( S \) is year of schooling \((h = \exp(10\% \times S) \) is defined as human capital). The regression results for (6) are reported in Table 10. We can see that human capital is undoubtedly very important in explaining income per worker, however, the coefficient of DFF is still very significant.\(^{58}\) Hence, the evidence from the micro data is in accord with the evidence we find in the macro data in Table 9, if we take into account that the DFF is negatively correlated with the use of intermediate non-agricultural input \( X \).

<table>
<thead>
<tr>
<th></th>
<th>DFF</th>
<th>Z</th>
<th>h</th>
<th>Adj. ( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-4362.90****</td>
<td>15.89</td>
<td>4156.04****</td>
<td>0.1832</td>
</tr>
<tr>
<td></td>
<td>(1285.10)</td>
<td>(13.80)</td>
<td>(1152.44)</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: Standard errors are reported in parentheses. \(****\): 0.1% significance.  
Source: Authors’ own calculation based on 2006 Farmer’s Credit Situation Survey.

A.4 Development Accounting

The production function based on which we conduct our development accounting exercises is as follows:

\[ \frac{GDP_a}{L_a} = A_a \left( \frac{Z}{L_a} \right)^{1-\sigma-\gamma} \left( \frac{K_a}{L_a} \right)^{\gamma} h^\sigma, \]

where agricultural capital income share \( \gamma = 0.25 \) and agricultural labor income share \( \sigma = 0.39 \) are chosen from the average of the estimations by Chow (1993) and Cao and Birchenall (2013). The data of \( GDP_a, \) land \( Z, \) physical capital \( K_a, \) human capital \( h, \) and labor \( L_a \) are the same as in Section 2.1 and 2.2.

\(^{58}\) The human capital is significant because this is a regression in individual level.
How successfully can the factor-only model explain labor productivity differences? The first is the log variance ratio ($\text{success}_1$), which is defined as the ratio of log variance in output per worker generated by the factor-only model divided by the actual log variance. Log variance ratio is kind of sensitive to outliers, so we also check another less sensitive measure which is inter-percentile differential ratio ($\text{success}_2$). In particular, we compare the 90-10 ratio of output per worker generated by the factor-only model with the actual 90-10 ratio.\footnote{We use the values of average of each decile. The results calculated from single points are similar.}

B Additional Results

In this section we conduct several sensitivity analyses and alternative experiments to test the robustness of our quantitative results. First we explore the role of province-specific variables ($A$, $\kappa$, $Z/N$, and $\theta$) and conduct some counterfactual analyses. Then we conduct several sensitivity tests for important parameters ($\alpha$ and $\sigma$) to see whether they are crucial to the results. Finally we use another target to redo the calibration for financial friction parameter $\lambda$ and our model with this alternative calibration still produces similar quantitative results.

B.1 Variations in $\theta$, $\kappa$, $Z/N$

We do not allow each province to have its specific land-to-employment ratio $Z/N$, labor market friction $\theta$, and agricultural productivity parameter $\kappa$ in previous analysis. It turns out that the quantitative results replicate the data well, which confirms that $\lambda$ and $A$ are crucial for agricultural labor productivity differences. It is also of our interest that how other parameters such as $\theta$, $\kappa$ and $Z/N$ affect our quantitative results.

Here we allow all provinces to have their province-specific calibrated values of parameters $A$, $\kappa$, $Z/N$, $\theta$, and $\lambda$ and then simulate the model. We calibrate these parameters in the same way as in Section 4.1 except that province-specific parameters are calibrated by using data of the corresponding province. It turns out that our model fits the data quite well. Quantitative results in this section are basically quite similar to our previous results. Figure 14 shows the results when $\theta$, $\kappa$, $Z/N$ are calibrated to match the data in corresponding provinces.

The qualitative results do not change at all while the quantitative results match the data better. For example, the elasticity (the slope) between $L_a/N$ and $X/Y_a$ is $-2.08$ while our model predicts a slope of about $-1.83$, which is 88 percent of the relationship between agricultural labor productivity and intermediate input-output ratio from the data. The middle panel shows a positive correlation between $X/Y_a$ and $Y_a/L_a$, and the slope of the fitted line generated from the simulated results of the model are almost identical to that of the data.
Figure 14: Model vs Data: Provincial Specific $A$, $\kappa$, $Z/N$, $\theta$, and $\lambda$

*Note:* Both axes are in log scale.

*Source:* Agricultural intermediate input and output are from on 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of *China Statistical Yearbook* using the method from Brandt and Zhu (2010).

The elasticity of aggregate output per worker $Y/N$ and intermediate input share $X/Y_a$ is 92 percent of that shown in the data.

**B.1.1 The role of $\theta$**

Figure 15 shows the correlation between the ratio of agriculture to non-agriculture wage $1 - \theta$ and two key macro variables: agricultural intermediate input share $X/Y_a$ and agricultural productivity $Y_a/L_a$. The positive correlation between the agricultural intermediate input share and the wage gap between the two sectors $1 - \theta$ indicates the substitution between labor and intermediate inputs. Lower wages provide strong incentive to use more labor and less intermediate inputs for agriculture production, and vice versa.

We also compare the factor differences of important endogenous variables ($L_a/N$, $Y_a/L_a$, $Y/N$) generated by the two frictions $\lambda$ and $\theta$, since both of them could affect agricultural output. First we allow provinces to have their specific economy-wide productivity $A$, agricultural productivity parameter $\kappa$, and land-to-employment ratio $Z/N$. Then we see how the calibrated model can explain the differences in $L_a/N$, $Y_a/L_a$, and $Y/N$ across provinces.
In order to isolate the contribution of financial frictions, labor market frictions and both, we consider 4 different versions of the model:

1. Model with no frictions: all provinces share the same parameter value of $\theta$ and $\lambda$ ($\theta_j = \theta_{Beijing}$ and $\lambda_j = \lambda_{Beijing}$);

2. Model (1) with labor market friction $\theta$ ($\lambda_j = \lambda_{Beijing}$) only;

3. Model (1) with financial friction $\lambda$ ($\theta_j = \theta_{Beijing}$) only;

4. Model (1) with both frictions.

The quantitative results are presented in Table 11. We report the ratio of the equilibrium outcomes between the average value of richest decile provinces and that of the poorest decile provinces. Since we choose $\lambda$ to match $X/Y_a$ in the data and they match perfectly, so there is no need to report the factor differences in $X/Y_a$. The model with no frictions can only explain limited factor differences of $Y_a/L_a$ (58 percent), and $Y/N$ (77 percent). These differences mainly come from variations of $A$, $\kappa$, and $Z/N$. After we introduce the differences in borrowing constraint parameter (version 3), the model can explain additional 26 percent
Table 11: Effects of Borrowing Constraint

<table>
<thead>
<tr>
<th></th>
<th>( L_a / N )</th>
<th>( Y_a / L_a )</th>
<th>( Y / N )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rich/poor</td>
<td>rich/poor</td>
<td>rich/poor</td>
</tr>
<tr>
<td>Data</td>
<td>0.06</td>
<td>0.45</td>
<td>7.49</td>
</tr>
<tr>
<td>(4) Varying in ( \lambda, \theta, \kappa, \frac{Z}{N}, A )</td>
<td>0.04</td>
<td>0.29</td>
<td>7.38</td>
</tr>
<tr>
<td>(3) Varying in ( \lambda, \kappa, \frac{Z}{N}, A )</td>
<td>0.05</td>
<td>0.28</td>
<td>6.26</td>
</tr>
<tr>
<td>(2) Varying in ( \theta, \kappa, \frac{Z}{N}, A )</td>
<td>0.04</td>
<td>0.21</td>
<td>5.37</td>
</tr>
<tr>
<td>(1) Varying in ( \kappa, \frac{Z}{N}, A )</td>
<td>0.05</td>
<td>0.20</td>
<td>4.36</td>
</tr>
</tbody>
</table>

Note: Each column is ranked separately.

Source: Agricultural intermediate input and output are from the 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010). Land size is calculated from 2009 Investigation and Evaluation of Cultivated Land Quality.

and 6 percent factor differences in \( Y_a / L_a \) (84 percent), and \( Y / N \) (83 percent), respectively. In other words, the factor difference in agricultural labor productivity of the model with financial frictions is approximate 1.44 times of that difference in the model with no frictions. The amplification effect of financial friction is only 7 percent, again confirming the importance of province-wide productivity parameter in accounting for aggregate labor productivity.

The model which includes both financial friction and labor market distortion (version 4) is able to account for 84 percent of the factor difference in aggregate productivity across provinces in China, which is 7 percent better comparing to the model with no frictions (version 1). With both frictions, the model now can explain 99 percent of the factor differences in \( Y_a / L_a \), which matches the observed differences much better and leaves the unexplained factor to only 1.01 and 1.19, respectively. We should also notice that when we put both frictions together, the model can explain more than the sum of the factor differences generated by each friction alone in each column.

The model with both frictions accounts for 64 percent of the cross-province agricultural employment share difference, while the model without frictions can only account for 38 percent of that difference. The economy with only one friction (version 2 or 3) can account for 44 percent or 59 percent of that difference, respectively. Although labor market friction has a direct effect on labor allocation and financial friction has an indirect effect, the model with financial friction only is able to account for a larger fraction of the differences than that generated by the model with labor market friction only. Models with financial frictions only outperforms models with labor market friction only in terms of explaining cross-province differences of key variables. We conclude that credit market imperfection is more important in accounting for productivity differences than labor market imperfection.
B.1.2 The role of $\kappa$ and $Z/N$

Figure 16 shows the correlation between $Z/N$ and aggregate GDP per worker.\(^{60}\) We can see that $Z/N$ is relatively larger in poor provinces than in rich poor provinces.

\[ \text{Figure 16: Land to Employment Ratio} \]

*Note:* Horizontal axis is in log scale.

*Source:* GDP data is adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of *China Statistical Yearbook* using the method from *Brandt and Zhu (2010)*. Land size is calculated from 2009 *Investigation and Evaluation of Cultivated Land Quality*.

Figure 17 shows the variation of $\kappa$ across provinces. We can see that $\kappa$ does not vary systematically from poor provinces to rich provinces.

To determine the role of $\kappa$ and $Z/N$, we conduct the following 4 experiments:

1. Values of province-wide productivity $A$, agriculture-specific productivity parameter $\kappa$, and land-to-employment ratio $Z/N$ are set to be the same across provinces and equal to those of our benchmark economy: $A = A_{Beijing}$, $\kappa = \kappa_{Beijing}$, and $\frac{Z}{N} = \frac{Z}{N_{Beijing}}$.

2. Land-to-employment ratio $Z/N$ differs only: $A = A_{Beijing}$, $\kappa = \kappa_{Beijing}$, and $\frac{Z}{N}$ is calibrated to match the data of the corresponding province.

3. Agriculture-specific productivity $\kappa$ differs only: $A = A_{Beijing}$, $\frac{Z}{N} = \frac{Z}{N_{Beijing}}$, and $\kappa$ is calibrated to match the data of the corresponding province.

\(^{60}\)Aggregate GDP per worker is in log scale.
4. Economy-wide productivity $A$ differs only: $\kappa = \kappa_{Beijing}, \frac{Z}{N} = \frac{Z}{N}_{Beijing}$, and $A$ is calibrated to match the data of the corresponding province.

We also compare these results with the data and the model with variations in $\lambda, \theta, \kappa, Z/N,$ and $A$). Table 12 reports the results.

First by setting the same $A$, $\kappa$, and $Z/N$ for all provinces (experiment 1), we can have a better view of the frictions in our model. This experiment can generate 34 percent of the factor differences in $Y_a/L_a$ only by the differences in $\lambda$ and $\theta$ across provinces. Agricultural employment share is not affected too much so that the result can not give a good explanation in the factor differences in $Y/N$.

In experiment 2 we find that all the results do not change too much comparing to experiment 1, but overall they are similar. This confirms the empirical results in Section 2.2 that land is not important in explaining agricultural output. When we change $\kappa$ in experiment 3 we find that all the results increase a very small amount. Then we can conclude that $\kappa$ and $Z/N$ are not crucial in explaining productivity differences across provinces. This is not surprise because $\kappa$ does not vary systematically from poor provinces to rich provinces. And remember that land is never significant in determining the output in our previous empirical analysis.
Table 12: Experiments with Different $A$, $\kappa$, and $Z/N$

<table>
<thead>
<tr>
<th></th>
<th>$L_a/N$</th>
<th>$Y_a/L_a$</th>
<th>$Y/N$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rich/poor</td>
<td>rich/poor</td>
<td>rich/poor</td>
</tr>
<tr>
<td>Data</td>
<td>0.06</td>
<td>0.45</td>
<td>7.49</td>
</tr>
<tr>
<td>Baseline Model</td>
<td>0.04</td>
<td>0.29</td>
<td>7.38</td>
</tr>
<tr>
<td>$A = A_{Beijing}$, $\kappa = \kappa_{Beijing}$, $Z/N = Z/N_{Beijing}$</td>
<td>0.03</td>
<td>0.09</td>
<td>2.57</td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A = A_{Beijing}$, $\kappa = \kappa_{Beijing}$</td>
<td>0.02</td>
<td>0.06</td>
<td>2.82</td>
</tr>
<tr>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A = A_{Beijing}$, $Z/N = Z/N_{Beijing}$</td>
<td>0.03</td>
<td>0.13</td>
<td>3.98</td>
</tr>
<tr>
<td>(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa = \kappa_{Beijing}$, $Z/N = Z/N_{Beijing}$</td>
<td>0.04</td>
<td>0.28</td>
<td>7.34</td>
</tr>
<tr>
<td>(4)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Each column is ranked separately.

Source: Agricultural intermediate input and output are from on 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010). Land size is calculated from 2009 Investigation and Evaluation of Cultivated Land Quality.

Experiment 4 confirms our prediction that the economy-wide TFP parameter $A$ is important to explain productivity differences besides the frictions. In this experiment with only different $A$, the model can generate 62 percent, 98 percent and 84 percent differences in $L_a/N$ and $Y/N$ of the observations from the data.

B.1.3 Changes in $A$ and $N$

Economy-wide TFP and total employment are two important parameters. Here we conduct two experiments: increase $A$ and $N$ by the same percentage in each province. All parameters are not recalibrated after the adjustment.\footnote{Since most the parameters will not be affected, the results are quite similar even if we recalibrate all the parameters.} Table 13 reports the results.

First we increase $A$ by 10 percent, 20 percent, and 50 percent for all provinces. Due to the decline of income effect, we can see the factor differences decrease when $A$ increases, but not too much. Then we increase $N$ by 2 percent, 5 percent, and 10 percent (employment cannot grow too fast). Due to the diminishing role of $Z/N$, the role of $A$ and $\lambda$ increase. Given the limited change in $N$, the overall effect is also limited.

Table 13 reports the results.
Table 13: Effects of Borrowing Constraint: Increase in $A$ and $N$

<table>
<thead>
<tr>
<th></th>
<th>$L_a/N$</th>
<th>$Y_a/L_a$</th>
<th>$Y/N$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rich</td>
<td>poor</td>
<td>rich/poor</td>
</tr>
<tr>
<td>10% ↑ in $A$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varying $A, \lambda$</td>
<td>0.04</td>
<td>0.23</td>
<td>6.16</td>
</tr>
<tr>
<td>Varying $A$ only</td>
<td>0.04</td>
<td>0.16</td>
<td>3.97</td>
</tr>
<tr>
<td>Varying $\lambda$ only</td>
<td>0.04</td>
<td>0.06</td>
<td>1.60</td>
</tr>
<tr>
<td>20% ↑ in $A$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varying $A, \lambda$</td>
<td>0.04</td>
<td>0.21</td>
<td>6.15</td>
</tr>
<tr>
<td>Varying $A$ only</td>
<td>0.04</td>
<td>0.15</td>
<td>3.96</td>
</tr>
<tr>
<td>Varying $\lambda$ only</td>
<td>0.04</td>
<td>0.06</td>
<td>1.60</td>
</tr>
<tr>
<td>50% ↑ in $A$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varying $A, \lambda$</td>
<td>0.03</td>
<td>0.17</td>
<td>6.12</td>
</tr>
<tr>
<td>Varying $A$ only</td>
<td>0.03</td>
<td>0.12</td>
<td>3.94</td>
</tr>
<tr>
<td>Varying $\lambda$ only</td>
<td>0.03</td>
<td>0.05</td>
<td>1.60</td>
</tr>
<tr>
<td>2% ↑ in $N$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varying $A, \lambda$</td>
<td>0.05</td>
<td>0.25</td>
<td>6.17</td>
</tr>
<tr>
<td>Varying $A$ only</td>
<td>0.05</td>
<td>0.17</td>
<td>3.97</td>
</tr>
<tr>
<td>Varying $\lambda$ only</td>
<td>0.04</td>
<td>0.07</td>
<td>1.60</td>
</tr>
<tr>
<td>5% ↑ in $N$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varying $A, \lambda$</td>
<td>0.05</td>
<td>0.25</td>
<td>6.17</td>
</tr>
<tr>
<td>Varying $A$ only</td>
<td>0.05</td>
<td>0.18</td>
<td>3.97</td>
</tr>
<tr>
<td>Varying $\lambda$ only</td>
<td>0.05</td>
<td>0.07</td>
<td>1.60</td>
</tr>
<tr>
<td>10% ↑ in $N$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varying $A, \lambda$</td>
<td>0.05</td>
<td>0.26</td>
<td>6.18</td>
</tr>
<tr>
<td>Varying $A$ only</td>
<td>0.05</td>
<td>0.18</td>
<td>3.97</td>
</tr>
<tr>
<td>Varying $\lambda$ only</td>
<td>0.05</td>
<td>0.07</td>
<td>1.61</td>
</tr>
</tbody>
</table>

Note: Each column is ranked separately.

Source: Agricultural intermediate input and output are from 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010). Land size is calculated from 2009 Investigation and Evaluation of Cultivated Land Quality.

B.2 Sensitivity of $\alpha$ and $\sigma$

In Section 4.1, we briefly discuss the effects of the parameterization or estimation we choose. Here we focus on the role of the key parameters in our model: $\alpha$ and $\sigma$. The key mechanism in our model is the tradeoff between the use of agricultural intermediate input and labor. Therefore the explanation power of our model is governed by the intermediate input income share $\alpha$ and the labor income share $\sigma$. We consider several alternative values for $\alpha$ and $\sigma$ and in each case we recalibrate the model following the previous procedure.

Changes in $\alpha$ would change the relative importance of intermediate input and hence the role of financial frictions. The value $\alpha = 0.5252$ we use in the previous analysis keeps the borrow-
ing constraint of the province with the highest intermediate input share exactly indifferent between binding and not binding. Therefore it is the lowest value we can use to make sure all provinces are binding. Higher $\alpha$ would increase the explanation power of our model but the improvement should be little because there is not too much room. We run experiments for $\alpha = 0.55, 0.60, 0.65$. Table 14 shows the results. The first four columns show the factor differences in $L_a/N, Y_a/L_a$ and $Y/N$, and the last three columns show the elasticities of them with respect to $X/Y_a$. The baseline model refers to the model with all the variations in $\lambda, \theta, \kappa, Z/N$, and $A$. In general the results are quite similar for this range of $\alpha$ and there are no large differences of change in any of the values. For example, when we decrease $\alpha$ to 0.50, the factor differences in both $Y_a/L_a$ and $Y/N$ only drop 0.05 comparing to the baseline model, and elasticities barely change.

Table 14: Experiments with Different $\alpha$ and $\sigma$

<table>
<thead>
<tr>
<th></th>
<th>$L_a/N$</th>
<th>$Y_a/L_a$</th>
<th>$Y/N$</th>
<th>elasticity of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rich</td>
<td>poor</td>
<td>rich/poor</td>
<td>N</td>
</tr>
<tr>
<td>Data</td>
<td>0.06</td>
<td>0.45</td>
<td>7.49</td>
<td>5.89</td>
</tr>
<tr>
<td>Baseline Model</td>
<td>0.04</td>
<td>0.25</td>
<td>6.17</td>
<td>4.87</td>
</tr>
<tr>
<td>$\alpha$ differs only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha = 0.65$</td>
<td>0.04</td>
<td>0.29</td>
<td>7.35</td>
<td>4.98</td>
</tr>
<tr>
<td>$\alpha = 0.60$</td>
<td>0.04</td>
<td>0.29</td>
<td>7.35</td>
<td>4.96</td>
</tr>
<tr>
<td>$\alpha = 0.55$</td>
<td>0.04</td>
<td>0.29</td>
<td>7.36</td>
<td>4.94</td>
</tr>
<tr>
<td>$\sigma$ differs only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma = 0.75$</td>
<td>0.05</td>
<td>0.30</td>
<td>7.35</td>
<td>5.01</td>
</tr>
<tr>
<td>$\sigma = 0.70$</td>
<td>0.04</td>
<td>0.29</td>
<td>7.35</td>
<td>4.99</td>
</tr>
<tr>
<td>$\sigma = 0.60$</td>
<td>0.04</td>
<td>0.29</td>
<td>7.35</td>
<td>4.96</td>
</tr>
<tr>
<td>$\sigma = 0.55$</td>
<td>0.04</td>
<td>0.28</td>
<td>7.48</td>
<td>4.88</td>
</tr>
</tbody>
</table>

Note: Each column is ranked separately.

Source: Agricultural intermediate input and output are from on 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010). Land size is calculated from 2009 Investigation and Evaluation of Cultivated Land Quality.

Increasing $\sigma$ would decrease the explanation power of fixed land input, hence the overall results of our model become better. When we change $\sigma$ in an acceptable interval $[0.55, 0.75]$, all the results would change as we expected, but overall they are very similar. For example, when we decrease $\sigma$ from 0.75 to 0.60, the factor differences of $Y_a/L_a$ and $Y/N$ only drop 0 and 0.05, respectively, and all the elasticities we examine only drop 0.01.

52
B.3 Calibration Using Agricultural Debt

Now we use agricultural short-term (one year and below) loan at the end of year 2006 as $X$ to calibrate $\lambda$.\(^{62}\) For Beijing, this value is $\lambda_{\text{Beijing}} = 0.5665$. Figure 18 shows a positive correlation between the agricultural intermediate input share and borrowing constraint parameter $\lambda$.\(^{63}\) Again we consider 4 different versions of the model and report the ratio of the equilibrium outcomes between the average value of richest decile provinces and that of the poorest decile provinces (see Table 15).

In general, the results in Table 15 are similar as in Table 5 and Table 11 in which we use observed $X$ to calibrate $\lambda$. The factor differences generated by this calibration are overall similar when we use an indirect target to calibrate financial friction parameter. Comparing to the results of the direct target method, the results of agricultural employment share barely change. In Table 11 the model with both frictions can explain 99 percent and 84 percent of the factor differences in $Y_a/L_a$ and $Y/N$, here the model can still explain 95 percent and 80 percent of the factor differences in $Y_a/L_a$ and $Y/N$ respectively, i.e., only 4 percent less for

\(^{62}\)The loan data is only reported at the end of each year. The outstanding loan at the end of 2006 are used for production and will be paid back in 2007. See Appendix A.1.8.

\(^{63}\)In this figure, Shanghai is excluded because the agricultural loan value is extremely low.
Table 15: Calibration Using Agricultural Debt

<table>
<thead>
<tr>
<th></th>
<th>$L_a/N$</th>
<th>$Y_a/L_a$</th>
<th>$Y/N$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rich</td>
<td>poor</td>
<td>rich/poor</td>
</tr>
<tr>
<td>Data</td>
<td>0.06 0.45</td>
<td>7.49 5.89</td>
<td></td>
</tr>
<tr>
<td>Varying in $\lambda, \theta, \kappa, Z/N, A$</td>
<td>0.04 0.28</td>
<td>7.12 4.69</td>
<td></td>
</tr>
<tr>
<td>Varying in $\lambda, \kappa, Z/N, A$</td>
<td>0.05 0.31</td>
<td>6.39 4.75</td>
<td></td>
</tr>
<tr>
<td>Varying in $\theta, \kappa, Z/N, A$</td>
<td>0.04 0.22</td>
<td>5.37 4.54</td>
<td></td>
</tr>
<tr>
<td>Varying in $\kappa, Z/N, A$</td>
<td>0.05 0.22</td>
<td>4.36 4.57</td>
<td></td>
</tr>
<tr>
<td>Varying in $\lambda, A$</td>
<td>0.06 0.26</td>
<td>5.20 4.71</td>
<td></td>
</tr>
<tr>
<td>Varying in $A$</td>
<td>0.05 0.19</td>
<td>3.98 4.54</td>
<td></td>
</tr>
<tr>
<td>Varying in $\lambda$</td>
<td>0.04 0.11</td>
<td>2.48 1.08</td>
<td></td>
</tr>
</tbody>
</table>

Note: Each column is ranked separately.

Source: Agricultural intermediate input and output are from on 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010). Land size is calculated from 2009 Investigation and Evaluation of Cultivated Land Quality.

both indicator. Hence we can conclude that the strong correlation between financial friction and intermediate input share is also confirmed by the evidence from macro data (aggregate agricultural debt).

### B.4 Comparative Statics

In Proposition 2, we derive qualitative comparative statics of several key variables with respect to financial frictions. Due to the non-linearity, we are not able to derive comparative statics of several key variables, in particular, intermediate input use and total input. Here we provide quantitative comparative statics on these two variables based on numerical solutions. Calibrated values of all parameters except financial friction parameter $\lambda$ are used to derive numerical solutions and conduct quantitative comparative statics.

The results are presented in Figure 19. Labor input decreases and agricultural productivity increases as financial frictions become less severe, consistent with our qualitative results in Proposition 2. Moreover, intermediate input monotonically increases when financial frictions become less severe. The intuition is as follows. Intermediate input is constrained and depressed. If the financial constraint is relaxed, then intermediate input must increase. This in turn leads to a decrease of labor input through general equilibrium effects: more labor is needed to produce more non-agricultural good to meet intermediate input demand and consumption demand. Agricultural productivity increases because of a reduction of labor input and an increase of output due to an increase of intermediate input.

More interestingly, the total input goes down when financial frictions become less severe.
but goes up if the financial friction parameter is large enough. This is a direct consequence of how intermediate input and agricultural labor input respond to changes of financial situation in equilibrium. Intermediate input use will increase whenever there is a small improvement of financial situation. The magnitude of increase depends on the demand of agricultural good. Non-agricultural good market clearing condition dictates that agricultural labor input must decrease. The magnitude of reduction depends on the demand of non-agricultural good. Therefore, how total input responds to changes of financial frictions has to do with the nonhomothetic preference. The reason is as follows. Under this preference, when financial frictions are very severe, relative expenditure on agricultural good is heavily biased due to the subsistence requirement of consumption on agricultural good so that non-agricultural consumption is quite small. When financial frictions become less severe and hence income becomes higher, relative expenditure on agricultural good decreases, meaning that non-agricultural consumption will increase more than agricultural consumption.

First, let’s consider the case where the financial frictions are severe. Suppose financial friction become less severe locally. Non-agricultural consumption will increase a lot more than the increase of agricultural consumption. Hence, the effect of agricultural labor input reduction on total input dominates the effect of intermediate input increase. This explains why total input goes down as $\lambda$ becomes larger locally when financial frictions are severe. However, as argued above, magnitude of the increase of non-agricultural good consumption due

---

*Figure 19: Change of Input Cost*

*Source: Model simulation.*
to increases of $\lambda$ decreases if $\lambda$ is large enough. An improvement of financial situation will have a smaller effect on non-agricultural good consumption relative to agricultural good consumption. Hence, when $\lambda$ is large enough, an increase of $\lambda$ leads to a smaller increase of non-agricultural good consumption. This, combining with our explanation on how intermediate input use changes as $\lambda$ changes, means the effect of agricultural labor input reduction on total input decreases as $\lambda$ goes up. Therefore, when $\lambda$ becomes large enough, the effect of agricultural labor input reduction on total input will be dominated by the effect of intermediate input increase. This is what happens in the bottom left panel of Figure 19.
C Model Discussion

C.1 An Alternative Specification of the Financial Constraint

Assume that a fraction $\phi$ of wage bill has to be paid upfront. Under this assumption, farmers have to borrow more than $X$ to finance the wage bill which has to be paid in advance. The incentive-compatible constraint for the representative farmer is

$$\max_{L_a} \{ p_a Y_a - (1 + r^*) \phi w_a L_a - (1 - \phi) w_a L_a \} \geq (1 - \lambda) \max_{L_a} \{ p_a Y_a - (1 - \phi) w_a L_a \}.$$  

The first-order condition for the problem on the equilibrium path (left-handed side) is

$$\sigma (1 - \alpha) p_a \frac{Y_a}{L_a} = (1 + r^*) w_a.$$  

Obviously from the above optimality condition, variations of $\phi$ could generate variations of labor productivity even if wage has no variation.

C.2 Equilibrium Analysis

A competitive equilibrium is a set of allocations $\{L_a, L_n, c_a, c_n, X\}$, a set of prices $\{p_a, w_a, w_n\}$, and profits for firms in the agricultural sector, such that: (i) given prices and profits, $\{c_a, c_n\}$ solve the utility optimization problem of the representative household; (ii) given prices, $\{L_a, L_n, X\}$ solve the profit optimization problem of the representative firms in each sector; (iii) condition (3) holds so that the representative household is indifferent between working in the two sectors; and (iv) labor market clears:

$$N = L_a + L_n;$$  

(v) agricultural goods market clears:

$$Y_a = Nc_a;$$  

(vi) non-agricultural goods market clears:

$$Y_n = Nc_n + X.$$  

Optimization of the representative farmer means the following two conditions hold:

$$\sigma (1 - \alpha) p_a \frac{Y_a}{L_a} = w_a.$$  

57
\[ X = \min \left\{ \frac{\lambda [1 - \sigma (1 - \alpha)]}{1 + r^*} p_a Y_a, \alpha p_a Y_a \right\}. \]  

(11)

Since our focus is to account for the cross-province agricultural labor productivity differences, a few variables in the competitive equilibrium are key to the analysis: the intermediate input ratio \( X / Y_a \), the share of employment in agricultural \( L_a / N \), labor productivity in the agricultural sector \( Y_a / L_a \), and aggregate labor productivity \( Y / N \). Production function of the agricultural sector suggests the following equality:

\[ \frac{Y_a}{L_a} = (\kappa A)^\sigma \left( \frac{Z}{N} \right)^{1-\sigma} \left( \frac{X}{Y_a} \right)^{1-\alpha} \left( \frac{L_a}{N} \right)^{\sigma-1}. \]  

(12)

According to this equality, the labor productivity in the agricultural sector depends positively on the intermediate input ratio \( X / Y_a \) and negatively on the employment share in the agricultural sector \( L_a / N \). Assuming the borrowing constraint of the representative farmer is binding in the equilibrium. Then from the optimality condition (10), the binding version of the borrowing constraint (11)\(^{64}\), and the decomposed version of agricultural production function (12), by eliminating the price of agricultural goods, we have the following two equilibrium conditions:

\[ \frac{X}{Y_a} = \left[ \frac{\lambda (1 - \sigma (1 - \alpha)) (1 - \theta)}{\sigma (1 - \alpha) (1 + r^*)} \right]^{1-\alpha} \left( \frac{L_a}{N} \right)^{(1-\alpha)(1-\sigma)}, \]  

(13)

\[ \frac{Y_a}{L_a} = A^{\sigma + \alpha (1-\sigma)} \kappa^{(1-\alpha)} \left[ \frac{\lambda (1 - \sigma (1 - \alpha)) (1 - \theta)}{\sigma (1 - \alpha) (1 + r^*)} \right]^{\alpha} \left[ \frac{(Z / N)}{(L_a / N)} \right]^{(1-\alpha)(1-\sigma)}. \]  

(14)

Market clearing conditions, optimality conditions of the households’ utility maximization problem, optimality condition (10) and binding version of (11) give the following equilibrium condition:

\[ \left[ a + \frac{\lambda (1 - \sigma (1 - \alpha)) (1 - \theta)}{\sigma (1 - \alpha)} \left( \frac{1 - \theta}{\sigma (1 - \alpha)} \right) \left( \frac{L_a}{N} \right) \right] \left( \frac{L_a}{N} \right) = \frac{(1 - \theta) (1 - a)}{\sigma (1 - \alpha)} \frac{a}{(Y_a / L_a)} + a. \]  

(15)

Equilibrium conditions (13), (14) and (15) could be solved for three key variables we are interested in: \( Y_a / L_a, L_a / N \), and \( X / Y_a \). The last variable we are interested in is the aggregate

\(^{64}\)Note that the cutoff value in the borrowing constraint is captured by \( \alpha \) and \( \sigma \). Calibrated values of these two parameters show that the borrowing constraint is binding for all the richest provinces.
GDP per worker which could be obtained as follows:

\[
\frac{Y}{N} = \frac{\text{GDP}_a + \text{GDP}_n}{N} = \frac{p^*_a Y_a - X + A (N - L_a)}{N} = \frac{Y_a L_a}{\bar{a} N} \left( \frac{p^*_a - X}{Y_a} \right) + A \left( 1 - \frac{L_a}{N} \right),
\]

(16)

where \( p^*_a \) is the PPP-adjusted relative price of the agricultural goods. Our quantitative analysis is based on equilibrium conditions (13), (14), (15) and (16).

C.2.1 An Analytical Example

Before switching to the quantitative exercise, in this section, we consider a special case which gives explicit expressions to our key variables. In particular, assume \( a = 0 \). In this case, since the utility weight is zero so that only the subsistence agricultural goods consumption is needed, i.e. \( c_a = \bar{a} \). This assumption simplifies the model in the sense that we are able to have closed-form solution to our key variables: intermediate input-output ratio \( (X/Y_a) \), agriculture employment share \( (L_a/N) \), and labor productivity in agriculture \( (Y_a/L_a) \). Base on these closed-form expressions, we could see the mechanism that how the two frictions as well as other important parameters such as economy-wide productivity affect agricultural and aggregate productivity.

Agricultural good market clearing condition requires

\[
Y_a = N \bar{a}.
\]

Rewriting the condition obtains

\[
\frac{Y_a}{L_a} = \bar{a} \left( \frac{L_a}{N} \right)^{-1}.
\]

(17)

The production function and the above condition leads to the following analytical solutions:

\[
\frac{L_a}{N} = \frac{1}{\kappa A} \left( \frac{\bar{a}}{(Z/N)^{1-\sigma} (X/Y_a)^{a/(1-a)}} \right)^{1/\sigma},
\]

(18)

\[
\frac{Y_a}{L_a} = \kappa A \left( \frac{(Z/N)^{1-\sigma} (X/Y_a)^{a/(1-a)}}{\bar{a}^{1-\sigma}} \right)^{1/\sigma},
\]

(19)

\[
\frac{X}{Y_a} = \left[ \frac{\lambda \left( 1 - \sigma (1 - a) \right) (1 - \theta)}{\sigma (1 - a) (1 + r^*) \kappa} \right]^{\sigma (1-a)/(\sigma (1-a) + (1-\sigma) (1-r^*))} \left( \frac{\bar{a}}{Z/N} \right)^{\frac{(1-\sigma)(1-a)}{\sigma (1-a) (1+r^*) \kappa}}.
\]

(20)

According to our calibration, \( a \) is close to zero and much smaller than other parameters. The numerical solution of the benchmark model confirms our intuition in this simplified model.
Equation (18) and (19) state the negative relationship between agricultural employment share and intermediate input-output ratio and the positive relationship between agricultural labor productivity and the intermediate input ratio. This is consistent with our empirical findings in Section 2 and confirms our intuition that more severe financial frictions lead to less use of intermediate inputs, more use of labor (larger agricultural employment share), and consequently smaller agricultural labor productivity. Moreover, in this simplified case, financial frictions and labor market frictions affect employment share in agriculture and agricultural productivity only through intermediate input-output ratio.

Financial frictions limit farmers’ ability of purchasing intermediate inputs and indirectly driving up the price of intermediate inputs.\footnote{Mathematically, if the binding borrowing constraint becomes tighter, then the Lagrangian multiplier becomes larger. Hence the shadow price of intermediate input is higher.} This incentivizes farmers use labor to substitute intermediate inputs.\footnote{Manuelli and Seshadri (2014) explains the slow tractor adoption in the US agriculture through a similar mechanism.} A lower degree of financial frictions (larger $\lambda$) means that farmers are able to buy more intermediate inputs than they used to be.

Other crucial parameters such as economy-wide productivity, agricultural specific productivity, and per capital land endowment also affect productivity as well as employment shares. Higher productivity parameters reduce the required labor used in agricultural production for subsistence requirement and hence reduce employment share in agriculture and increase agriculture productivity. Differently, higher land-to-employment ratio requires less intermediate inputs and labor to produce the same level of agricultural output, which results in lower employment share, lower intermediate input-output ratio, and hence larger productivity in the agricultural sector.

\section*{C.3 Proof of Proposition 2}

Intermediate input labor ratio is a linear function of $\lambda$:

\[
\frac{X}{L_a} = \frac{\lambda A (1 - \theta) [1 - \sigma (1 - \alpha)]}{\sigma (1 - \alpha) (1 + r^*)}.
\]

This is a consequence of the Cobb-Douglas production function. Equilibrium conditions (13), (14) and (15) imply labor inputs are decreasing with respect to $\lambda$. Then condition (14) implies labor productivity increases as $\lambda$ increases. This completes the proof.
C.4 The Equilibrium of the Financially Unconstrained Economy

If in an economy the financial constraint is not binding (for large value of $\lambda$), the equilibrium is determined by the following equations.

\[
\frac{X}{Y_a} = \left[ \frac{\alpha (1 - \theta) (1 - \sigma)}{\sigma (1 - \alpha) (1 + r^*)} A^{1-\sigma} \frac{\lambda (1 - \sigma)}{K^{\sigma} (Z/N)^{1-\sigma}} \right]^{1-\alpha} \left( \frac{L_a}{N} \right)^{(1-\alpha)(1-\sigma)}
\]

\[
\frac{Y_a}{L_a} = A^{\sigma + \alpha (1 - \sigma) \kappa^{\sigma} (1-\alpha)} \left[ \frac{\alpha (1 - \theta) (1 - \sigma)}{\sigma (1 - \alpha) (1 + r^*)} \right]^{\alpha} \left( \frac{(Z/N)}{(L_a/N)} \right)^{(1-\alpha)(1-\sigma)}
\]

\[
= \left[ a + \left( 1 - a \left( 1 - \frac{\alpha}{(1 + r^*)} \right) \right) \right] \frac{1 - \theta}{\sigma (1 - \alpha)} \left( \frac{L_a}{N} \right)
\]

\[
= \frac{a + (1 - a) \left( 1 - \frac{\alpha}{(1 + r^*)} \right)}{\sigma (1 - \alpha)} \left( \frac{L_a}{N} \right) + a.
\]

C.5 Proof of Proposition 3

As shown in equation (18), (19) and (20), as financial frictions become less severe, the use of labor decreases and the use of intermediate inputs increases. Total input, which is the sum of values of the two inputs, decreases with respect to financial frictions parameter $\lambda$. To see this, the first-order conditions implies

\[
T = (1 + r^*) X + w_a L_a
\]

\[
= [\lambda (1 - \sigma (1 - \alpha)) + \sigma (1 - \alpha)] p_a Y_a
\]

\[
= \left( \frac{\lambda (1 - \sigma (1 - \alpha)) + \sigma (1 - \alpha)}{\sigma (1 - \alpha)} \right) w_a L_a.
\]

$T$ is total input. It follows total input is smaller when financial frictions are less severe. Intermediate input labor ratio follows the same argument in Proposition 2 since the first-order conditions do not change in the simplified model. Since agricultural output is constant and the use of labor inputs is decreasing in $\lambda$, agricultural labor productivity increases as $\lambda$ increases. This completes the proof.
D Model with Physical Capital

D.1 The Model

The benchmark model does not include physical capital in both agriculture and non-agriculture. Physical capital is considered as one of the most important factors for production in both sectors and hence potentially affects productivity differences. In particular, if physical capital variation in agriculture across different regions is large, then the benchmark model will over predict the importance of financial frictions in agriculture.\(^{68}\) In this section, we extend the benchmark model by introducing physical capital into both sectors. The basic setting is the same as in the benchmark model except that production in both sectors need physical capital as one of inputs. Before we describe the extended model with physical capital, assume that the aggregate physical capital endowment of the representative household is \(K\). We only highlight the differences between the benchmark model and the extended model with physical capital.

D.1.1 Non-Agricultural Production

The non-agricultural good is produced by both labor and capital:

\[
Y_n = K_n^\eta (AL_n)^{1-\eta}, \tag{21}
\]

where \(Y_n\) and \(L_n\) are as before and \(K_n\) denotes physical capital used as an input in non-agriculture. The economy-wide productivity parameter \(A\) is labor-augmenting. \(\eta\) denotes the income share of physical capital in non-agriculture.\(^{69}\)

D.1.2 Agricultural Production

In addition to modern intermediate inputs and labor, the agricultural production uses physical capital. The production function is

\[
Y_a = X^\alpha \left( Z^{1-\sigma-\gamma} K_a^{\gamma} (\kappa AL_a)^{\sigma} \right)^{1-\alpha}, 0 < \sigma < 1, 0 < \alpha < 1, \kappa > 0, \tag{22}
\]

where \(Y_a, K_a,\) and \(L_a\) are output of agricultural good, physical capital and labor, respectively. \(Z\) is land and \(X\) is intermediate input supplied by the non-agricultural sector. Assume that land is a fixed factor in the production function so that labor and intermediate input both

---

\(^{68}\)See Priyo (2012) for a detailed explanation.

\(^{69}\)The Cobb-Douglas production function in non-agriculture is broadly used in macro literature.
exhibit decreasing returns to scale. $\sigma$ captures the income share of labor while $\gamma$ captures the income share of physical capital in agriculture.

### D.1.3 Household

The representative household gains utility from consuming the agricultural good $c_a$ and non-agricultural good $c_n$. Assume that the representative household owns the profits of agriculture production. Physical capital and labor endowment are supplied inelastically so that the total physical capital and labor supply is equal to $K$ and $N$. The preference of each household is still represented by a Stone-Geary utility function:

$$U = a \log(c_a - \bar{a}) + (1 - a) \log c_n, \ 0 \leq a < 1,$$

where $\bar{a}$ is subsistence level of agricultural good consumption and $a$ is a utility weight over the two goods.

### D.1.4 Frictions

Two frictions are considered to be barriers that keep farmers from efficiently using intermediate inputs: labor market friction and financial friction. First, we still have:

$$w_a = (1 - \theta) w_n, \ 0 \leq \theta < 1. \ (23)$$

Second, financial friction in agriculture is embedded in the model as follows. Before production, farmers (producers of agricultural goods) have to buy intermediate goods in advance. However, the farmer is assumed to be too poor to purchase any intermediate goods upfront. The only way to finance the purchase of intermediate goods is to borrow from the financial intermediary through financial contracts. The farmer have limited commitment and he could default. In particular, after the production, the farmer could renege on the financial contracts. In that case, only a fraction $1 - \lambda$ of the output net of payments for labor and physical capital could be kept by the farmer. Denote the exogenous interest rate, physical capital rental rate, wage rate in the agricultural sector, and the relative price of agricultural goods by $r^*$, $r$, $w_a$ and $p_a$. For any $X$, an incentive-compatible financial contract requires the following condition holds:

$$\max_{L_a, K_a} \left\{ p_a Y_a - w_a L_a - r K_a \right\} - (1 + r^*) X \geq (1 - \lambda) \max_{L_a, K_a} \left\{ p_a Y_a - w_a L_a - r K_a \right\}. $$

The above condition states that a farmer must end up with no less economic resources when he fulfills his credit (left-hand side) than when he defaults (right-hand side). The degree
of financial friction is captured by parameter $\lambda$. Larger $\lambda$ means smaller default value and hence less severe financial friction in agriculture.

The farmer is financially constrained if and only if

$$\lambda \leq \frac{\alpha}{1 - (\sigma + \gamma)(1 - \alpha)}.$$  

(24)

Again, intuitively, larger $\lambda$ implies the degree of financial friction is smaller. If degree of financial friction is smaller enough ($\lambda$ larger than the above cutoff value), the farmer will not be financially constrained (optimal level of intermediate inputs will be attained).

D.1.5 Competitive Equilibrium

A competitive equilibrium is a set of allocations $\{L_a, L_n, K_a, K_n, c_a, c_n, X\}$, a set of prices $\{p_a, r, w_a, w_n\}$, and profits for firms in the agricultural sector, such that: (i) given prices and profits, $\{c_a, c_n\}$ solve the utility optimization problem of the representative household; (ii) given prices, $\{L_a, L_n, K_a, K_n, X\}$ solve the profit optimization problem of the representative firms in each sector; (iii) condition (23) holds so that the representative household is indifferent between working in the two sectors; and (iv) labor market clears:

$$N = L_a + L_n;$$  

(25)

(v) physical capital market clears:

$$K = K_a + K_n;$$

(vi) agricultural goods market clears:

$$Y_a = Nc_a;$$  

(26)

(vii) non-agricultural goods market clears:

$$Y_n = Nc_n + X.$$  

(27)

D.2 Quantitative Analysis

The section quantitatively analyzes the role of financial frictions in generating labor productivity gap. We calibration the model with physical capital through a similar way as in Section 4.1. We then conduct basically what we do in Section 4.2 to derive quantitative results and evaluate this extended model.
D.2.1 Calibration

We assume all provinces are financial constrained. We consider Beijing as our benchmark economy and calibrate most of the model parameters to match its key aspects in 2007. We need to determine the following parameters: $\theta$, $Z/N$, $A$, $\eta$, $\gamma$, $\kappa$, $\sigma$, $a$, $\bar{a}$, and $r^*$. $\alpha$, $r^*$ and $Z/N$ are the same as in Section 4.1.

Agricultural capital income share $\gamma = 0.25$ and agricultural labor income share $\sigma = 0.39$ are chosen from the average of the estimations by Chow (1993) and Cao and Birchenall (2013). Non-agricultural capital income share $\eta = 0.57$ is calculated from input-output table.

Financial friction parameter $\lambda$ is calibrated based on the following equation

$$
\lambda = \frac{1 + r^*} {1 - (\sigma + \gamma) (1 - a)} p_a Y_a, \\
(28)
$$

where $X$ is the intermediate non-agricultural input for each province in 2007 and $i$ is the index for provinces. For Beijing, $\lambda = 0.6352$.\(^{71}\)

We do not have accurate or reliable data of farmers’ wage income in China, hence we measure the wage gap $\theta$ indirectly by using average labor productivity.\(^{72}\) Given (23) we have

$$
1 - \theta = \frac{w_a}{w_n} = \frac{\sigma (1 - a) \frac{p_a Y_a}{L_a}}{(1 - \eta) \frac{Y_n}{L_n}}.
$$

For Beijing $\theta = 0.8311$.

$A$ is the economy-wide TFP parameter, which is calibrated using non-agricultural TFP based on the production function (21). The non-agricultural TFP of Beijing in 2007 is 68.46.

Given $a$, $\sigma$, and $A$, $Z/L_a$, $X/Y_a$ from the data, we choose $\kappa$ to match the agricultural output per worker for each province, for Beijing it is $\kappa = 4397493$.\(^{73}\) We follow the same way as before to calibrate $a$ and $\bar{a}$, and $a = 0.002$, $\bar{a} = 1797.0$. We use PPP-adjusted capital data in 2006 to calculate capital-to-employment ratio $K/N$. For Beijing it is $K/N = 29227970$.

\(^{70}\)We cannot use the value of $\sigma$ in Section 4.1 because here we have to differentiate labor income share with capital income share.

\(^{71}\)Notice that $\gamma$ is calibrated to match the income share data directly so that it has nothing to do with whether or not the borrowing constraint is binding. Hence, we adopt the calibration strategy that is used in the benchmark model. Since the calibrated cutoff value (0.7545) is larger than the calibrated value of $\lambda$, our strategy works here. The calibration of $\lambda$ in all other provinces is the same due to the fact that Beijing has the largest intermediate input-output ratio.

\(^{72}\)See Restuccia, Yang and Zhu (2008).

\(^{73}\)We have huge $\kappa$ because agricultural capital is very small comparing to non-agricultural capital in the data (the mean of cross-province $K_a/K_n$ is about 5 percent). Large non-agricultural capital decreases calibrated $A$ and small agricultural capital increases agricultural TFP, which results in huge $\kappa$. 
Capital per worker varies systematically across provinces, i.e., rich provinces accumulate more capital than poor provinces (see Figure 20).

Figure 20: Capital per Worker (log-scale)

Table 16 reports all the calibrated parameters.

D.2.2 Quantitative Findings

Again we examine two measures: elasticity among endogenous variables, and factor differences of these endogenous variables. In the following, we conduct several experiments by varying different parameters.

First, we only vary $\lambda$ to match observed $X/Y_a$ and keep other parameters unchanged. Figure 21 shows the scatter plot between the endogenous variables. Comparing to the model without capital (Figure 6), the elasticities of $Y_a/L_a$ and $Y/N$ with respect to $X/Y_a$ are almost the same. But the model can explain more of the change in agricultural employment share when including physical capital. The elasticity (absolute value) increases from 0.56 to 0.85, which is a 13 percent improvement.

In Figure 22 we vary both $\lambda$ and $A$ and keep other parameters unchanged. Comparing with Figure 8, we may conclude that financial frictions and economy-wide efficiency are not able to explain what are observed in the data. The reason is that we are working on a model with physical capital. Capital endowments are different across provinces. In particular, capital
Table 16: Calibration of Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>0.0687</td>
<td>Land-to-employment ratio of Beijing in 2007</td>
</tr>
<tr>
<td>K</td>
<td>29227970</td>
<td>Capital-to-employment ratio of Beijing in 2006</td>
</tr>
<tr>
<td>A</td>
<td>68.459</td>
<td>Non-agricultural labor productivity of Beijing in 2007</td>
</tr>
<tr>
<td>(\kappa)</td>
<td>4397493</td>
<td>Agricultural labor productivity of Beijing in 2007</td>
</tr>
<tr>
<td>(\sigma)</td>
<td>0.39</td>
<td>\textit{Chow} (1993); \textit{Cao and Birchenall} (2013)</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>0.25</td>
<td>\textit{Chow} (1993); \textit{Cao and Birchenall} (2013)</td>
</tr>
<tr>
<td>(\eta)</td>
<td>0.5663</td>
<td>Non-agricultural input-output table</td>
</tr>
<tr>
<td>a</td>
<td>0.5252</td>
<td>The highest intermediate input share in 2007</td>
</tr>
<tr>
<td>a</td>
<td>0.0020</td>
<td>Long-run share of agricultural employment</td>
</tr>
<tr>
<td>(\bar{a})</td>
<td>1797.0</td>
<td>Agricultural employment share of Beijing in 2007</td>
</tr>
<tr>
<td>(\theta)</td>
<td>0.8311</td>
<td>Wage gap of Beijing in 2007</td>
</tr>
<tr>
<td>(\lambda)</td>
<td>0.6352</td>
<td>Intermediate non-agricultural input of Beijing in 2007</td>
</tr>
<tr>
<td>(r^*)</td>
<td>0.0156</td>
<td>One year real interest rate in 2007</td>
</tr>
</tbody>
</table>

Source: GDP data is adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of \textit{China Statistical Yearbook} using the method from \textit{Brandt and Zhu} (2010). Land size is calculated from \textit{2009 Investigation and Evaluation of Cultivated Land Quality}. Provincial sectoral capital is from \textit{Wu} (2009).

endowment in rich provinces is relatively more than poor provinces (see Figure 20). More physical capital discourages the use of labor and agricultural employment share will smaller. This leads to a larger agricultural and aggregate productivity. Hence, when we assume all provinces have the same capital endowment as in the benchmark economy which is rich, the model is not able to generate enough employment share in poor provinces and hence productivity are much higher in poor provinces. To this end, we conduct another experiment involving different physical capital endowment in different provinces. In Figure 23 we vary \(\lambda, K,\) and \(A\) and keep other parameters unchanged. In Figure 23, we see clearly that the quantitative performance of our model is much better than in the previous experiment.

Again, we see from the first experiment, the model is not able to generate enough aggregate productivity differences. The reason is the same as we explain in Section 4.2.1: the income effect is too small when we keep \(A\) the same across provinces at the level of the benchmark economy. If we start from a poor benchmark economy, the model is able to generate a much larger productivity difference. To see this, we conduct the following experiment. In Figure 24 we vary only \(\lambda\) but use the poorest province as the benchmark economy, that is keeping \(A\) the same across provinces at the level of the poorest province. The results are quite similar to Figure 9. Financial frictions generate fewer differences in productivity and agricultural employment share because we keep physical capital unchanged.

Table 17 shows the factor differences. In the model without physical capital, \(A\) captures the effect that physical capital has in the model with physical capital. Hence, if we vary \(A\) and
keep other parameters unchanged, we are not able to get large differences in agricultural employment share and productivity. If we vary both \( A \) and \( K \), the model is able to generate differences that are comparable to what we get by varying \( A \) only in the model without physical capital (See Table 5). Financial frictions again greatly amplify the effect of economy-wide efficiency and physical capital.

Table 17: Effects of Borrowing Constraint

<table>
<thead>
<tr>
<th></th>
<th>( \frac{L_d}{N} )</th>
<th>( \frac{Y_d}{L_d} )</th>
<th>( \frac{Y}{N} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rich</td>
<td>poor</td>
<td>rich/poor</td>
</tr>
<tr>
<td>Data</td>
<td>0.06</td>
<td>0.45</td>
<td>7.49</td>
</tr>
<tr>
<td>Varying in ( \lambda ), ( A ) and ( K )</td>
<td>0.03</td>
<td>0.24</td>
<td>7.02</td>
</tr>
<tr>
<td>Varying in ( A ) and ( K )</td>
<td>0.03</td>
<td>0.15</td>
<td>4.64</td>
</tr>
<tr>
<td>Varying in ( A ) only</td>
<td>0.04</td>
<td>0.08</td>
<td>2.36</td>
</tr>
</tbody>
</table>

Note: Each column is ranked separately.
Source: Agricultural intermediate input and output are from on 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010). Land size is calculated from 2009 Investigation and Evaluation of Cultivated Land Quality. Provincial sectoral capital is from Wu (2009).

Agricultural employment share and relative price of agricultural good are not directly targeted in our calibration. This allows us to evaluate our model by comparing the model-
Figure 22: Model vs Data: Different $\lambda$, $A$ (log-scale)

Note: Both axes are in log scale.

Source: Agricultural intermediate input and output are from the 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010).

generating data and the real data. We vary $\lambda$, $A$ and $K$ and keep other parameters unchanged. $\lambda$, $A$ and $K$ are calibrated to match the data of the corresponding provinces. Figure 25 plots the comparison of agricultural employment share while Figure 26 plots the comparison of relative price of agricultural good. Our model successfully predicts the patterns on agricultural employment share and relative price of agricultural good in the data.

To sum up, by including physical capital, the quantitative results are similar although our model has more explanatory power. Comparing to the results generated by the model without capital (Table 5), the model with capital is able to generate more factor differences in $Y_a/L_a$ and $Y/N$. Financial frictions alone could account for a large fraction of the differences in agricultural employment share and productivity. The amplification effect of financial frictions is quite large.
Figure 23: Model vs Data: Different $\lambda, K, A$ (log-scale)

Note: Both axes are in log scale.

Source: Agricultural intermediate input and output are from on 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010).
Figure 24: Model vs Data: Different $\lambda$ only, Poorest as Benchmark (log-scale)

Note: Both axes are in log scale.

Source: Agricultural intermediate input and output are from on 2007 China Regional Input-Output Tables. GDP and output data are adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010).
Figure 25: Agricultural Employment Share

Note: Horizontal axis is in log scale.

Source: GDP data is adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010).
Figure 26: Relative Price: $p_a$

Note: Both axes are in log scale. $p_a = \frac{PPP_a}{PPP_n}$.

Source: PPP data is based on authors’ own calculation. GDP data is adjusted for PPP based on authors’ own calculation. Employment data is adjusted from various issues of China Statistical Yearbook using the method from Brandt and Zhu (2010).