2002

2002-8 Earnings and Labor Mobility in Rural China: Implications for China's WTO Entry

Terry Sicular

Yaohui Zhao

Earnings and Labor Mobility in Rural China: Implications for China's WTO Entry

by

Terry Sicular and Yaohui Zhao

Working Paper # 2002-8 **December 2002**

RBC Financial Group
Economic Policy Research Institute
EPRI Working Paper Series

Department of Economics
Department of Political Science
Social Science Centre
The University of Western Ontario
London, Ontario, N6A 5C2
Canada

Earnings and Labor Mobility in Rural China:

Implications for China's WTO Entry

Terry Sicular
and
Yaohui Zhao

Revised, December 2002

* The authors are Associate Professor of Economics, University of Western Ontario, and Professor of Economics, Beijing University. We wish to thank Xinzheng Shi for his assistance, Will Martin, Tom Hertel and Martin Ravallion for their helpful comments, and the World Bank for financial support. The views expressed here, and any errors or omissions, are attributable only to the authors.

1. Introduction

The literature on trade liberalization in developing countries contains divergent views regarding the impact of liberalization on employment, incomes, and poverty. While most studies find aggregate welfare gains, they disagree over the distribution of these gains among households. One view is that trade liberalization generates broad-based employment gains across regions for skilled and unskilled labor, and that consequently income gains are shared widely (Dollar and Kraay, 2001). Under these conditions trade liberalization contributes to reductions in inequality and poverty. The alternative view is that employment and income gains go disproportionately to the already better-off groups, with negative implications for inequality and perhaps also for the poor (Rodrik, 2000).

Reality likely lies somewhere in between, with the outcome depending on specific conditions in the country in question. First, the impact of liberalization depends on the level and structure of pre-liberalization trade barriers, which determine the sectors that gain and lose. A differential impact on sectors holds implications for the distribution of gains among regions, skill levels, and income groups. Second, it depends on the pre-existing distribution of assets, that is, of land, capital, and human capital. So, for example, if trade liberalization benefits agriculture, and if land holdings are highly concentrated, then inequality could increase. Third, the distributional impact of trade liberalization depends on the flexibility of domestic markets, especially (but not only) for labor. Gains from trade liberalization are less likely to be shared equally where labor markets are segmented and barriers hinder labor mobility across sectors. Thus for specific household groups such as the poor, the impact of trade liberalization depends critically on local market conditions and household endowments (Winters, 2000).

These considerations are relevant to the impact of WTO entry and concomitant trade liberalization on employment and incomes in China. In these regards the Chinese case has some interesting features. China's labor markets have historically (under socialism) been inflexible and highly segmented. Domestic economic reforms have allowed greater labor mobility, but many observers believe that substantial institutional barriers to labor movement persist. Also, in China certain assets such as land and education, while not equally distributed, are nevertheless relatively equally distributed by developing country standards. These two features would have counterbalancing effects in that the former would tend to cause the gains from WTO entry to be concentrated while the latter would tend to cause the gains to be shared more broadly compared to similar liberalizations in other countries.

This paper examines the microeconomic determinants of rural employment and incomes in China. Using survey data, we estimate income, wage, and labor supply functions for rural households in China. Since the households derive income from agriculture and sideline family businesses and since labor hired in such activities is fairly rare (see Bowlus and Sicular, forthcoming), we must impute shadow wages derived from self-employment. Together, the income, wage, shadow wage, and labor supply functions empirically describe household income generation from employment.

Our analysis fills a gap in the literature. While the literature examining employment and earnings in rural China is now quite substantial, most studies of China's rural employment analyze the determinants of occupational status, that is, whether or not individuals participate in different types of work such as wage jobs or non-agricultural sidelines (examples are Hare, 1994, 1999a, 1999b, Knight and Song, 1997, 1999; Michelson and Parish, 2000, Parish, Zhe and Li, 1995, Rozelle et al.,1999, and Zhao, 1999a, 1999b). Relatively few studies estimate rural labor supply *per se*, that is, hours or days worked (examples are Knight and Song, 1997, and Yao, 1999). Even fewer studies estimate labor supply as a function of wages, the relationship of greatest interest here.

To our knowledge the only study that estimates labor supply for rural China and includes a measure of the wage as an explanatory variable is by Meng (2000). Like Meng, we estimate time worked as a function of wages and other variables. Unlike Meng, and indeed unlike the relevant literature for other developing countries (Jacoby, 1993, Skoufias, 1994, etc.), our labor supply functions allow for the possibility that wages or shadow wages, and labor's response to these wages, can differ depending on the type of wage. That is, in our analysis household labor supply is a function of not just one but of multiple wages. Thus it is possible for labor supply to be more responsive to the market wage than to the agricultural shadow wage, or vice versa. Finally, while other studies estimate total labor supply, we estimate total labor supply and also its components, in this case labor supply to household agricultural production, to household non-agricultural production, and to wage employment. By estimating the components of labor supply, we obtain information about how wages and other variables influence the composition of employment.

Our empirical results provide some of the underlying parameters needed to understand the effects of trade liberalization on levels of employment and earnings. The income generation functions give estimates of the impact of agricultural versus non-agricultural employment on income from labor. We find that non-agricultural employment generates substantially more income per hour worked than does agricultural employment. This result is consistent with the findings of other studies (e.g., Knight and Song, 1997; Meng, 2000; Michelson and Parish, 2000), which generally conclude that income inequality among households reflects differences in access to higher-paying, off-farm jobs.

We take this analysis one step further and decompose the income gap between richer and poorer households. Our decomposition reveals that most of the income gap is accounted for not by differences in hours worked in different occupations, but by differences in estimated parameters. That is, poorer households are poorer not so much because they supply less labor to non-agricultural jobs, nor because they are less educated (although both these characteristics apply), but because the returns to the labor they supply in each occupation and the returns to their education and other characteristics are lower than the returns received by richer households. These findings suggest that WTO entry's impact on income distribution will depend not only on how it affects wages and employment structure on average, but also on how it affects the distribution of wages and of the returns to education among households.

Our labor supply estimates provide information on how work hours, and also the composition of work hours, would respond to changes in wages for agricultural and non-agricultural work. We find that wages do not have a significant effect on the total number of hours worked. They do, however, significantly influence the composition of hours worked. Our estimates of labor supply by type of employment—in household agriculture, household non-agriculture, and off-farm wage jobs—indicate that most own- and cross-wage elasticities are well below one. This suggests that while labor moves among these types of employment, mobility is limited.

A notable exception to this pattern is the cross-elasticity of wage employment with respect to the agricultural shadow wage. This elasticity is significant and large, indicating that differences in labor supply to off-farm employment among households is driven by differences in the returns to labor in household farming. Put differently, high market wages do not "pull" labor out of agriculture; rather, low marginal returns to work in agriculture "push" labor into wage employment. We discuss these results more fully below.

We begin in section 2 with an overview of aggregate trends in China's rural employment, earnings, and labor markets, with special attention to institutions and policies that affect labor mobility and the distribution of earnings. In section 3 we describe the dataset. Section 4 examines household income generation on average and analyzes the income gap between richer and poorer households. Section 5 contains our econometric estimates of wage and labor supply functions.

Throughout this paper our focus is on rural households. WTO entry can affect urban as well as rural households, but the institutional setting and economic behavior of these two types of households are substantially different, and they require separate analysis. Analysis of rural households is important because the rural sector contains most of China's population and also most of its poor. Despite increased urbanization, 64% of the population and 74% of employed persons are still classified as rural (NBS, 2001, p. 37, 39). Average per capita income in rural areas is only about one-third that in urban areas, and evidence suggests a widening of the urban-rural income gap in recent years (Yang and Zhou, 1999). A disproportionate share of China's poor population is located in rural areas.[1] Recent estimates for 1999 by Chen and Wang (2001) report a poverty rate (using the $1/day poverty line) of 24.9% in rural areas, versus only 0.5% in urban areas.

2. Rural employment and earnings in China in the 1990s: The Aggregate Picture

A review of aggregate trends in rural employment and earnings in China provides a broad context in which to interpret our microeconomic results. Since aggregate trends reflect the impact of major policy reforms (including domestic market liberalization as well as trade liberalization in advance of WTO entry), they also provide some clues regarding the potential impact of WTO entry on employment and earnings.

Developments in the 1990s on both the supply and demand side suggest movement towards fuller employment and higher earnings in rural areas. On the supply side, in the 1990s China's labor force grew slowly both overall and in rural areas. Due to population planning policies and a marked fertility decline in the 1970s, cohorts entering the labor force in the 1990s were smaller than those in the 1970s and 1980s. New cohorts entering the labor force should remain small or even decline further in the coming years due to strict population control policies adopted in the 1980s.

On the demand side, macroeconomic growth combined with policy liberalization generated new job opportunities for rural workers. Rural non-agricultural employment grew substantially. While the official data on such employment is problematic, it provides a rough indication of trends. Rural non-agricultural employment includes employment in TVEs and also in private and individual enterprises. During the 1990s employment in these enterprises grew, on average, 5.4% annually. In absolute terms, during the 1990s the increase in rural enterprise employment exceeded 66 million jobs, and by 1999 this employment was

[1] World Bank (2001) gives official Chinese estimates of the number of rural poor declining from 65 million in 1995 to 42 million in 1998. UNDP (1998) and NBS (1998) give Chinese official estimates of the number of urban poor that range from 10 to 15 million in the mid-1990s. These numbers may not be entirely comparable, but they roughly indicate that 75-85% of the poor are rural.

equivalent in number to one third of the rural labor force.

A notable aspect of growth in rural enterprise employment is that nearly half of it was due to the expansion of private and individual enterprises. By 1999 private and individual rural enterprises employed over 45 million people, equivalent to about 10% of the rural labor force. While some of this growth could be the result of reclassification of collective TVEs as private businesses, it likely also reflects the growing importance of household-based non-agricultural activities.

Also notable during the 1990s was a rise in rural-urban migration. Following a relaxation of restrictions on labor movement, rural-urban migration appears to have grown substantially. Data on migration are spotty, and definitions of what constitutes migration differ (see Wu and Zhou, 1996; Rozelle *et al.*, 1999), but most estimates suggest at least a doubling of the number of migrants between the late 1980s and mid- or late 1990s. Sources suggest that by the mid- or late 1990s the number of migrants (excluding commuters) probably exceeded 50 million, or about 10% of the number of rural employed persons (Zhao, 1999b; Wu and Zhou, 1996).

Altogether, then, growth in employment by rural TVEs, by private and individual enterprises, and through migration increased from perhaps 130 million in the early 1990s to roughly 230 million in the late 1990s. By the late 1990s, non-agricultural employment of rural residents had risen from less than 30% to nearly 50% of the number of rural employed persons.

This substantial expansion of non-agricultural employment has spurred some debate regarding the nature of labor markets in rural China. Some authors have argued that labor markets are now fairly open and competitive, with considerable labor mobility (Rawski and Mead, 1998). Others, however, argue that while open and competitive labor markets are emerging in some regional pockets, and while generally the direction of change has been towards more open, competitive conditions, rural labor markets nevertheless continue to be imperfect, and institutional and administrative barriers persist (Parish, Zhe and Li, 1995;

Knight and Song, 1999). These issues are relevant to our analysis and are discussed further below.

Discussion of employment cannot be complete without mention of the largest employer, agriculture. Official statistics reveal cycles in agricultural employment. The number of rural employed persons rose to more than 340 million in 1991-92, fell to less than 325 million in 1995-97, and then rose again to almost 330 million in 1999-2000 (National Bureau of Statistics, 2000). Interestingly, the upswings in these cycles occurred at the same time as downswings in rural enterprise employment, and vice versa, which suggests an inverse relationship between the two kinds of employment.

The official data on rural employment, however, are problematic. They simply count the number of people by primary occupation. They do not capture the fact that many rural workers engage in multiple occupations and that hours worked in any particular occupation can fluctuate over time. Some studies indicate that changes in hours worked in agriculture have been significant (Rawski and Mead, 1998; World Bank, 2001). Estimates of agricultural employment based on hours worked indicate that in the mid-1990s when rural enterprise employment was expanding rapidly, agricultural labor days worked actually rose (World Bank, 2001). These estimates indicate that growth in non-agricultural employment and agricultural employment can occur concurrently. The relationship between non-agricultural and agricultural employment is important to our analysis, as the willingness and ability of rural residents to supply labor hours to different sectors can affect the impact of trade liberalization.

What have been the effects of the above trends in employment on rural incomes and inequality? On average the real per capita income of rural households rose during the 1990s. Between 1990 and 1999 net income per capita rose about 70 percent. This growth was derived from multiple sources. Wages contributed the largest share of the increase (38%), followed by agriculture (29%) and household non-agricultural sidelines and businesses

(22%).[2] While non-agricultural sources dominated, the fact that agriculture contributed nearly one third of the increase in income is notable given that, according to the official data, agricultural employment supposedly declined.

The average trends outlined above mask changes in distribution among poorer and richer households. The general consensus is that rural income inequality increased during the 1990s, while poverty decreased. These apparently contradictory developments reflect income growth for the poor, but faster income growth for richer groups.

A recent World Bank study outlines some key aggregate factors underlying the recent decline in rural poverty, some of which are relevant from the perspective of China's WTO entry (see World Bank, 2001). First, aggregate growth in GDP appears to be important. Nationally, the reduction in poverty coincided with a period of rapid GDP growth, and poverty reduction occurred more quickly in those regions that experienced the most rapid aggregate growth. Second, the composition of growth matters. The rate of poverty reduction has been faster in regions where agricultural growth more or less kept pace with growth in other sectors. This reflects the fact that agriculture is the primary source of employment and income for the poor. These points suggest that the effect of WTO entry on poverty will be influenced by its impact both on aggregate growth and the sectoral composition of that growth.

3. The Survey Sample

For our empirical analysis we use data from the China Health and Nutrition Survey (CHNS). CHNS data were collected through an independent survey conducted by an international team of researchers collaboratively sponsored by the Carolina Population Center at the University of North Carolina at Chapel Hill, the Institute of Nutrition and Food Hygiene, and the Chinese Academy of Preventive Medicine. Data are available for four years, 1989, 1991, 1993 and 1997. The survey used a multistage, random cluster sampling

---

[2] Income data in the text are deflated using the rural consumer price index and are taken from National Bureau of Statistics, various years; National Bureau of Statistics Rural Social and Economic Survey, 2000.

method and covers about 3800 households with 14,000 individuals in nine provinces with different geographic and economic characteristics.[3] The CHNS also includes information on community-level variables such as market prices, health facilities, and social services.[4]

The CHNS survey data are useful here because they include detailed information on incomes and hours worked in different occupations as well as on a wide range of relevant individual, household and community characteristics. The sample includes as household members migrant workers who work and live out of town but whose earnings and expenses are considered part of the household's. The data do not, however, allow us to distinguish between work and income from migrant versus local employment. Thus while we are able to investigate mobility among types of work (in household agricultural production, household non-agricultural production, and wage employment), we cannot investigate geographical mobility.

For our estimations we use the most recent or 1997 data, and we drop urban households. That is, we use a sub-sample of households that includes only those that reside in rural and suburban villages and in county towns. We include suburban villages and county towns because they have close ties with rural areas, and a significant portion of the population in these areas holds a rural *hukou*. This sub-sample covers 3,239 households containing 8,590 working-age adults. Since we are interested in labor supply, we further restrict our sample to households that have positive levels of labor time for at least one household member. This reduces the sub-sample to 2,998 households with 8,326 working-age adults. The number of observations actually used in our analysis varies among regressions depending on the extent of participation in the activity being analyzed and on the prevalence of missing values in relevant variables.

Table 1 contains descriptive statistics for the non-urban sub-sample of the CHNS survey. It also gives some comparable statistics from the 1997 NBS official rural household

---

[3]Provinces covered are Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong.

[4]Detailed information about the CHNS is available at the website www.cpc.unc.edu/china/home.html.

survey, where available. Household size and structure are similar for the NBS data and for CHNS sub-sample. Income levels for the CHNS sub-sample is somewhat lower than for the NBS survey, but this could reflect differences in how income is calculated. The structure of income also differs. In particular, agricultural income is noticeably higher in the NBS survey than in the CHNS samples. This could reflect the inclusion of households in suburban villages and county towns in the CHNS sub-sample, but not in the NBS rural survey. For both the NBS and CHNS samples the major sources of income are agriculture and wage employment.

The CHNS data provide some information about relative earnings in different sectors. Dividing average earnings by hours worked for each sector suggests that the returns to labor in agriculture is lower than in other occupations. Average net earnings per hour worked in agriculture are roughly 1.2 yuan, as compared to 2.2 yuan in non-agricultural sidelines and 2.8 yuan in wage employment. As seen below, when we use regression analysis to estimate the returns to labor, these earnings differentials become larger.

4. Income generation functions

A common approach to analyzing the determinants of income is to estimate an income generation function, where net income is a function of labor inputs, land and capital assets, and other household or regional characteristics that contribute to the generation of earnings. This approach takes income generation as a simple, linear accounting relationship where income equals the sum of household labor and other assets times the returns to those assets. The regression coefficients provide estimates of the marginal returns to each asset. We use this approach both to examine income generation on average and also to explore differences between poorer and richer subgroups.

The first columns (Model I) of Table 2 contain results from income generation regressions for all households in our CHNS non-urban sub-sample with at least one working adult. The dependent variable is the sum of net earnings from the three occupations—net earnings (revenue minus non-labor variable costs) from household agricultural production, net earnings from household non-agricultural production, and wage income. Explanatory variables include actual labor hours in each of household agricultural production, household non-agricultural production, and wage employment, as well as of education, the ratio of males to females, age, land and equipment assets to capture non-labor inputs that contribute to income from household production, and provincial dummy variables. Mean values of these variables are shown in Table 3.

Of central interest are the returns to labor in different occupations. The estimated coefficient for agricultural hours of work is positive but small and statistically insignificant. The coefficient for labor hours in non-agricultural household production is 1.69, and that for hours in market employment is 2.30, both statistically significant. These results indicate that income differences among households are generated more by non-agricultural work than by agricultural work; furthermore, wage employment brings the highest returns.[5] These results are consistent with the findings of other studies for China (Knight and Song, 1997; Meng, 2000; Michelson and Parish, 2000).

Our analysis also shows that education has a significant effect on household income. An additional year of schooling for the most educated worker in the household increases household net income by more than 330 yuan, roughly 4% of average net income. Again, this result is consistent with findings in other studies.

The average age of household workers has a significant, small positive effect on household earnings. We do not find a significant effect from the sex ratio of the workforce. As expected, land and equipment assets have significant, positive effects. An additional *mu* of land increases net income by 75 yuan or 1% of mean income, and an additional 1000 yuan

[5] Note that for agricultural and non-agricultural household production, these estimated returns are substantially lower than average earnings per hour worked given earlier (1.2 and 2.2 yuan, respectively). This could reflect that in household production the average and marginal returns to labor differ, and also that the income generation function controls for non-labor inputs. The estimated return to wage labor hours is also somewhat lower than average earnings (2.8 yuan).

of equipment assets increases net income by 151 yuan or 2% of mean income. Location of residence is also important, as some provincial dummy variables are significant and large.

We explore differences between poorer and richer households using a modified Oaxaca-type decomposition. This involves first estimating income generation functions for richer and poorer sub-groups. Models II and III in Table 2 contain the results of income generation equations for households in the top 30% and bottom 30% of the net earnings distribution. The two groups have strikingly different coefficients. The returns to labor, education, land and equipment are all significantly higher for the top 30% than the bottom 30%. For the bottom 30%, the returns to agricultural labor and to non-agricultural labor in household production are negative, albeit small in magnitude. This would be consistent with a standard surplus labor story. Land and equipment also have negative, small coefficients. This could be due to heterogeneity in the quality of productive characteristics or could reflect the lower marginal returns to these assets for poorer households. In contrast, for the top 30% coefficients on the three types of labor and also on land and equipment are all positive and significant.

The Oaxaca decomposition combines these estimated coefficients with mean values of the variables for the two groups. Mean values of the variables appear in Table 3. Here one also finds substantial differences between the richer and poorer households. Net earnings of the richest 30% are nearly twenty times that of the poorest 30%. Maximum education in richer households is 2.8 years more than in poorer households. Land assets are slightly larger for poor households (although not corrected for land quality), but equipment assets are substantially smaller.

Employment also differs. Labor time for the poorer group is overwhelmingly in household agriculture. The richer group works in all three sectors, but wage employment dominates. Also notable is the fact poorer households work fewer hours in total than do the richer households. Total work time for the poorer group is 409 days as compared to 601 days for the richer group. Work days per adult in the poor group average 149, as compared to 222 in the rich group.[6] Time worked in total and per adult is nearly 50% higher in the rich group than in the poor group.

Table 4 gives the results of our Oaxaca-type decomposition. While conventional decomposition combines the contribution of regional dummy variables together with that of other characteristics, we separate out the effects of regional differences on the income gap. The overall difference of 16,168 yuan in mean net earnings is thus decomposed into three components: (1) the income gap due to differences in levels of household productive characteristics other than regional location, (2) that due to differences in returns to these household productive characteristics, and (3) that due to regional factors.

The results show that all three components contribute to the income gap, but the differences in the returns to productive characteristics are most important. This component, which reflects differences between the two groups in the estimated coefficients, accounts for 64% of the income gap. The largest contributor here is the difference between the rich and poor in the return to education, which by itself explains 21% of the income gap. Differences in the returns to age, agricultural labor, and wage employment are also important.

Differences between the two groups in levels of observable productive characteristics explain 21% of the income gap between the two groups. The largest contributor here is the difference in wage employment hours, which accounts for 17% of the income gap. The richer households have fewer labor hours in agriculture (see Table 3), thus the quantity difference in agricultural hours serves to reduce the income gap by 4%. Differences in education levels explain 4% of the gap. Differences in mean endowments of land and productive equipment, and in household structure and age, are relatively unimportant.

The third component is the income gap attributable to regional factors. Regional location, like other household characteristics, contributes to the income gap in two ways: first, the returns to region of residence or estimated coefficients for the provincial dummy

---

[6]A workday is assumed to contain eight hours. On average the poor households contained 2.74 adults and the rich households 2.71 adults.

variables differ between the rich and poor groups, and second, the regional distribution of the poor differs from the regional distribution of the rich. Together these regional factors explain about 15% of the overall income gap, and virtually all of this is due to differences in the returns to location.

In summary, our decomposition analysis shows that poorer households are poor relative to the richer households not only because they have different characteristics than richer households, but also because the returns to their characteristics are lower. Here the key contributors to the income gap are lower returns received by poor households for education, agricultural labor, wage employment, and age (a proxy for work experience). Also notable is the lower supply of wage labor by poorer households.

5. New estimates of labor supply and allocation

The above estimates and decomposition indicate that the returns to labor and the pattern of employment are important factors underlying income differences. This raises questions regarding what explains the level and pattern of labor supply, and also about why the returns to labor vary among sectors and households. Clearly these two questions are interrelated, as household labor supply may depend on the returns to labor, and the returns to labor (especially in household production) may depend on household production decisions, including decisions on the allocation of labor. Here we pursue these questions by empirically analyzing the determinants of the returns to labor and of labor supply, and we employ instrumental variable methods that address the endogeneity of key explanatory variables.

We begin with a model of time allocation by rural households. The family is assumed to have endowments of workers, land, and other assets. The household may allocate its resources among three possible income-generating activities: agricultural production, non-agricultural production, and market or wage employment. In the first two activities the household organizes production using certain quantities of family assets and labor and perhaps also purchased inputs, hired assets, and hired labor. In market employment the household faces an exogenous wage rate set by the market.

Labor supply decisions are the outcome of household utility maximization, where utility is a function of the consumption of leisure and goods and the budget constraint depends on full income, which includes profits from household production as well as the value of the household's time endowment. Where markets function well and factors are perfectly mobile, households maximize utility by first maximizing profits at market prices, and then deciding on optimal levels of leisure and goods consumption given market prices and their endowments. Labor allocation to the different activities—hours worked in agricultural production (H_A), non-agricultural production (H_N) and wage employment (H_M)[7]—are the outcome of this decision process. In theory the allocation of time among these activities should equalize the returns to labor in them. Optimal leisure consumption is the difference between the household's time endowment and total time worked in all the activities.

Where markets are imperfect, where transactions costs or barriers to mobility are present, or where households have preferences for certain types of work, the returns to labor in different activities may no longer be equal. In such situations, moreover, the returns to labor in household production may not be observable. These returns or shadow wages will be endogenous and a function of both production- and consumption-side variables. Households thus may simultaneously face three different prices of labor, specifically, a shadow wage for labor in agricultural production (W_A), a shadow wage for labor in non-agricultural production (W_N), and an observed wage for market employment (W_M).

Since factor markets in China are likely to be imperfect, we adopt an empirical strategy that allows for this possibility. Specifically, we use the approach of Jacoby (1993), who notes that at the household optimum, the household's shadow price of labor would equal the value of the marginal product of labor in household production. Labor supply, then, can be modeled as a function of the shadow wages. We thus specify labor supply as a function of

[7] Note that wage employment includes wage labor in agriculture, but in our sample hired farm labor is a small proportion of total wage employment.

shadow wages, as well as of other relevant variables. Also, since we are interested in labor supply to three activities, we have three different labor supply functions. These take the form

$$H_i = f^i(\hat{W}_A, \hat{W}_N, W_M, I, X), \text{ i=A, N, M,} \tag{5.1}$$

where $\hat{W}_A, \hat{W}_N, W_M$ are the shadow wage in agricultural production, the shadow wage in non-agricultural production, and the wage rate in market employment, respectively. I is household non-labor income, and X is a vector of household characteristics. Total labor supply is the sum of these three labor supplies.

Empirical specification

The specification of our labor supply functions basically follows that of Jacoby and others,[8] except that we include multiple wages. The functions for labor supply to household agricultural production, household non-agricultural production, and market employment are, respectively,

$$\log H_A = \alpha_0 + \alpha_A \ln \hat{W}_A + \alpha_N \ln \hat{W}_N + \alpha_M \ln W_M + \alpha_I I + \alpha_X X + \varepsilon_A \tag{5.2}$$

$$\log H_N = \beta_0 + \beta_A \ln \hat{W}_A + \beta_N \ln \hat{W}_N + \beta_M \ln W_M + \beta_I I + \beta_X X + \varepsilon_N \tag{5.3}$$

$$\log H_M = \gamma_0 + \gamma_A \ln \hat{W}_A + \gamma_N \ln \hat{W}_N + \gamma_M \ln W_M + \gamma_I I + \gamma_X X + \varepsilon_M \tag{5.4}$$

Overall labor supply (H) is the sum of labor supply into three sectors. Alternatively, it can be estimated directly as

$$\log H = \delta_0 + \delta_A \ln \hat{W}_A + \delta_N \ln \hat{W}_N + \delta_M \ln W_M + \delta_I I + \delta_X X + \varepsilon \tag{5.5}$$

The signs of the effects of the three wage rates on total labor supply depend on the relative magnitude of substitution and income effects.

The estimation of equations (5.2-5.5) requires knowledge of wage or shadow wage rates in all activities. For those engaged in agricultural or non-agricultural self-employment, shadow wages can be calculated using estimated parameters from household production functions (Jacoby, 1993). We therefore begin with estimates of production functions in

---

[8] Other studies taking this approach include Skoufias (1994) and Abdulai and Regmi (2000).

section 5.1. Section 5.2 presents our estimates of the wage and shadow wage functions. We use these estimates to project (shadow) wages for households that do not participate in one or more activities. These projected (shadow) wages are employed in our analysis of labor supply, which appears in section 5.3.

5.1 Estimation of production functions

We use the standard Cobb-Douglas functional form to estimate the production functions for agricultural and non-agricultural activities. For each of these activities the production function takes the form

$$\ln Q = \sum_{j=1}^{n} \alpha_j \ln M_j + \sum_{k=1}^{m} \gamma_k Z_k + \varepsilon \tag{5.6}$$

where Q is the total value of output produced by the household. M is a vector of production inputs, which includes family labor hours, the value of variable costs (including materials and hired labor),[9] the value of fixed capital, and land used in agricultural production. Z is a vector of other control variables, including maximum education of household workers, [10] the average age of household workers, and provincial dummy variables.

Each production function is estimated using three estimation methods, OLS, IV and selectivity-bias corrected estimation. In the IV or 2SLS estimations, labor and variable cost inputs are treated as endogenous. Instruments used in both the agricultural and non-agricultural production functions are exogenous variables for the household and the community, including household composition variables (number of working-age adults, children, and the elderly), and local market prices of vegetables, pork, chicken, and gasoline. In the non-agricultural production function we also include local market prices of honey-combed coal briquet, coal lumps, coal powder and liquefied natural gas, and the market wage for unskilled labor as instruments. The selectivity-bias corrected equation treats whether the household engages in the activity as an endogenous decision. Here the identification variable

---

[9] The CHNS data does not distinguish hired labor from other variable inputs.

[10] We also ran specifications that included the average education of household workers along with maximum education, and also by itself. In almost all cases the coefficient on average education was not significant.

is the number of dependents in the household, and we use the Heckman method (Heckman, 1979). As noted in tables 5 and 6, the IV estimation does not pass the Hausman joint exogeneity test for either agricultural or non-agricultural production. Selectivity terms in the selectivity-bias correction equations (lambda statistics reported for Model III) are also statistically insignificant for both equations.

The first column of Table 5 reports OLS estimates of the agricultural production function. The results indicate that variable inputs are the dominant contributor to agricultural output, with an elasticity of 0.46. Land is the next most important contributor to agricultural production with the elasticity being 0.32. Labor hours have an elasticity of 0.19. The coefficient for agricultural equipment is also significant. These estimates are similar in magnitude to those in other studies (e.g., Yang, 1997; Li and Zhang, 1998). The coefficients of the above inputs add up to 0.999, which indicates that the technology displays constant returns to scale. Note also that education contributes positively to agricultural output—an additional year of education increases output by 2.1 percentage points. Comparing columns (1)-(3) reveals that the instrumental variable and selectivity-bias corrected estimates are very similar to the OLS estimates.

The first column of Table 6 reports OLS estimates of the non-agricultural production function. The elasticity of labor inputs is 0.63, more than triple the labor coefficient for agriculture. Output value is less responsive to variable costs in non-agricultural production than in agricultural production. The coefficient of equipment is the same as in agriculture, 0.033. Education is statistically insignificant. Interestingly, the non-agricultural production function displays decreasing returns to scale. This suggests that rural households may encounter difficulties in expanding the scale of household non-agricultural production (or perhaps higher income households under-report such income). As in Table 5, comparing columns (1)-(3) shows that the selectivity-bias corrected estimates are similar to the OLS estimates. In the IV regression the coefficients on labor and variable costs differ from those in the other specifications, but the Hausman test indicates that the overall the results are not significantly different from OLS.

While the Hausman tests do not provide strong support for preferring the IV results to OLS, below we nevertheless use the results from the IV specifications to calculate shadow wages. We prefer the IV results because in theory output, labor hours and variable costs are jointly and endogenously determined, and so in principle IV methods are needed to obtain unbiased results.

5.2 Estimation of wages and shadow wages

We are interested in learning how rural households respond to differentials between agricultural and non-agricultural wages. Although we observe hourly wage rates received by employed workers, the same is not true for workers engaged in agricultural or non-agricultural household production. We follow the approach of Jacoby (1993) and derive shadow wages from the production function estimates. Specifically, we calculate the shadow wage rates, or marginal products, of family labor hours from the IV estimates of the Cobb-Douglas production functions in tables 5 and 6 as follows:

$$\hat{W}_j = \frac{\hat{\alpha}_j Y_j}{L_j}, \text{ j = agriculture, non-agriculture.} \tag{5.7}$$

Y is the value of output and L is family labor hours.

Equation (5.7) allows us to calculate average shadow wages for households that engage in agricultural or non-agricultural production. These shadow wages are derived from household production functions and thus are at the household, rather than individual, level.

Households compare expected wage differentials among sectors when deciding their sectoral labor allocation. After a selection is made, only wages in the actually selected sectors are observed. To estimate the expected wage differential, we need the counterfactual wage rates for households not participating in a sector. In order to obtain counterfactual wage rates, we regress the shadow wages computed as outlined above against a range of household characteristics that potentially affect the productivity of workers. These include all exogenous variables in the production functions, the variables used to instrument labor, and variable costs. We then use the results from these regressions to predict shadow wages for all households.

Note that the projected shadow wages are used later in our labor supply functions, and so some of the explanatory variables here will serve as instrument variables in our labor supply regressions. Therefore, because all second-stage variables should be included in the first-stage regressions, we also include all exogenous variables from the labor supply equations in the shadow wage equations. These variables include non-labor income, the number of dependents, education and age of workers, health status of household members, and a few community level variables that measure the prevalence of non-farm activities.

Table 7 reports estimates of the shadow wage equations for agricultural household production and non-agricultural household production. We use OLS regression because, after including all the variables mentioned above, it is difficult to find additional identification variables for use in 2SLS. The agricultural shadow wage function has good explanatory power and the results are reasonable. As expected, agricultural equipment and land are significant in enhancing agricultural labor's marginal productivity. Maximum schooling is also positive and highly significant.

The explanatory power of the non-agricultural shadow wage function is low, perhaps reflecting the smaller number of observation and heterogeneity of labor inputs in non-agricultural activities. Here the most significant variable is non-agricultural equipment.

As is the case for employment in household production, not all households or individuals participate in wage employment. We therefore also need to predict market wages for non-participants. In addition, market wages are reported at the individual level, but in order to analyze household labor supply we need a measure of the expected wage at the household level. To address these issues, we first estimate a wage equation for all workers participating in wage employment. We then use these estimates to predict wage rates for all workers in the sample, after which we aggregate for each household the predicted individual wage rates of family members to obtain a household-level wage.

Table 8 gives the estimates of wage functions for the employed workers. Our specification follows the standard approach in the labor literature, with some additional variables relevant to the China case added. The dependent variable is the log of hourly wages. Independent variables include variables capturing individual human capital or labor quality (education, age, age squared and health status), individual characteristics (marital status and sex), and regional dummy variables. We allow for possible sample selection bias by applying the Heckman method, but the results show that selectivity is not important.[11]

The employed wage function indicates that education receives a rate of return of 3%. This effect is statistically significant at 1% level. The age-earning profile has a concave shape, with maximum wage reached at the age 48. Marriage has a positive but statistically insignificant effect on wage. Given marital status, female workers earn 20.3% less than their male counterparts. Those who have urban hukou status earn 11.0% more than those who do not. Using the coefficients from the selectivity-bias corrected regression, we predict wages for all workers. To obtain a household-level wage, we calculate the weighted average of predicted wages for the workers in each household, where the weights are based on each worker's hours of work.

Table 9 gives the mean predicted shadow wages and market wages of all households, and, for comparison, computed shadow wages and observed market wages of participating households. Not surprisingly, the mean wages of participating households are higher than those predicted for all households. This reflects that participation is more likely at higher wage rates. The returns to labor are lowest for agricultural production, followed by non-agricultural production and wage employment. These numbers are broadly consistent with the estimated returns to work time from the income generation regressions reported above.

---

[11] Estimated coefficients for Models I and II are nearly identical despite the significance of the number of dependents in sector choice. Note that lambda is also highly insignificant. This indicates that there is no selectivity bias.

5.3 Estimation of labor supply functions

Our household labor supply functions follow the form of equations (5.2) to (5.5). For the dependent variable we use hours per adult rather than total hours so as to avoid possible correlation between household labor supply and the number of working-age adults in the household. All specifications are estimated using 2SLS (IV), with the wage and shadow wage equations reported above serving as the first-stage regressions. Since a large number of households supply zero hours to non-agricultural production and market employment, for these functions we also run Tobit regressions and report the marginal effects for all households and the standard errors of the marginal effects.

Our results appear in Table 10. The first columns give estimated coefficients for total labor supply per adult. These results indicate that total labor supply is not sensitive to the marginal returns to labor in any of the three sectors. The income effect is negative and significant at the 10% significance level, indicating that leisure is a normal good. Interestingly, total labor supply per adult is higher the more dependents (elderly and children) a household has. This suggests that the need to earn income to support dependents outweighs the need to spend non-earning time caring for dependents.

While the wage variables do not have a significant effect on total labor supply, they affect the allocation of labor among different activities. The shadow wage for work in household agricultural production is, as expected, positive for agricultural and negative for non-agricultural and market labor supply. It is only significant, however, for market employment, where it has a fairly large coefficient. A one percent increase in the shadow agricultural wage decreases labor supply to wage employment by 2.6%.

The effect of the shadow wage for work in non-agricultural household production is negative and significant for agricultural labor and positive and significant for both non-agricultural and market labor supply. In other words, households with higher returns to labor in non-agricultural sidelines work fewer hours in agriculture, and more hours in both non-agricultural and market employment. Here the elasticities are small for agriculture and non-agriculture, and close to one for market employment. The signs on the coefficient for the market wage are the same as those for the shadow non-agricultural wage, but this variable is only significant in the equation for non-agricultural labor supply.

Overall these results reveal complementarity between employment in non-agricultural household production and off-farm wage jobs, but substitution between agriculture and both types of non-agricultural employment. Furthermore, the estimated coefficients on wage variables are larger in magnitude and more significant in the non-agricultural labor supply equations than in the agricultural labor supply equation. In other words, agricultural labor supply does not respond much to wage levels, while the supply of labor to non-agricultural forms of employment is more sensitive to wage levels.

From the wage elasticities we can calculate the elasticity of labor transfers in response to relative wage differentials among sectors.[12] Subtracting equation (5.2) from (5.3) gives the following expression:

$$\begin{aligned}\log(H_N / H_A) &= (\hat{\beta}_0 - \hat{\alpha}_0) + (\hat{\beta}_A - \hat{\alpha}_A)\ln \hat{W}_A + (\hat{\beta}_N - \hat{\alpha}_N)\ln \hat{W}_N + (\hat{\beta}_M - \hat{\alpha}_M)\ln W_M + \ldots \\ &= (\hat{\beta}_0 - \hat{\alpha}_0) + (\hat{\beta}_A - \hat{\alpha}_A + \hat{\beta}_N - \hat{\alpha}_N)\ln \hat{W}_A + (\hat{\beta}_N - \hat{\alpha}_N)\ln(\hat{W}_N / \hat{W}_A) \\ &\quad + (\hat{\beta}_M - \hat{\alpha}_M)\ln W_M + \ldots \end{aligned} \tag{5.8}$$

Similarly, subtracting (5.2) from (5.4) gives

$$\begin{aligned}\log(H_M / H_A) &= (\hat{\gamma}_0 - \hat{\alpha}_0) + (\hat{\gamma}_A - \hat{\alpha}_A + \hat{\gamma}_M - \hat{\alpha}_M)\ln \hat{W}_A \\ &\quad + (\hat{\beta}_N - \hat{\alpha}_N)\ln \hat{W}_N + (\hat{\beta}_M - \hat{\alpha}_M)\ln(W_M / \hat{W}_A) + \ldots \end{aligned} \tag{5.9}$$

Using these expressions and the results in Table 10, we calculate the marginal response of $\log(H_N / H_A)$ to $\log(\hat{W}_N / \hat{W}_A)$ as 0.626 (=0.50+0.126), and the marginal response of $\log(H_M / H_A)$ to $\log(W_M / \hat{W}_A)$ as 0.645 (=0.465+0.180). These numbers imply that a one percent increase in the ratio of either non-agricultural wage to the agricultural wage raises the ratio of that non-agricultural work to agricultural work by about 0.6%.

[12] We thank Martin Ravallion for suggesting this approach.

Note that these numbers assume that the change in the wage ratio is due to a rise in the non-agricultural shadow or market wage, with the agricultural shadow wage remaining constant. Due to the presence of cross-wage effects, shifts caused by changes in the agricultural shadow wage (holding the non-agricultural and market wages constant) would be different. Here the relevant calculations are as follows: the marginal response of $\log(H_A / H_N)$ to $\log(\hat{W}_A / \hat{W}_N)$ is 0.24 (=0.07+.168), and the marginal response of $\log(H_A / H_M)$ to $\log(\hat{W}_A / W_M)$ is 2.67 (=0.07+2.60). In other words, a one percent increase in the ratio of the agricultural shadow wage to the non-agricultural shadow wage would increase the ratio of labor in agricultural to non-agricultural production by 0.24%. For an increase in the ratio of the agricultural shadow wage to the market wage, the response would be much larger, 2.7%.

These numbers indicate that, for the most part, the effects of relative wages on sectoral labor allocation are not overly large or significant (relative to the standard errors). The one exception is the case where a change in the agricultural shadow wage causes a change in the ratio of the agricultural shadow wage to the market wage. In this case, shifts in labor allocation between agriculture and off-farm employment may be relatively large.

These findings suggest that labor movement out of agriculture into other sectors is not driven by the "pull" of rising non-agricultural wages, but would be driven by the "push" of lower agricultural wages. Why would these pull and push effects differ? One explanation is that for rural families income from agriculture provides the base income that ensures that the household achieves some acceptable or minimum standard of living. If agricultural income falls short—for example, in the event of drought, which would reduce overall agricultural income as well as the returns to labor in agriculture—households may feel compelled to seek income from other sources. They do not typically seek that additional income from non-agricultural self-employment, because such activities usually require start-up time and some initial investment. Casual wage jobs, however, entail lower initial costs and generate cash income more quickly. Field interviews with rural households who have experienced shortfalls in agricultural income indeed indicate that these households often send family members out to find casual wage jobs, sometimes seasonal, in construction, services, and other unskilled occupations.

Several non-wage variables have significant and interesting effects on labor allocation. Non-labor income is significant and negative for labor supply to both agricultural and non-agricultural household production, but is significant and positive for market employment. This result could reflect that higher income increases access to market employment by, for example, providing the resources needed to cover initial and search costs associated with obtaining wage jobs. It could also reflect that households gain some status or utility from off-farm work. The magnitudes of these income effects, however, are fairly small. Not surprisingly, the number of dependents significantly increases labor supply to both types of household production, but reduces labor supply to off-farm jobs. One more dependent in the household increases working hours that an adult devotes to agricultural production by 12%, increases hours devoted to non-agricultural production by 17%, and reduces working hours in market employment by 21%.

Interestingly, several other variables that are not significant for total labor supply have effects on the allocation of labor among activities. Education reduces hours worked in agricultural production by 4%, increases hours in non-agricultural production by 8%, and increases hours in market employment by a substantial 29%. Age does affect labor supply to agricultural and non-agricultural production, but it significantly reduces hours supplied to market employment. Health has a significant impact on labor allocation. If all family members are in good health, then the family supplies more labor to off-farm jobs and less to agriculture.

Community-level variables, while not significant in explaining total labor hours, also affect the composition of labor supply. We use a dummy variable for the availability of TVEs and the number of self-employed household business as indicators of the prevalence of non-farm activities in local communities. As expected, both variables are significant in reducing labor allocation into agricultural production and raising participation in non-agricultural production and market employment. Also, market prices for fuel and rice are negative and in some cases significant for non-agricultural and wage labor supply.

In general, our estimates indicate that labor supply to agriculture and labor supply to wage employment have opposite responses to most explanatory variables. In other words, increases in wage employment would tend to be accompanied by decreases in agricultural employment, and vice versa. This conclusion is consistent with aggregate statistics that show such opposing trends in aggregate employment in the two sectors. Labor supply to non-agricultural household businesses, however, for some variables moves with agriculture and for others with wage employment. Thus, depending on which determinants of labor supply are driving aggregate trends, we might expect to observe concurrent growth in agricultural and non-agricultural household employment, or growth in one and decline in the other.

6. Conclusions

The impact of trade liberalization in China depends not only on the aggregate sectoral and price shifts that follow liberalization, but also on the microeconomic response of households and individuals to those shifts. In this paper we analyze the microeconomics of household earnings and employment in rural China. Using household survey data with broad regional coverage, we examine household incomes, wage determination, and labor supply in total and among sectors.

Our analysis of income generation provides information about earnings from labor. We find that the returns to non-agricultural labor hours are higher than for agricultural labor hours, a result that is consistent with other studies in the literature. We then decompose the income gap between the top and bottom thirds of the income distribution. The decomposition reveals that poorer households are poor relative to the richer households not only because their employment patterns and other characteristics differ from those of richer households, but also because the returns to their labor and other characteristics are lower. Thus the problem is not only that poorer households have lower levels of off-farm employment, but also that the returns to their work and also to other characteristics are lower than for richer households.

Key contributors to the income gap are the lower returns received by poor households for education, for agricultural labor, and for market employment. Also important is the fact that poorer households work fewer hours worked in higher-paid market employment. These results suggest that analysis of the impact of trade liberalization should consider not only the fact that the level and composition of employment differ between income groups, but also that the returns to employment in each occupation can differ.

Our analysis of household labor supply provides some insights into how changes in wages following WTO entry might affect patterns of employment and thus earnings. Total labor supply is not sensitive to changes in wages, which suggests that any changes in wages resulting from WTO entry would not have much impact on overall employment. We also find that, for the most part, the allocation of labor between self-employment in agriculture, self-employment in non-agriculture, and off-farm wage jobs is not overly sensitive to changes in wages. In particular, labor supply to both agricultural and non-agricultural production is inelastic with respect to most wages. Put differently, changes in wages do not cause large changes in the amount of labor supplied to household production. These findings are consistent with the view that in labor is not overly mobile.

Off-farm labor supply is more responsive, at least to the agricultural shadow wage. A lower agricultural shadow wage is associated with substantially increased employment in off-farm jobs. Thus if WTO entry raised wages for off-farm jobs, it would not "pull" labor into wage jobs; however, if WTO entry reduced the returns to agricultural labor, it might "push" labor into wage jobs.

While in general labor supply is not overly sensitive to wages, it is sensitive to other variables. Household characteristics such as education, health, the number of dependents, and regional location have significant effects on the level and/or sectoral composition of labor supply. Higher education and better health are associated with less work time in agriculture and more work time in non-agricultural types of work. More dependents are associated with more work time overall, as well as with more work time in household production and less in off-farm wage employment. These findings suggest that labor mobility among sectors is significantly influenced by variables such as education, health, and demographic structure. Such variables are not directly affected by WTO entry, at least in the short term.

References

Abdulai, Awudu and Punya Prasad Regmi, 2000, "Estimating Labor Supply of Farm Households under Nonseparability: Empirical Evidence from Nepal," *Agricultural Economics* 22: 309-320.

Bowlus, Audra and Terry Sicular, forthcoming, "Moving toward Markets? Labor Allocation in Rural China," *Journal of Development Economics*.

Chen, Shaohua and Yan Wang, 2001, "China's Growth and Poverty Reduction: Recent Trends between 1990 and 1999," mimeo, The World Bank, July.

Dollar, David and Aart Kraay, 2001, "Trade, Growth and Poverty," World Bank Policy Research Working Paper 2615.

Hare, Denise, 1994, "Rural Nonagricultural Activities and Their Impact on the Distribution of Income: Evidence from Farm Households in Southern China," *China Economic Review* 4(1): 59-82.

Hare, Denise, 1999a, "'Push' versus 'Pull' Factors in Migration Outflows and Returns: Determinants of Migration Status and Spell Duration among China's Rural Population," *Journal of Development Studies* 35(3): 45-72, February.

Hare, Denise, 1999b, "Women's Economic Status in Rural China: Household Contributions to Male-Female Disparities in the Wage Labor Market," *World Development* 27(6): 1011-29, June.

Heckman, James, 1979, "Sample Selection Bias as a Specification Error," *Econometrica* 47(1): 153-162, January.

Jacoby, Hanan, 1993, "Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra," *The Review of Economic Studies* 60(4): 903-921.

Knight, John, and Lina Song, 1997, "Chinese Peasant Choices: Migration, Rural Industry or Farming," mimeo, Oxford Applied Economics Discussion Paper No. 188, October.

Knight, John, and Lina Song, 1999, *The Rural-Urban Divide: Economic Disparities and Interactions in China*, Studies on Contemporary China, Oxford and New York: Oxford University Press.

Li, Tianyou and Junsen Zhang, 1998, "Returns to Education under Collective and Household Farming in China," *Journal of Development Economics* 56(2): 307-335, August.

Meng, Xin, 2000, *Labour Market Reform in China*, Cambridge: Cambridge University Press.

Michelson, Ethan and William Parish, 2000, "Gender Differentials in Economic Success: Rural China in 1991," in Barbara Entwisle and Gail E. Henderson, eds., *Redrawing Boundaries: Work, Households and Gender in China*, Berkeley: U.C. Press, pp.134-56.

National Bureau of Statistics Rural Social and Economic Survey Team, 2000, *Zhongguo Nongcun Zhuhu Diaocha Nianjian 2000* (China Rural Household Survey Yearbook 2000), Beijing: China Statistics Press.

National Bureau of Statistics, 1998, *Zhongguo Shehui yu Keji Fazhan Baogao* (China Social and Science and Technology Development Report), Beijing: China Statistics Press.

National Bureau of Statistics, 2001, *China Statistical Abstract 2001*, Beijing: China Statistics Press.

National Bureau of Statistics, various years, China Statistical Yearbook, Beijing: China Statistics Press.

Parish, W.L., X. Zhe and F. Li, 1995, "Nonfarm Work and Marketization of the Chinese Countryside," *China Quarterly* 143: 697-730.

Rawski, T.G. and R.W. Mead, 1998, "On the Trail of China's Phantom Farmers," *World Development* 26(5): 767-781.

Rodrik, Dani. 2000. "Comments on 'Trade, Growth, and Poverty,' by David Dollar and Aart Kraay," Mimeo: http://ksghome.harvard.edu/~.drodrik.academic.ksg/papers.html.

Rozelle, Scott, Guo Li, Minggao Shen, Amelia Hughart, John Giles, 1999, "Leaving China's Farms: Survey Results of New Paths and Remaining Hurdles to Rural Migration," *China Quarterly* 158: 367-393, June.

Skoufias, Emmanuel, 1994, "Using Shadow Wages to Estimate Labor Supply of Agricultural Households," *American Journal of Agricultural Economics* 76: 215-227, May.

UNDP, 1998, *China Human Development Report: Human Development and Poverty Alleviation*, 1997, Beijing: UNDP.

Winters, L. Alan, 2000, "Trade, Trade Policy, and Poverty: What are the Links?" Centre for Economic Policy Research Paper No. 2382.

The World Bank, 2001, *China: Overcoming Rural Poverty*, Washington, D.C.: The World Bank.

Wu, Harry X. and Li Zhou, 1996, "Research on Rural to Urban Labour Migration in the Post-Reform China: A Survey," Chinese Economy Research Unit Working Paper 96/4, University of Adelaide, February.

Yang, Dennis Tao and Hao Zhou, 1999, "Rural-Urban Disparity and Sectoral Labour Allocation in China," *Journal of Development Studies* 35(3): 105-133, February.

Yang, Dennis Tao, 1997, "Education and Off-Farm Work," *Economic Development and Cultural Change* 45(3): 613-632, April.

Yao, Yang, 1999, "Rural Industry and Labor Market Integration in Eastern China," *Journal of Development Economics* 59(2): 463-96, August.

Zhao, Yaohui, 1999a, "Labor Migration and Earnings Differences: The Case of Rural China," *Economic Development and Cultural Change* 47(4): 767-82, July.

Zhao, Yaohui, 1999b, "Leaving the Countryside: Rural-to-Urban Migration Decisions in China," *American Economic Review* 89(2): 281-286, May.

Table 1. Descriptive statistics, CHNS, 1997

	NBS rural household survey, 1997	CHNS non-urban households		
Variable		Mean	Std.	Obs
Household size	4.35	3.83	1.42	3239
Number of adults	2.79	2.67	1.29	3212
Number of dependents per capita		.30	.26	3212
Number of male adults		1.36	.80	3212
Number of female adults		1.31	.76	3212
Number of male labor (hour>0)		1.18	.73	3239
Number of female labor (hour>0)		1.10	.72	3239
Mean age of working age adults		35.39	14.30	3238
Max years of schooling for family worker		8.43	3.21	2990
Cultivated Land (mu)		6.84	12.14	1807
Household total income (yuan)		9094	9871	1966
Household net earnings (yuan)	9092	7774	9784	2160
Household net earnings from agricultural sectors (yuan)	5081	2069	5095	2380
Household net earnings from nonagricultural sectors (yuan)	1009	1267	4627	3080
Household earnings from employed work (yuan)	2240	3830	7835	3069
Household nonlabor income (yuan)		1397	2868	2808
Household total labor hours		3786	2874	2904
Household total labor hours put into household agricultural production		1803	2203	3054
Household total labor hours put into household nonagricultural production		572	1504	3170
Household total labor hours put into employed work		1359	2115	3122
Heilongjiang		.12		3239
Jiangsu		.12		3239
Shandong		.12		3239
Henan		.13		3239
Hubei		.12		3239
Hunan		.12		3239
Guangxi		.13		3239
Guizhou		.13		3239

Note: The NBS data are taken from the China Statistical Yearbook. The NBS rural survey sample includes only households in rural villages but in all provinces, while our sub-sample of CHNS data includes households in suburban villages and county towns but only in eight provinces. Note also that income is calculated differently for the NBS and CHNS samples. Also, NBS income data are per capita, and so household total income shown above is estimated as mean household size times mean income, which could create some bias because household size and income per capita are typically inversely correlated. Nevertheless, comparison of the CHNS means with the NBS means gives a rough indication of how the CHNS sample compares with the official national rural survey sample.

Table 2 Income generation functions
Dependent variable: Household net earnings

	Model I: All Households		Model II: Bottom 30%		Model III: Top 30%	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Agricultural production labor hours	0.194**	0.097	-0.110***	0.033	0.782**	0.308
Non-agricultural production labor hours	1.686***	0.135	-0.033	0.079	1.644***	0.326
Market employment labor hours	2.303***	0.111	0.369***	0.133	1.549***	0.288
Maximum education in workers	338.991***	70.282	43.250*	25.421	448.258**	198.092
Ratio of male labor to total labor	-0.779	8.735	1.068	3.149	-16.192	26.326
Land	74.918***	17.412	-24.737	16.212	49.615*	29.646
Total equipment(1000 yuan)	151.407***	16.615	-27.520*	15.578	154.165***	30.909
Mean age of family workers	41.672**	20.940	-9.714	6.970	52.507	64.501
Heilongjiang	900.866	709.122	967.179*	513.768	2757.233	4050.562
Jiangsu	1937.351***	695.930	1451.596***	483.547	4511.97	4046.863
Shandong	684.794	816.349	663.564	550.831	3579.502	4280.659
Henan	-972.873	725.083	1178.118**	478.962	4742.403	4580.596
Hubei	166.528	677.229	1458.811***	481.151	3216.005	4255.864
Hunan	2346.878***	880.881	1048.804*	556.058	4369.267	4302.067
Guangxi	-433.246	724.902	1446.324***	493.105	4162.543	4461.87
Guizhou	-	-	1589.135***	451.041	3775.855	4387.367
Constant	-2618.693**	1337.526	-	-	-	-
Adjusted R square	0.361		0.289		0.692	
Number of observations	1708		514		513	

Note: Guizhou is the omitted regional dummy variable in Model I. In Models II and III all regional dummy variables are included and the constant term suppressed to allow analysis of the regional contribution in the Oaxaca decomposition.

***, **, *: Significance at 1, 5 and 10 percentage levels.

Table 3. Descriptive statistics for variables used in the income generation functions, CHNS, 1997

	All households used in the income generation function		Households in the bottom 30%		Households in the top 30%	
Variable	Mean	Std.	Mean	Std	Mean	Std
Household net earnings (yuan)	7989	9686	897	1629	17065	12692
Labor hours in household agricultural production	1869	2247	2921	2294	957	1835
Labor hours in household nonagricultural production	567	1532	213	931	849	1963
Labor hours in market employment	1560	2085	134	548	3004	2250
Household total labor hours	3995	2495	3268	2293	4810	2548
Max years of schooling for family worker	8.61	3.16	7.16	3.15	9.93	2.94
Ratio of male labor to total labor	0.53	0.22	0.52	0.22	0.54	0.20
Equipment (1000 yuan)	2.21	11.54	0.82	4.58	4.05	17.57
Cultivated Land (mu)	4.73	12.05	4.73	5.50	4.24	18.89
Mean age of family workers	40.0	9.87	40.8	11.34	38.7	8.61
Heilongjiang	.16		.10		.20	
Jiangsu	.15		.09		.25	
Shandong	.09		.04		.13	
Henan	.11		.16		.05	
Hubei	.14		.16		.09	
Hunan	.07		.04		.11	
Guangxi	.11		.12		.09	
Guizhou	.17		.28		.08	
Number of observations	1708		514		513	

Table 4 Decomposition of household net earnings differentials between households belonging to the top and bottom 30% of the income distribution

	Difference between top and bottom 30%	Percent of total difference
Household net earnings	16168	100
Amount due to		
(1) Differences in productive characteristics	3419	21.1
(2) Differences in returns to productive characteristics	10267	63.5
(3) Differences in regional location and returns	2483	15.4
(1) Differences due to productive characteristics		
(a) Agricultural production labor hours	-659	-4.1
(b) Non-agricultural production labor hours	512	3.2
(c) Market employed labor hours	2751	17.0
(d) Education	681	4.2
(e) Ratio of male labor to total labor	-21	-0.1
(f) Land	-6	-0.03
(g) Total equipment (1000 yuan)	205	1.3
(h) Mean age	-45	-0.3
(2) Differences due to returns to productive characteristics		
(a) Agricultural production labor hours	1730	10.7
(b) Non-agricultural production labor hours	890	5.5
(c) Market employment labor hours	1852	11.5
(d) Education	3460	21.4
(e) Male/female worker ratio	-913	-5.6
(f) Land	333	2.1
(g) Total equipment(1000 yuan)	442	2.7
(h) Mean age	2473	15.3
(3) Differences in regional location and returns		
(a) Differences in regional location	-88	-0.5
(b) Differences in returns to regional location	2571	15.9

Table 5 Agricultural production function
Dependent variable: Log output value

	Model I: OLS		Model II: IV		Model III: Correct for sample selection bias	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Log labor hours[a]	0.191***	0.034	0.219	0.199	0.179***	0.035
Log variable costs[a]	0.458***	0.031	0.430**	0.207	0.446***	0.030
Log equipment	0.033***	0.009	0.034***	0.011	0.033***	0.009
Log land	0.317***	0.051	0.328***	0.118	0.317***	0.052
Maximum schooling of labor	0.021**	0.009	0.022*	0.011	0.023**	0.009
Mean age of labor	0.001	0.003	0.001	0.003	0.002	0.003
Dummy: Telephone present in local areas	0.106*	0.061	0.099	0.064	0.112*	0.061
Heilongjiang	0.092	0.114	0.108	0.174	0.087	0.113
Jiangsu	-0.088	0.095	-0.064	0.160	-0.101	0.095
Shandong	-0.176	0.138	-0.159	0.194	-0.183	0.137
Henan	-0.638***	0.100	-0.633***	0.133	-0.635***	0.101
Hubei	0.007	0.088	0.004	0.107	0.029	0.088
Hunan	0.147	0.154	0.175	0.226	0.175	0.156
Guangxi	-0.320***	0.093	-0.310***	0.099	-0.274***	0.093
Constant	2.475***	0.323	2.444**	0.982	2.632***	0.325
Lambda	-		-		-0.080	0.096
Adjusted R square	0.464		0.461		-	
Log-likelihood	-		-		-1414.124	
Number of observations	1085		1079		1118	

Note: "a" indicates endogenous variables. Instrument variables used are family composition variables (number of working age adult, number of children, and the number of elderly) and market prices for vegetables, pork, chicken and gasoline. The Hausman statistic for the joint endogeneity test is 1.39, and the P-value is 0.994.

***, **, *: Significance at 1, 5 and 10 percentage levels.

Table 6 Nonagricultural production function
Dependent variable: Log output value

	Model I: OLS		Model II: IV		Model III: Correct for sample selection bias	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Log labor hours[a]	0.626***	0.039	0.253	0.302	0.662***	0.048
Log variable costs[a]	0.168***	0.012	0.323***	0.089	0.156***	0.015
Log equipment	0.033***	0.010	0.020	0.016	-0.025	0.034
Maximum schooling of labor	0.008	0.014	0.003	0.018	-0.021	0.024
Mean age of labor	-0.009**	0.004	-0.010*	0.005	-0.002	0.006
Dummy: Telephone present in local areas	0.074	0.096	-0.084	0.173	0.097	0.112
Heolongjiang	0.049	0.205	0.206	0.260	0.361	0.349
Jiangsu	0.191	0.141	0.290	0.178	0.168	0.199
Shandong	-0.248	0.186	-0.387	0.236	-0.132	0.258
Henan	0.028	0.123	0.071	0.149	-0.071	0.156
Hubei	-0.023	0.141	-0.127	0.178	-0.128	0.213
Hunan	0.100	0.132	0.092	0.160	0.240	0.234
Guangxi	-0.122	0.111	-0.113	0.133	-0.402**	0.167
Constant	3.049***	0.343	5.177***	1.873	3.517***	0.629
Lambda	-		-		-0.408	0.256
Adjusted R square	0.610		0.453		-	
Log-likelihood	-		-		-957.878	
Number of observations	514		512		1651	

Note: "a" indicates endogenous variables. Instrument variables are: family composition variables (number of working age adult, number of children, and the number of elderly), the market prices of honey-combed coal briquet, coal lumps, coal powder and liquefied natural gas, and the local market wage rate for the unskilled. The Hausman statistic for the joint endogeneity test is 4.47, and the P-value is 0.346.

***, **, *: Significance at 1, 5 and 10 percentage levels.

Table 7 Shadow wage equations in household agricultural/non-agricultural production
Dependent variable: Log shadow wage

	Household agricultural production (OLS)		Household non-agricultural production(OLS)	
	Coefficient	Standard error	Coefficient	Standard error
Agricultural equipment(1000 yuan)	0.018***	0.007	0.036	0.025
Non-agricultural equipment(1000 yuan)	0.013***	0.005	0.017***	0.005
Land	0.012***	0.003	-0.016	0.012
Number of family working age adult	-0.042	0.032	0.046	0.058
Number of dependents	0.038	0.036	0.080	0.064
Ratio of number of family male labor to total labor	-0.0002	0.002	0.001	0.004
Family non-labor income(1000 yuan)	0.003	0.019	-0.044	0.033
Maximum schooling of labor	0.040***	0.013	0.006	0.028
Mean age of labor	0.002	0.004	-0.001	0.008
Dummy: All family members are healthy	0.157**	0.071	0.103	0.133
Dummy: Telephone present in local areas	-0.008	0.089	0.231	0.155
Market price of kerosene	0.153	0.158	0.468*	0.259
Market price of rice most commonly used	-0.010	0.030	0.116	0.121
Dummy: TVEs present in local areas	0.110	0.081	-0.088	0.143
Number of self-employed household enterprises in local areas	0.004*	0.002	-0.001	0.003
Heilongjiang	0.917***	0.152	0.010	0.435
Jiangsu	0.615***	0.175	0.362	0.319
Shandong	0.585***	0.223	0.215	0.376
Henan	-0.452***	0.132	0.091	0.221
Hubei	-0.076	0.114	-0.034	0.248
Hunan	0.459	0.325	0.132	0.459
Guangxi	-0.200	0.151	0.073	0.245
Constant	-2.228***	0.426	-1.801**	0.821
Adjusted R square	0.257		0.064	
Number of observations	1010		275	

***, **, *: Significance at 1, 5 and 10 percentage levels.

Table 8 Employed wage equation
Dependent variable: Log hourly wage

	Model I: OLS		Model II: Selectivity bias corrected[a]	
Independent variable	Coefficient	Standard error	Coefficient	Standard error
Female	-0.203***	0.029	-0.203***	0.033
Marriage status	0.012	0.046	0.012	0.058
Education level	0.030***	0.005	0.030***	0.009
Age	0.029***	0.011	0.029**	0.012
Age squared	-0.0003**	0.0001	-0.0003**	0.0001
Health status	0.066*	0.040	0.066*	0.039
Urban Hukou	0.110***	0.029	0.110***	0.029
Heilongjiang	-0.084	0.065	-0.084	0.064
Jiangsu	0.181***	0.058	0.181***	0.069
Shandong	-0.057	0.061	-0.057	0.070
Henan	-0.236***	0.072	-0.236***	0.073
Hubei	-0.043	0.064	-0.043	0.063
Hunan	0.229***	0.063	0.229***	0.065
Guangxi	-0.002	0.063	-0.002	0.062
Constant	-0.018	0.189	-0.018	-
Lambda	-		0.00005	0.084
Adjusted R square	0.116		-	
Log-likelihood	-		-5493.72	
Number of observations	1910		8123	

a. The identification variable is the number of dependents.
***, **, *: Significance at 1, 5 and 10 percentage levels.

Table 9 Summary statistics for wage rates and shadow wages
Unit: yuan/hour

Variable	Obs.	Mean	Std. Dev.
Shadow wage for agricultural production (W_A)	1079	0.623	1.110
Shadow wage for agricultural production, predicted ($\hat{W}_A$)	1525	0.427	1.847
Shadow wage for non-agricultural production (W_N)	512	1.235	1.687
Shadow wage for non-agricultural production, predicted ($\hat{W}_N$)	1525	0.908	3.212
Employed wage rate (individual level) (W_M)	1910	4.013	37.789
Employed wage rate, predicted (individual level) ($\hat{W}_{Mi}$)	8270	2.239	0.554
Employed wage rate. predicted (household level) ($\hat{W}_M$)	2755	2.271	0.526

Table 10. Household labor supply equations

	Log total hours		Log agricultural hours		Log non-agriculture self-employment hours				Log market employment hours			
	2SLS		2SLS		2SLS		Tobit unconditional		2SLS		Tobit unconditional	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error	Marginal effect	Standard error	Coefficient	Standard error	Marginal effect	Standard error
$\ln\hat{W}_A$	-0.105	0.086	0.070	0.099	0.112	0.310	-0.168	0.417	-1.290***	0.338	-2.600***	0.496
$\ln\hat{W}_N$	0.049	0.057	-0.126**	0.066	0.381**	0.205	0.500*	0.279	0.518**	0.224	1.262***	0.320
$\ln\hat{W}_M$	0.072	0.109	-0.180	0.125	0.717**	0.386	0.932**	0.461	0.553	0.427	0.465	0.377
Log (non-labor income)	-0.009*	0.005	-0.032***	0.006	-0.053***	0.019	-0.048***	0.018	0.116***	0.021	0.108***	0.020
Number of dependents	0.112***	0.020	0.123***	0.023	0.232***	0.070	0.173***	0.062	-0.259***	0.077	-0.211***	0.074
Dummy: TVEs present in local areas	0.019	0.045	-0.194***	0.052	0.183	0.160	0.221	0.152	1.064***	0.177	1.199***	0.170
Number of self-employed household enterprises in local areas	-0.001	0.001	-0.007***	0.001	0.002	0.004	0.002	0.003	0.012***	0.004	0.014***	0.004
Market price of kerosene	-0.132	0.090	-0.050	0.103	-0.023	0.312	-0.027	0.282	-0.295	0.344	-0.598**	0.338
Market price of common rice	-0.023	0.016	0.001	0.019	-0.097*	0.058	-0.132*	0.069	-0.116*	0.063	-0.191***	0.065
Maximum education level of worker	0.006	0.008	-0.039***	0.009	0.063**	0.029	0.076***	0.030	0.247***	0.032	0.287***	0.034
Ratio of male labor to total labor	-0.001	0.001	0.001	0.001	-0.001	0.004	-0.001	0.004	-0.003	0.005	-0.001	0.004
Mean age of family worker	0.003	0.002	0.004	0.003	0.002	0.008	-0.002	0.008	-0.014	0.009	-0.017**	0.009
Dummy: good health for family	-0.019	0.043	-0.126**	0.050	0.220	0.154	0.154	0.145	0.223	0.168	0.346**	0.164
Heilongjiang	-0.030	0.134	-0.234	0.155	-1.051**	0.481	-0.776***	0.297	1.715***	0.527	6.421***	1.823
Jiangsu	0.027	0.111	-0.329***	0.128	-0.389	0.393	-0.294	0.292	2.230***	0.431	3.553***	0.983
Shandong	0.152	0.114	-0.513***	0.130	-0.720*	0.392	-0.465**	0.235	2.576***	0.435	4.719***	1.149
Henan	-0.163***	0.075	-0.315***	0.087	0.474*	0.269	0.266	0.341	-0.520*	0.294	-0.904***	0.207
Hubei	0.208***	0.064	0.331***	0.074	-0.567**	0.228	-0.524***	0.149	-0.167	0.250	-0.217	0.222
Hunan	0.126	0.165	0.168	0.190	-0.258	0.582	-0.271	0.399	0.266	0.641	1.745	1.107
Guangxi	0.283***	0.086	0.372***	0.098	1.295***	0.301	1.012**	0.434	-0.394	0.331	-0.564**	0.232
Constant	6.957***	0.316	7.321***	0.364	0.394	1.122	-	-	-1.866	1.229	-	-
Adjusted R-square	0.096		0.193		0.100		-		0.186		-	
Pseudo R-square	-		-		-		0.056		-		0.082	
Observations	1360		1404		1430		1430		1421		1421	

Note: Working hours are per adult. ***, **, *: Significance at 1, 5 and 10 percentage levels.