MARKETS, HUMAN CAPITAL, AND INEQUALITY:

EVIDENCE FROM RURAL CHINA

Dwayne Benjamin
University of Toronto

Loren Brandt
University of Toronto

Paul Glewwe
University of Minnesota

Guo Li
World Bank

August, 1999
This version: March, 2000

This is a revised version of a paper prepared for the IEA World Congress panel sessions on "Global Inequality: Where Are We and Where Are We Headed," Buenos Aires, August 23-27, 1999. We thank Richard Freeman and conference participants for comments. Benjamin and Brandt thank the SSHRC for financial support.

Abstract

Market reforms are generally credited with the rapid growth enjoyed by China's rural sector. This growth has not been without some cost, however, as inequality has also increased. Estimates suggest that the Gini rose from less than 0.20 to over 0.40 during this period. In this paper we go behind these numbers to explore the nature and causes of this inequality. To begin, we find that a considerable share of rural inequality is driven by local differences in household incomes, as opposed to regional income differences, that have been the focus of the previous literature. We then examine inter-household income differentials at the village level, exploring the links between education, market development, non-agricultural employment, and household income. To address these questions, we draw on a recently collected data set from Northeast China, that was collected by two of the authors in collaboration with Chinese colleagues in Hebei and Liaoning provinces in 1995. For purposes of comparison, we also draw on the Chinese Health and Nutrition Survey. We find that indeed, increasing rates of return to education and unevenly developed non-agricultural business opportunities contribute to the high levels of inequality in the countryside. Of most interest, however, is the implication that simultaneous improvements in educational attainment and off-farm market-development would allow more households to share in the rapid growth in rural China.

JEL Classification: 015, D3, P00

I. Introduction

Beginning in the 1980s, almost all of the socialist countries replaced their planned economies with economic systems that relied heavily on market forces to determine the production and allocation of goods and services. This transformation has affected the lives of nearly 2 billion people. Historically, the two main arguments in favour of planned economies were that they are more productive in the long run (because they avoided the inherent instability of market forces, and were able to mobilize more resources for investment than a decentralized system), and that they provide a more equitable distribution of income. The experience of both socialist and market economies in the 20th century decisively rejects the first argument; it would be hard to find observers of almost any persuasion who claim that planned economies are more productive or more efficient than market economies. Yet the second argument may well be valid; planned economies may indeed be more equitable than market economies. This raises the possibility that some societies may wish to retain at least some of the policies of planned economies, despite their inefficiencies, to maintain a more equitable distribution of income. Consequently, for countries that abandoned planning in favour of the market an important policy issue is the extent to which this policy shift has increased inequality.

This paper examines inequality in China, the first socialist economy to begin the transformation from a planned to a market economy, using data collected in rural areas of Northeast China in 1995. As explained below, official statistics from China show a substantial increase in inequality in rural areas since the introduction of the market economy, yet there is very little evidence on the *causes* of this increased inequality and the *mechanisms* through which it increased. This paper uses an unusually rich data set to examine these deeper questions.

We begin by reviewing a number of conceptual issues linking the transition from planned to market-based determination of incomes to inequality. In this discussion we highlight the potentially crucial roles that human capital, such as formal education, and market institutions might play in affecting the degree of inequality associated with transition. After providing a brief review of the existing literature

on inequality in China, we turn to our own empirical explorations. Rather than trying to summarize the sources of inequality in a country as vast as China, we focus on households in a 6 counties in Liaoning and Hebei. What we sacrifice in breadth of sample is more than made-up in the detailed household-level information.

We confirm some of the findings of the existing literature, namely that non-agricultural income has been the driving force in increases in inequality, and that the educated have been most able to take advantage of economic transition. However, we also show that the existing literature has under-emphasized a number of important points. First, most of the inequality we observe is within villages, not driven by differences in income levels across locations. This is especially true of the role of non-agricultural income. Second, the role of human capital is paramount in determining the evolution of inequality, especially in the way it interacts with unevenly developed local market opportunities that characterize economic transition. Third, differences in capital accumulation (savings) across households will contribute even more to inequality in the future. Finally, on a practical note, several important data, measurement, and econometric issues associated with evaluating inequality have been ignored, which underlines the importance of survey design and data collection in future research on inequality in transition economies. Measuring the consequences of economic transition requires household surveys that can keep up with the changing nature of household economic activity. This point has been made before by Ravallion and Chen (1996), and our results confirm their insight.

II. Economic Transition and Inequality

A. Institutional Background: Transition and China's Rural Reform

The reform of China's rural sector dates from the late 1970s. The key to the reforms was implementation of the Household Responsibility System (HRS), which marked the re-introduction of family farming to China's agricultural sector after nearly twenty five years of collectivized agriculture. Under HRS, households received land-use rights and residual income rights, in return for meeting tax and

obligatory sales quotas to the state. The land-use rights were typically distributed to households on an egalitarian basis, using allocation rules based on household size and demographic composition. Ownership, however, remained vested in the village, or former collective, with a majority of villages electing to redistribute use rights to existing and newly formed households every 4-6 years.[1] By 1982, 98 percent of Chinese rural households were producing under HRS.[2]

Simultaneously, price and marketing reforms were carried out that heralded a long-run decline in government planning of procurement and sales in the farm sector, and allowed for new entry in both input and output markets.[3] By the early 1990s, procurement of agricultural products by the state commercial system had fallen to less than a third of total farm sales, down from over seventy percent in the late 1970s. Restrictions on household sideline activities in agriculture were also relaxed. Combined, these reforms increased the returns to a household's labor in the farm sector, and in the process sparked an enormous boom in agriculture growth. Between 1978-84, agricultural output grew 7.7 percent per year, or two and a half times the rate between 1952-78.[4]

According to Ho (1994), the incentive effects of HRS had an unintended consequence of making between 25-30 percent of the rural labor force redundant in farming. The policy imperative of finding ways to absorb this surplus labor, fiscal decentralization, and a relaxation of restrictions on entry into industry spurred the development and growth of local enterprise under local government initiative.[5] Between 1980 and 1995, output in township and village enterprises (TVEs) grew at an annual rate of nearly 18%, or nearly twice that in the state sector, while employment increased over the same period from 30 million to 128.6 million.[6] By 1995, output by these firms constituted a quarter of total industrial

---

[1] For an extended discussion of the allocation of land rights, see Turner, Brandt and Rozelle (1999).

[2] See Justin Lin (1994)

[3] See Sicular (1995), and Park and Rozelle (1998).

[4] Justin Lin (1992), citing Ministry of Agriculture estimates.

[5] See Byrd and Gelb (1990) and Naughton (1995).

[6] State Statistical Bureau (1998).

output. Beginning in the mid-1990s, many of these firms were privatized, and the growth of privately owned enterprises in China's rural sector accelerated under a less discriminatory regime.

Finally, long-standing restrictions on mobility and migration from rural-to-urban and rural-to-rural areas were gradually relaxed beginning in the mid-1980s.[7] Estimates differ, but one commonly cited estimate (Zhao (1999)) puts the number of migrants in China's urban areas at over 50 million. Much of this migration appears to be temporary, with a high percentage of the migrants ultimately returning to the countryside. Through their effect on wage earnings and remittances, these trends have potentially important implications for household behavior and the distribution of income.

In summary, the reforms of the 1980s and 1990s yielded two main changes in the way that households earn a living. First, the market — however imperfectly developed — replaced egalitarian-motivated, administrative "assignment" of income to households. Second, the reforms contributed to rapid increases in agricultural productivity, and the emergence of non-agricultural sources of income: economic development, in short. As we proceed, we will try to separate the impact of these two features of economic transition on income inequality, conceptually if not empirically.

B. Transition and Inequality

Even if every Chinese family got richer after the reforms, it is difficult to imagine inequality falling with a move to market-based income determination. In fact, official statistics show that income inequality in rural China increased significantly the last decade and a half. Results based on China's (annual) Rural Household Survey, show that the Gini coefficient increased from 0.21 in 1978 to 0.32 in 1994.[8] Khan and Riskin (1998) report results based on data collected in a nationally representative survey for 1988 and 1995, using standard international definitions for household income. They suggest even higher levels of inequality, but confirm the rising trend: The Gini increased from 0.34 in 1988 to 0.42 in

[7] See the collection of papers edited by Lorraine West and Yaohui Zhao (1999).

[8] Ministry of Agriculture (1994)

1995.

So inequality went up, but this tells us very little about *why* inequality rose, and leaves important questions unanswered. For example, is the increase transitory or permanent? Are the "losers" in transition likely to fall further behind as the economy grows and becomes more market-oriented? How much of the inequality is the consequence of markets *per se*, as opposed to a product of the accompanying growth and economic development? How much of the inequality is due to "market" based policies, and how much is due to remaining vestiges of socialist organization? How important are regional or sectoral aspects of the income distribution? Are there any obvious government policies that can lessen some of the diverging trends? Is the rising inequality – regional or otherwise– a threat to the political sustainability of reforms? Clearly, we cannot address all of these questions. Nevertheless, it is worth spending a short time tracing the economic avenues by which transition can increase inequality. We do this with a simple model that helps identify the key relationships that we want to explore in addressing these questions.

Imagine an economy with two villages A and B, with N_A and N_B people. The distribution of income can be described by:

$$F(y_{1A}, y_{2A}, ..., y_{N_A A}, y_{1B}, y_{2B}, ...y_{N_B B}) \equiv F(y)$$

where y_{hv} is household h's income in village v, and y denotes the vector of everyone's income. In the most general framework, household income will depend on household endowments, x, and local economic conditions, including institutions, I, that determine how endowments convert to income:

$$y_{hv} = g(x; I_v)$$

The function $g(\bullet)$ summarizes the mapping between endowments, village conditions, and household income. Perhaps the simplest form this function can take is that under a perfect markets model:

$$y_{hv} = w_v^M \, ' x$$

where endowments of villagers are valued at common village market prices, w_v^M, and households earn the

full value of their endowments. We can take this model one step further by having factor prices equal across the two villages, in which case income inequality is driven entirely by the inequality of factor endowments. In this framework we can also calculate how much of the total inequality is due to differences in endowments between the villages.[9]

We can think of economic transition as a movement from a function $g(x; I_v)$ with planned-economy (and collectivist) institutions, I_v^C, to a perfect markets-based set of institutions, I_v^M.[10] While no simple model is likely to hold during the transition period, we can imagine the changes in inequality that result from the movement from a collectivist economy with one set of prices, to a market economy with another set of prices.

For simplicity, we begin by ignoring any changes in "endowments" that result from transition.[11] In a rural economy (under collectivization), the key factors of production are: land, labour, human capital (education), physical capital, political connections, and good luck. Under collectivization, land was collectively owned, and implicitly equally distributed. This is also true under the current transition, where land use-rights are distributed to households on an essentially constant per-capita basis. Per-capita labour is similarly distributed, then as now; the distribution of human capital is less clear. Physical capital, being collectively owned, was more equally distributed under collectivization. Political connections were not equally distributed, nor was good luck (fortune).

Under collectivization, household incomes were determined at the brigade (and team) level.

---

[9] This is the "ecological" model of inequality described by Riskin (1994), whereby Chinese inequality is driven by differences in geographic endowments. The contrasting model (or emphasis) is provided by the "socioeconomic" model, described by Riskin, which emphasizes local institutions.

[10] This is a simplification: markets were not totally absent in the collectivist period, and in the market-based economy, there remain restrictions on price-setting behavior.

[11] Of course, endowments are likely to change with transition as well. Some reshuffling of endowments like land (or other entitlements) may occur as a direct consequence of government policy. As well, individuals will respond to the changed economy by acquiring factors of production, like human or physical capital. Our simplification is really a glossing over of the fact that there are few true "endowments" that correspond to economist's usual theoretical notion.

Factor prices (work points) were set so that returns were relatively constant across agricultural tasks. Similarly, household incomes were adjusted on the basis of the number of dependents ("to each according to his needs."). The returns to human capital were further depressed by restricting the ability of households to engage in non-agricultural activities, where returns to education are typically highest. A simple representation of this institutional setting is: $y_{hv} = g(x; I_v^C) = w_v^c{}'x + g'z$, where w_v^C is the vector of endowment prices under collectivization, and (is the return to non-productive characteristics, z (needs).

Thus, we expect a highly egalitarian distribution of village income, with village-level income inequality deriving only perhaps from differences in household misfortune, such as the illness of household members, or political connections. Under this institutional framework, it thus seems reasonable to view the primary source of inequality in the rural economy as deriving from across-village differences in endowments.

How might the distribution of income change with a movement to market pricing of endowments?[12] Since land is still (fairly) equitably distributed under the Household Responsibility System, that will not likely be a source of changes in inequality. Similarly, the distribution of labour remains essentially equal among households. The main changes will occur in returns to human capital, returns to physical capital, and the return to household fortune. While they do not belong directly in a market economy, it is also likely that political connections will help determine relative success.

There are at least a couple of dimensions in which the returns to human capital might increase. First, under the HRS, better farmers can retain their surpluses, and so the implicit return to human capital in farming will rise. If this human capital is unequally distributed, we expect an increase in overall inequality. Of course, given the constraints on farm size, there are limits to how much inequality can be generated by differences in farming ability. Much more important, with reform restrictions on household

[12] For an excellent description of the impact of transition on "social differentiation" (inequality), applied in a comparative perspective to China and Vietnam, see Van Luong and Unger (1998).

engagement in non-agricultural activities were gradually relaxed. It is here where we expect the returns to skill to matter most. Skilled individuals can set up their own businesses, or work for others in non-agricultural pursuits, such as TVE's. To the extent that these opportunities are unequally distributed across households, because of inequality of human capital, connections, or good fortune, then we expect to see an increase in village inequality. With fixed distributions of endowments, this inequality will be reflected in changes in the factor prices. Applied to the entire rural economy, we expect to see an increase in inequality both between and within villages. Some villages, because of advantageous locations, might see greater non-agricultural development.

Of course, transition concerns the entire process of change, not the instantaneous movement to a new set of institutions. One prediction of a more gradual movement to a market-based, less agricultural economy is the emergence of a sector-based "Kuznets curve." Overall inequality will be higher, as some regions' average incomes rise through the taking advantage of the new non-agricultural opportunities. This might also be true within-villages, as some households move into these more lucrative opportunities sooner than others. Only after a period of time, through the effect of migration, trade, or other factor mobility, will income differences between regions and across households be diminished.

The possibility of spatially uneven rural development, especially driven by the growth of township and village enterprises is prominent in the literature.[13] Through the early 1990s, TVEs were the most rapidly growing segment of the industrial sector, and by 1997 absorbed more than 125 million workers.[14] Conventional wisdom, represented for example by Rozelle (1994), suggests that the growing concentration of these enterprises in richer parts of China led to widening inter-regional differences. Fiscal decentralization, restrictions on inter-regional factor mobility and protectionism tended to magnify these gaps.

---

[13] See Hare (1994), Khan and Riskin (1998), Guojia Jiwei (1995), Knight and Song (1993) and Knight and Li (1997).

[14] Statistical Yearbook of China (1998)

There are a number of shortcomings to this work. First, some of the analysis, such as Howes and Hussain (1994), and Rozelle (1994), is carried out using per capita welfare measures at the aggregate level, (e.g. the village, township, county, or provincial level). This has two effects. On the one hand, it effectively underestimates inequality because it ignores any differences arising from differences among households within these administrative units. In other words, it focuses solely on inter-local differences. At the same time, conclusions drawn from these data about the contribution of various factors, e.g. TVEs, to overall inequality are valid only insofar as most of the inequality is arising from differences in mean incomes between these units.

Still, the exercise can be informative about trends in inequality. We conduct this exercise ourselves in Figure 1. Here we calculate the country-wide (rural) Gini coefficient, assuming that everyone in a province has the same income. This eliminates within-province inequality by construction, so that all inequality is due to regional income differences. The figure shows similar trends for inequality of both income and consumption. The Gini rises from around 0.10 in 1980 (at the beginning of the reforms) to 0.15 in 1997. This shows that Chinese inequality rose by 50 percent because of widening regional disparities. However, if we take other measures of total rural inequality reported earlier as benchmarks, then at most one half (0.10/0.21 and 0.15/0.32) of national inequality can be attributed to income differences between provinces. This fraction is even smaller if we use benchmarks based on most other recent household surveys, where the current Gini's are over 0.40. Thus, at least half, and probably more, of current inequality is due to income differences among "neighbouring" households.

Our back-of-the-envelope exercise yields similar results to more formal studies of spatial inequality. For example, Tsui (1991) confirms our estimates of the broad trends in provincial income inequality for the earliest part of our sample. Regarding the contribution of cross-province inequality to the total, Hussain, Lanjouw, and Stern (1994) in a sample drawn from 10 provinces, find that only 15 percent of inequality of household income was due to provincial income differences. Cheng (1996) employs data from 5 provinces, and finds under 40 percent of the inequality due to cross-province income

differences. Clearly, one expects that such results will be sensitive to the provinces examined, and the degree of measurement error in the household data. We pursue these question later, with our own empirical explorations. The main point we make here is that uneven spatial economic development, due to TVE's or other types of industrialization, only yields a half-story, at best, about links between transition and inequality. We must look at individual provinces and villages in order to appreciate more fully the impact of institutional change.

Other studies, e.g. Hare (1994), Kahn and Riskin (1998), used household-level data to explore the possible impact of uneven rural development on income inequality. These papers focus on the changing composition of rural household income (possibly due to TVE's). While these studies do not separate the spatial and non-spatial dimensions, they verify the important contribution of income from non-agricultural sources to inequality. But, with the exception of Hare (1994), no attention is given to *why* income from alternative sources may differ at the household level. This is extremely important if certain kinds of income are more disequalizing that others, and for the design of policies that may help reduce inequality. Also, interpretation is difficult when income from alternative income sources is lumped together, e.g. income from off-farm wages is combined with that from family-run businesses. Insofar as these sources of income are less than perfectly correlated with each other, grouping them together may hide important aspects of emerging inequality and their links to household attributes and the external economic environment with which these households interact.

While it is simple to think of transition as a quick movement from collective to market-based factor prices, it is also misleading. Markets are likely to be imperfect for some time, so the function mapping endowments to incomes is likely much messier than suggested by the market clearing model. Instead of valuing household endowments at market factor prices, they will be valued at shadow prices:

$$y_{hv} = \overline{w}_{hv}' x$$

where $\overline{w}_{hv}$ is the vector of shadow factor prices. Shadow prices can depend on both village and

individual household characteristics, and may be written as:

$$\overline{w}_{hv} = h(x, I^T, w, e)$$

depending on the distribution of endowments, local institutions, market prices, and other factors. Some aspects of transition might make inequality worse than the case in which the movement to a new market clearing outcome was instantaneous. Shadow prices might also be unequally distributed in a village, so that those with more endowments have their favourable position compounded by advantageous prices. This will especially be the case if there is quantity rationing in the various factor markets, so that those households with poor endowments cannot get "full value" for their limited endowments of labour and land.

The most interesting possibility concerns the role of human capital. Not only might education be the key to accessing scarce non-agricultural opportunities, but it may play an independent role in helping households cope with dramatic economic change. Economic transition is almost certainly a process where information about income earning opportunities, and how best to take advantage of them, is highly imperfect. The more educated, either directly because of their education, or simply as a reflection of innate ability, may thrive in such an environment. This possibility was raised, most notably by Schultz (1975), but more recently by Rosenzweig (1995) in the context of development, and Orazem and Vodopivec (1995) in the context of transition. Indeed, given the nature of transition in China, the distinction between the pure effects of institutional transition and economic development are fuzzy, at best. Combined, we expect the distribution of human capital to interact with the extent of new economic opportunities in complicated ways that imply high returns to education, and increased inequality. On a related note, political connections can be viewed as a special kind of human capital, especially if there is political rationing of the economic opportunities themselves, or of the information about them.

At this point, we know very little about how returns to human capital in China's rural sector are evolving. Several recent studies, however, provide some preliminary estimates. First, in contrast to the

pre-reform period, the return to education in agriculture during reform is positive. Li and Zhang (1998), using household level data for Sichuan for 1990, find a return of about 3 percent to average years of schooling of family members in agriculture. Using the educational level of the household head, they find a similar return, with the return to education of those individuals schooled during the Cultural Revolution slightly lower. They do not find a significant return to schooling in the pre-reform period, consistent with the highly redistributive tendencies of the collective. Dennis Yang (1997) finds a similar return to education during the reform period, but argues that most of it is coming through its effect on the household's managerial capabilities as opposed to any effect it may have on household members' ability to perform routine tasks in agriculture. He finds a slightly lower return to schooling in off-farm wages.[15] More recently, An and Yang (1998) analyzed the returns to schooling in total household operations and that to agriculture and non-agriculture activity separately. They find a much higher return to total household activity, and suggest that as much as half of the effect of human capital on earnings is coming through its effect on the allocative decisions households are making across sectors. While these results are suggestive, no linkages to inequality are explored.

Finally, the role of physical capital is likely to be very different in a transition economy than in either a purely collective or purely market based one. Under collectivization, household accumulation of capital was severely restricted. With economic reforms, successful households can accumulate physical capital. With perfect capital markets, this will generate inequality on its own, to the extent that economic success is unequally distributed. However, if capital markets are imperfect, then the effect on inequality will be more severe. Unequal capital holdings might interact with unequal (or limited) access to capital, further compounding income inequality, especially to the extent that capital is a necessary input in the operation of non-agricultural (or even agricultural) businesses. We expect, therefore, the accumulation of

[15] The low return in the wage sector may reflect differences in the wage determination between workers who are assigned to jobs (non-market group) versus those who find jobs themselves (market group). Gregory and Min (1995) find a significant return to schooling for the latter group, but no return for the non-market job-holders. They are not, however, able to explain the assignment of workers to these sectors.

capital to combine with the distribution of other factors (especially education) to generate even more inequality of income from non-agricultural sources, and income inequality over all.

In summary, we can think of transition as the movement from administered to market-determined prices of factor endowments. In this way we can speculate on the likely consequences of transition on incomes and inequality. However, transition is more complicated than that. Imperfect development of factor markets is likely to combine with the changing distribution of endowments to worsen inequality.

III. The Data

We use two sources of household level data. Our main data set is the 1995 North and Northeast China Living Standards Survey (NNCLSS). We also use the China Health and Nutrition Survey (CHNS) for comparison.

A. The NNCLSS: Sampling Frame

This survey was carried out by the authors and Chinese colleagues in 1995 and extends to 780 households in 6 counties, 18 townships, and 30 villages in Hebei and Liaoning provinces (North-Northeast China). The survey provides detailed household level information on income, expenditure (disaggregated for farm and non-farm activity), labour supply and farm management. The basic structure of the survey was based on the World Bank's Living Standards Measurement Survey, described in more detail by Glewwe and Grosh (1998).

The six counties were not selected randomly, but chosen to correspond to the site of an intensive household-level investigation carried out by Japanese investigators in 1936 and 1937. Five villages in each county were selected, one of which had been fully enumerated in the 1930s. The other four villages in the county, including one from the same township as the administrative capital of the survey; one located in the same township as the village surveyed in the 1930s; and two drawn from a third township, were selected to try to obtain as representative cross-section in each county as possible. 130 households were surveyed in each county: Fifty from the village surveyed in the 1930s, and twenty from each of the

remaining four villages. Households were chosen randomly using the most recent village registry.

B. The CHNS: Sampling Frame

We also use data collected for rural households as part of the China Health and Nutrition Survey. While this survey is longitudinal, we only make use of the single cross-section for 1993 (from the survey carried out in 1994). A detailed description of these data (as well as the data itself) can be obtained at the website.[16] The CHNS data include the provinces of Lioaning, Shandong, Jiangsu, Henan, Hubei, Hunan, Guangxi, and Guizhou, and provides a more nationally representative survey than our 1995 data. Missing, however, are provinces from China's southeastern coast (Zhejiang, Fujian, and Guangdong). In each of the eight provinces, three rural villages in four counties, for a total of 96 villages were sampled. Total sample size (of rural households) is around 1800, with approximately 20 households sampled per village.

In Figure 2 we present a map that shows the locations of the provinces sampled in the two surveys. Only Liaoning is sampled in both surveys. On the same map, we show the per capita rural income levels in 1994. Most of the provinces have income levels in the middle two categories: 900-100; and 1100-1400 Yuan per year. Liaoning is in the richest category. So, this sample will likely understate the degree of heterogeneity in China as a whole, as it excludes both the very poorest, as well as the richest provinces. That said, it is difficult to imagine coming to any conclusions, statistical or anecdotal, that could generalize across all of China. Our conclusions will never be intended as China-wide statements. That said, each of the sample provinces has a population greater than most developing countries!

IV. Empirical Results

A. Descriptive Statistics

Table 1 provides descriptive statistics for the various sorts of income earned by households in our 1995 sample. For each source of income, we report unconditional means, including in our calculations

[16] http://www.cpc.unc.edu/projects/china/china_home.html

any zeroes and negative values as well as the proportion of households with non-missing observations on each source of income. As a measure of income dispersion, we show the Gini coefficient, calculated only over the observations with positive values. The numbers are calculated using all 778 households with complete income and expenditure information, including several possibly "big" outliers.

Average household, and household per capita, income in our sample are 13,071 and 3,510 RMB, respectively. These estimates are higher than those reported on the basis of the SSB's Rural Household Survey. Only 10% of this difference can be attributed to the inclusion of the imputed value of capital services. Other reasons for these differences are probably driven by differences in survey methodology. Our LSMS style survey is more comprehensive in its enumeration of income sources, and includes several built-in consistency checks with which to "double-check" estimated income. Turning to consumption, average household expenditure is considerably lower than income, but still at a level that suggests that previous estimates of living standards have understated household incomes. At 1994 exchange rates, our estimate implies a per capita income of US$ 400. Purchasing power parity calculations suggest per capita income three times higher.[17] The difference between income and expenditure also implies significant household savings, a point to which we return later.

Farming is the most important economic activity for these households: it accounts for nearly half of total income, and all but two percent of the households have positive farm income. In turn, nearly three-quarters of farm income comes from cropping income (mostly corn in Hebei, and rice and corn in Liaoning), with the remainder from sidelines such as vegetable gardens and greenhouses. Next important to farming, family-run businesses are the source of slightly more than twenty percent of total household income.[18] Nearly a third of all households, or 247 households, are involved in a total of 274 non-agriculture businesses of some kind, with ten percent of the households involved in more than one

[17] On this point, see Lardy (1994).

[18] Income from family-run businesses represents both profits plus the implicit returns to family labor. We do not try to break up these two components.

enterprise. A third of the enterprises are in commerce, with fifteen percent in construction and transportation. On average, 2.2 individuals work in these enterprises, with the average enterprise having been in operation for almost five and a half years. Wage income is the source of about 15% of total income, with three-quarters of that coming from male wages. Forty percent of the households report income from this source.[19] The remainder of income comes from animal husbandry and an "other" category, which includes such things as interest income, remittances, pension income, etc.

The Gini coefficient for household income for the entire sample is .40, and marginally lower for per-capita household income. Our estimate is similar to that obtained by Khan and Riskin (1998) in a nationally representative survey that used comparable measures of income. Among the sources of income, income from wages and farm income are the most equitably distributed. The latter is largely a product of a fairly egalitarian distribution of land at the village level. The most unequally distributed items, on the other hand, are income from family-run business, animal husbandry, and our "other" category. This suggests that we look to these items in our analysis of the factors underlying the increase in inequality.

Finally, household expenditure, a potentially superior estimate of long-run living standards, is considerably more equally distributed than income, with a Gini for household expenditure of 0.28. In Figure 3, we graph the kernel densities for these two measures, which reveals a much fatter upper tail for income. While some of this difference may be the usual product of measure error in income relative to consumption, it may also reflect genuine differences between income and consumption, and potentially important savings behaviour.

B. Spatial Dimensions of Inequality

Following on the issues raised in our discussion of Figure 1, we now explore how much of the

[19] Slightly more than a third (36.2 percent) of these jobs were in the villages in which these households lived; 43.1 percent were in either the county or township seat; and the remainder (20.7 percent) were in the cities. Most of these jobs households found on their own and were not "allocated" or "rationed".

inequality in our sample is at the local level, as opposed to differences in average incomes across regions. For the pre-reform era, it is generally believed that most of the inequality was eliminated within localities, but not across regions.[20] With reforms, we might expect inequality to increase within villages, as collectivist institutions eroded. However, we might also expect inequality to increase across regions as well, as some areas are in a better position to take advantage of non-agricultural and other opportunities arising as a product of market liberalization and reform. In the exercises below, we decompose inequality into that part due to differences between and within localities.

Table 2 reports decomposition estimates for a variety of measures of income and expenditure inequality, using a number of alternative definitions of "location" (village, county and province). Expenditure data are not provided for the CHNS data, and so we report results for income only for that sample.[21] We take as our benchmark the Gini coefficient for income or expenditure. We also show the variance of logs and the Theil index. The decompositions are most easily done and explained using the variance of logs: the R-squared from a regression of log income on location dummies provides an estimate of the percentage of the variation in log per capita income explained by location. The remainder represents the percentage arising from within-location variation. We also report the decomposition for the Theil index, which has a similar interpretation, but is generally a preferred inequality index.

The results consistently show that most of the inequality is within, not between, whichever inequality index is chosen. This means that while differences in such factors as the level of rural industrial development, soil quality, nearness to urban centers, etc., affect mean incomes across localities, most of the inequality still comes from differences among households *within* these locations. For the NNCLSS, 25 to 30 percent of the variation in both income and expenditure comes from between-village differences, with even lower percentages arising from differences at the county or provincial level. The

[20] See Putterman (1993) and Vermeer (1982).

[21] The public-use version of the CHNS data set for 1993 does not provide an estimate for household income. We constructed a measure using the definitions of alternative income sources provided for the 1989 and 1991 data.

CHNS data suggest a similar spatial pattern, despite the fact that the CHNS sampling frame was chosen to maximize differences in economic development across villages. The basic insight to be gained from these decompositions is that understanding inequality in rural China requires understanding income determination within villages.

C. Inequality and Sources of Income

Which income sources generate the most inequality? We decompose income inequality by source using a straightforward method employed by Shorrocks (1982, and 1983). With K sources of income, total income for household i is given by $y_i = \sum_{k=1}^{K} y_{ik}$. The vector of economy wide household incomes is given by $Y = Y_1 + Y_2 + ... + Y_K$, where Y and Y_k are vectors of household income of source k for the whole economy. Average income is given by: $\overline{Y} = \overline{Y}_1 + \overline{Y}_2 + ... + \overline{Y}_K$. An increase in the mean income of source k by 1 penny leads to an increase in the mean of total income by one penny. A first order estimate of the proportional increase in mean total income is given by:

$$W_k = \frac{m_k}{m}$$

where W_k is the share of income source k in total income (evaluated at the mean). So, a 1 percent increase in income source k will increase average income by W_k percent: The more important the income source is, the larger the increase in mean income will be.

To what extent does the same logic apply to measures of inequality? Does an increase in the dispersion of a Y_k lead to an increases of inequality in direct proportion to its income share, W_k? Stated differently, under what conditions can we write:

$$I(Y) = \sum_{k=1}^{K} W_k I(Y_k)$$

This simple decomposition can be done only if each source of income is perfectly correlated with total income, i.e., if everyone is a "clone" in terms of the structure of income, except in total income.[22] In such a case, an increase in the inequality of income source *k* will increase total income inequality in direct proportion to its share of total income.

However, to the extent that a particular type of income is earned by the rich, then an increase in inequality of this source of income is going to benefit the rich disproportionately, and so increase inequality in excess of its share of total income. On the other hand, an increase in inequality of income correlated with being poor will disproportionately benefit the poor, and so decrease inequality. Both the correlation of the source of income with total income, and its relative size will matter in any decomposition.

Shorrocks (1982) shows that under reasonable conditions, a decomposition of total inequality can be calculated, so that the fraction of total inequality deriving from Y_k (for any inequality index) is:

$$S_k = \frac{\text{cov}(y_k, y)}{\text{var}(y)}$$

How can we interpret S_k? First, we can compare S_k to zero. If S_k is negative, increases in the inequality of income source k will actually reduce inequality because of the income source earned by (and distributed to) the poor. Second, we can compare S_k to W_k. Since the rich tend to earn more income from all sources, increases in inequality of any income type increases overall income inequality. However, some sources may be relatively less disequalizing, and W_k is a useful benchmark. If S_k is greater than W_k, increases in the inequality of the distribution of Y_k can be viewed as disproportionately increasing income inequality.

[22] See Shorrocks (1983).

Estimation Issues

The parameter of interest is S_k, which can be estimated as the coefficient from a regression of y_{ik} on y_i:

$$y_{ik} = a_k + b_k y_i + e_{ik}$$

The error term is worth noticing. For starters, we only have a sample (not the population) and must recognize that there is sampling error. More seriously, the error term may contain less benign components. Measurement error is the most significant possibility. Assume that each source of income is measured as follows:

$$y_{ik}^{*} = y_{ik} + v_{ik}$$

where y_{ik}^{*} is the measured income, and y_{ik} (as before) is the true value. Then, because measured y_i is the sum of the various y_{ik}, there is a possibility of correlation between y_i and e_{ik}. This generates two possible biases. First, the plim of $\hat{b}_k$ may exceed b_k, because of the positive correlation between the common measurement error in components in y_{ik} and y_i. For example, if a household overstates its non-farm income, this will spill over to the households total income. We will then overestimate the true association between non-farm income and total income. Second, we may instead get the usual attenuation bias of measurement error, in which case we will understate b_k.

This potential statistical problem can be addressed most easily by instrumental variable (IV) techniques. In our case we use per capita expenditure as an instrument for per capita income, because it is (mostly) measured independently of income.[23] Even if the instrument is imperfect, this exercise allows

[23] We exclude the value of capital services from our instrument to minimize any mechanical overlap in the measurement of income and expenditure.

us to explore the possible sensitivity of our conclusions to measurement error. Intuitively, we have two separate indicators for a household's position in the income distribution, and thus an alternative way to estimate the correlation between the various income sources and total income.

For the NNCLSS, we calculate the decompositions over two samples: the full sample, and one with the richest household dropped (this household had income over 200,000). This household is rich because of high non-farm business income, so the decompositions are sensitive to inclusion of this household. The results are reported in Table 3. For each source of income, we report three columns of numbers: 1) The share of the source of income in total income (the benchmark W_k); 2) The raw correlation coefficient (i.e., the OLS coefficient from the regression of income of type k on total income); and, 3) The 2SLS estimate of this coefficient (i.e., the measurement error corrected correlation).

Looking at the full sample (the first three columns), most income sources contribute to inequality in approximate proportion to their share of income. Women's wages have a small negative (equalizing) contribution, while men's wages contribute less to inequality than their share. Crop income contributes much less than its share to inequality. Restrictions on land ownership, and the administrative allocation of land effectively place an upper bound on how much a household can earn from farming. In contrast, animal husbandry contributes more than its share, and non-farm family businesses contribute the most to inequality. These results suggest that it is income from these enterprises, not wages from TVE's, that is generating most of the inequality. The OLS and 2SLS coefficients do not differ significantly; however, the results are sensitive to the exclusion of the richest household. Dropping this household reduces the contribution of non-farm business income, but it does not change the overall story: In order to understand rural inequality we must understand non-agricultural income.

In the bottom panel of Table 3, we show comparable results using the CHNS. Even though the sampling frame is different, we find similar patterns to the NNCLSS. First, the income shares of each type are similar, even though the CHNS has some built-in gaps (such as the value of capital services). Crop income is the largest source of income, but contributes significantly less than its share to inequality.

On the other hand, both animal husbandry and non-agricultural businesses contribute significantly more to inequality than their shares suggest. Our NNCLSS results do not seem to be an artifact of the sample.

In Table 4, we combine the previous two exercises, and explore how the contribution to inequality by income source varies over space. Our questions are: 1) How much of the contribution of non-agricultural income is driven by spatial differences in non-agricultural development; and 2) Does the contribution of non-agricultural income to inequality itself vary across villages? In columns 2 and 3, we show the coefficients that correspond to those reported in Table 3, except we include village fixed effects. Here, we see that adding village fixed effects does not change the conclusions drawn from Table 3: differential incomes from non-agricultural sources contribute significantly more to within-village inequality than is expected on the basis of their share in income. F-tests for village interaction terms show that the contribution of non-agricultural income to village inequality does vary across villages – perhaps due to differences in local institutions. However, when the village interactions are estimated by 2SLS, we do not find statistically significant differences across villages. Perhaps the OLS result is due to the higher fraction of measurement error at the village level, or alternatively, our results reflect the imprecision with which the instruments can predict differences in income between villagers.

D. Household Income and Consumption

We begin with an overview in Table 5 of the relationship between household endowments and total income and consumption. In the next section we look in more detail at how these productive characteristics, like land and human capital, affect the composition of income.

We specify the land variables as a combination of both the amount of land that the household has been allocated, and a dummy for whether or not the household has any land. It appears that land allocation is partially endogenous to the type of income earned. Rich households with sizeable non-agricultural enterprises do not receive allocated land, or elect to return it, so the coefficient on whether a household has land is negative. However, once we account for this heterogeneity, it appears that an

increase in land holdings increases per capita income. Unless land were distributed disproportionately to the poor, this implies that an increase in land inequality leads to higher inequality (conditional on all other sources of income remaining unchanged). While not an endowment variable *per se*, if a household suffers an adverse production shock (measured in terms of the reduction in output from a normal year), its income falls significantly, but expenditures do not. One interpretation of this pattern of coefficients is that households are able to smooth their consumption relatively more than their income in response to bad luck. This suggests that the large gap between income and consumption may be related to savings (as opposed to only measurement error), and that these savings are used, at least partially, to smooth consumption.

Turning to the human capital variables, the coefficient on years of schooling implies a rate of return to schooling of around 6.6% for income, and 3.8% for expenditure.[24] Unless there is a complicated covariance of the measurement error in income and consumption with education, the difference between the income and expenditure coefficients may reflect a correlation between savings rates and education. Technical training also has a significant rate of return, for both income and consumption, but apprentice training is insignificant. In order to see whether the rate of return to schooling varies across villages, we estimated an ANOVA, controlling for the same set of variables as above. An F-test for the interaction terms between education and the village dummies is significant at the 5% level. Later on, we explore in more detail how these differences across villages in the rate of return to human capital affect income inequality.

E. Household Earnings Activities

In Table 6 we explore the connections between education and the composition of income. Specifically, we report estimates of Probit equations for whether a household engages in animal

[24] These estimates are well in line with those reported in Psacharopolous (1985).

husbandry, male wage labor, female wage labour, or non-agricultural farm businesses. The covariates are the same as in Table 5, including village fixed effects. The most important result is that education plays only a small role in determining *what* activities that households engage in. Those households raising animals tend to be slightly less educated, while those working off the farm for wages, especially wage labour are slightly more educated, with average household education being highly related to female participation in wage labour. Most interestingly, schooling does not seem to matter in determining whether a household is operating non-agricultural business. Having family members with either technical or apprentice training, on the other hand, increases the likelihood of having male members working off the farm, or the family engaged in non-agricultural businesses.

Table 7 shows the estimated "returns to schooling" for the various types of income. These regressions are estimated only over those households with positive earnings, so there may be selectivity biases on the various coefficients. Fixing this bias is impossible in the context of this paper. However, we can speculate about its likely consequences. Regarding the education coefficient, if there is a positive rate of return to schooling for a given source of income, and households are positively selected into an activity, then the OLS coefficients will be biased downwards (towards zero). These biases might not be large, given the absence of strong education effects on most activity in Table 6, but we must note that the selection is not just into a specific activity, but also on whether a household has positive earnings (as opposed to losses) in that activity, which might be more strongly selected. For crop income and animal husbandry, we find only very small, but positive rates of return. Our estimates are consistent with those obtained in the previous literature. However, we estimate high rates of return in wage labour and non-agricultural income. Moreover, aggregating all types of non-crop income, we estimate a rate of return of about 10%. This suggests that households have a rate of return to schooling, not just within each activity, but also in the construction of a portfolio of income types. The slightly higher return over all non-crop income might also reflect the fact that there is less selection bias in this equation, since fewer households are excluded. In summary, we find high rates of return to schooling in non-agricultural pursuits, though

entrance into these activities is less related to education.

F. Savings and Capital Accumulation

In most empirical studies of inequality, measurement considerations dominate the choice between income and consumption (expenditures) as the measure of living standards. On these grounds, we usually prefer expenditures, since they tend to be better measured in developing countries. Furthermore, long run living standards are probably best reflected in consumption levels, as even accurately measured income may contain a significant transitory component. Our intention was to follow conventional practice, and draw most of our conclusions from the expenditure results. However, the divergence between income and consumption may reflect more than measurement error and transitory income shocks. As this is an economy in transition, households may not have settled into long run consumption patterns. In particular, high income households may be re-investing (saving) a considerable portion of their income, possibly in response to poorly developed external credit markets. In this case, we understate the differences in long-run living standards between households if we focus on consumption alone. Current income may be disproportionately related to future income (and consumption) through current savings. Essentially, savings is a mechanism by which "luck" and economic success will be compounded in a transition economy, even more than that implied by equilibrium models, suggested (for example) by Deaton and Paxson (1994) for Taiwan. We explore this possibility in detail.

As we previously pointed out, there is a considerable gap between household income and expenditures. For any individual household, this is not surprising given the degree of measurement error in household surveys. While it is tempting to interpret the difference between income and consumption as savings, the difference ("savings") may be mostly noise. More surprising, however, is the extraordinary level of implied savings across the whole sample. Evaluated at the median, household income (excluding capital services) is 9,352 versus 6,449 in expenditures, implying household savings of 2,903– a savings rate of 31 percent! Usually, reporting-error leads to higher levels of expenditure than income (which is

under-reported), and negative "savings." We thus consider the possibility that some of this gap is genuine savings.

Our objective in this section is quite limited. We wish to evaluate whether, in fact, there appears to be savings "signal" in the difference between income and consumption, and thus that savings rates may be genuinely high. We already saw indirect evidence of savings in Table 5, where we noted the differential response of income and consumption to an agricultural production shock. This was consistent with savings being used for consumption smoothing (self-insurance). With imperfect markets and diminished redistributive institutions, and presumably riskier income streams, it is plausible that households now engage in more precautionary savings.

Our main cross-check is to evaluate the relationship between contemporaneous savings and the estimated stock of household wealth, beginning with some "back-of-the envelope" calculations. We construct estimates of the total investment by households in the accumulation of physical and financial assets. Financial assets consist of deposits in financial institutions, cash on hand, and net borrowing through formal and informal lending. We have year-end figures for each of these. We also have detailed data on durable goods expenditure, both that for household use, e.g. housing stock, and that used in agriculture and non-agriculture business activities. For these assets, we know the year purchased and the price paid. Almost all of this asset accumulation occurred in the 1980s and 1990s. In principal, total investment in asset accumulation should equal the sum of household savings over this period. These totals should not be confused with (but are related to) estimated net wealth, the current value of these assets. The difference between the two is depreciation and the effect of inflation on the value of these assets. Average total accumulated investment in productive and non-productive assets equals 25,856 yuan, and the median is 14,500, or almost 1.5 times estimated median income. Out of this total investment, forty percent represents the fixed investment in agriculture and non-agriculture activity. Given the significant increases in nominal incomes between 1980 and 1995, it takes savings levels on the order of 30% in order to generate the estimated level of wealth! Given the rapid development of family

enterprises and imperfect or nonexistent capital markets, such levels of savings might be necessary.

A more formal exploration is presented in Table 8. Define the "rate of savings," r_S, as:

$$\ln\left(\frac{\text{income}}{\text{expenditures}}\right) = \ln(pcy) - \ln(pcx) \equiv r_S$$

In column (I), we show the results of regressing r_S on the same set of explanatory variables as Tables 7. First note that the savings rate is positively related to the level of schooling. Perhaps more educated households have lower discount rates; perhaps they have better investment opportunities; or possibly they are better decision makers. Second, we confirm that adverse production shocks reduce savings.

In columns II and III, we link assets to income. While we expect the two variables are positively related, the more interesting question is whether the elasticity is greater than one. If so, it suggests that wealth increases more than proportionately with income, i.e., that wealth is highly "disequalizing" or associated with high incomes. This suggests that not only is wealth unequally distributed (its Gini coefficient is 0.58), but that high income households have a disproportionate share of wealth. In column II we report OLS estimates, while in column III we report 2SLS estimates, where we try to correct for measurement in ln (pcy) using ln (pcx) as an instrument (like we did in tables 3 and 4). The OLS and 2SLS results are sufficiently different to lead to suspicion of measurement error bias. Focussing on the 2SLS results, the estimated elasticity is 1.5 and significantly higher than 1.0, suggesting that if these assets yield future income, it will accrue to higher income earners, and further increase inequality.

In columns IV-VI we try to see whether there is any significant relationship between the "savings rate," and accumulated wealth. In column (IV), we estimate the "reduced form", observing that assets are highly related to the same variables as the "savings rate," most notably education and the production shock. In the next two columns, we estimate the "structural" asset equation by OLS and 2SLS, instrumenting the savings rate for measurement error (since we know that the difference between income and expenditure will be mostly noise). We use the indicator of the degree of production shock, farm size,

and whether there are young children or teenagers in the household as instruments for the savings rate. We thus assume that these instruments affect assets only through the savings rate. We use overidentification tests to confirm the internal consistency of the instruments, and also show that empirically there is no evidence of a correlation between the identifying instruments and the error term of the asset equation. The results in Column VI suggest that the rate of "savings" is significantly, positively related to the level of accumulated assets, with an elasticity of almost one, which suggests a proportional relationship between the savings rate and the level of wealth.

All of this evidence suggests that there is genuinely high savings by these households. This fact needs to be taken into account in any picture of inequality in the transition to a market economy. Given the significant rate of savings, especially by high income households, we might expect the level of inequality to be even higher in the future. Our results also suggest that the role of imperfectly developed capital markets may be greater than previously believed, not just for economic efficiency, but also for income inequality.

G. Overview of Village-level Inequality

In this final section, we summarize some of the key determinants of inequality within villages, focussing on possible inter-linkages between the distribution of human capital and market development. While we use regressions to conduct this summary exercise, we are especially cautious in interpreting the coefficients causally: almost certainly these village characteristics evolved simultaneously. Instead, the regression coefficients should be treated as slightly sophisticated tabulations. The dependent variable is the Gini coefficient of village inequality, where the village Gini's range from under 0.2 to over 0.5, with most lying on either side of 0.3. Our focus is on how the distribution of education interacts with market development. Before turning to the results, it is first worth asking in what dimensions educational attainment varies across the households within the sample.

Educational attainment increased significantly with the expansion in the basic education system

through much of the first three decades of the PRC.[25] Both literacy and average years of schooling increased, while differences between males and females narrowed considerably with each successive cohort.[26] There is considerable debate over the impact of the restructuring of the education system in the post-Mao era and the post-1978 economic reforms on educational attainment in rural China. On the one hand, the re-introduction of family farming with the HRS and the rapid development of off-farm opportunities increased the returns to labour, and thus the opportunity cost of families keeping their children in school. Decentralizing fiscal reforms, on the other hand, shifted much of the responsibility for primary school funding to rural communities. This gave rise to differences in funding levels and possibly school quality. In poorer areas with weaker tax bases, high school fees reportedly discouraged enrollment.

Some observers suggest that these developments contributed to growing differences across communities and regions in average levels of educational attainment. Possibly counteracting these effects are the potential effects of increased household income and an increase in the returns to schooling on the demand for education.[27]

We cannot address these issues here, but are interested in the potential effect of differences in educational attainment on inequality. First, differences in schooling levels across villages can be an important source of inter-village differences in income, especially in an environment in which the returns to education are positive. There are also differences in educational attainment levels across families within villages. These differences can be a product of differences across age-sex cohorts in educational attainment and differences in the demographic composition of households, or can arise from differences within age-sex cohorts in schooling levels. The source of the inequality in educational attainment has potentially important implications for our interpretation of inequality in China and for policy.

---

[25] Schooling was severely disrupted during the famine between 1959 and 1961, and again in the mid-1960s, which explains some of the differences between later cohorts.

[26] See Kai-yuen Tsui (1997), Lavely, Xiao, Li, and Freedman (1990), and Emily Hannum (1998).

[27] As noted in our previous discussion, returns to education appear to have been zero, prior to the reforms.

Using as our measure of household schooling the average number of years of completed schooling of all household members 16 years and older that are not currently enrolled, we carried out decomposition exercises analogous to those performed for income. We are interested in assessing the contribution of inter-village and inter-cohort differences to overall school inequality. Several findings emerge. Differences in average years of schooling across villages only explains about 10 percent of the differences in schooling levels across households in our sample. In other words, most of the differences are within villages, and either arise from differences across age cohorts or within age cohorts. An examination of our household data reveal that differences across age cohorts are very stark: average schooling increases from 1.8 years for individuals 61+; to 5.1 for individuals 46-60; 6.2 for age group 31-45; 7.1 for 21-30,; and 7.6 years for the cohort aged 16-20. [28] These differences are the source of another 25 percent of the variation in average household schooling. Together, these factors explain slightly more than a third of the variation in household schooling, leaving nearly two thirds unexplained by age-cohort or village.

Returning to the income inequality summaries, Table 9 shows the regression results, with a variety of control variables. In columns I and III, we show results from a "lean" specification with simple summaries of factor-inequality: the Gini for per capita land and the Gini for average household education inequality, as well as some simple market/institutional variables. The market variables are: 1) The estimated rate of return to schooling; 2) The distance to major markets (county seat); 3) The share of income from non-agricultural sources; 4) The share of the village labour force employed in TVE's; and 5) the fraction of output sold (as opposed to self-consumed). The village marketing variables (distance, employment in TVE's, and the share of crops sold to markets) are taken from a separate village questionnaire, while the remaining variables are estimated from the household survey. Very few coefficients in the lean specification are significant, though the sign patterns are in line with what we

[28] These numbers probably slightly underestimate differences in schooling levels across cohorts because these calculations exclude village migrants who tended to be both younger and better educated.

expect given a simple model linking income to factor endowments and prices. Income inequality is higher in those villages where a higher fraction of income is earned from non-agricultural activities; and lower the closer a village is to the county seat, the more commercialized is agriculture, and the higher the fraction of the labour force employed by TVE's. The coefficients in the consumption inequality equation are generally less significant than the income regressions.

The most interesting results appear when we add interactions between the human capital distribution and the "institutional" measures. One way to think of these variables is that they proxy various dimensions of market opportunities. It comes as no surprise that there are interactions between the distribution of human capital, and the degree of market opportunity (reflecting the level of economic development, or the extent of "transition"). Interpreting interaction effects can be tricky, so we discuss the coefficients in some detail. Consider first the distribution of education. To help with interpretation, note that inequality of education is negatively correlated with the average level of schooling; that is, in villages where average schooling levels are higher, the Gini of education is lower. The total effect of education on income inequality depends on all of the interaction effects.

We begin with the separate partial effects, which show where the inequality of schooling is more or less disequalizing. First, where land inequality is higher, the disequalizing effect of education is lower. This may be a product of the land allocation process, where higher land inequality results from land being allocated from the more educated households with better off-farm opportunities. Second, there is a strong positive interaction between the distribution and returns to schooling: where education is unequally distributed, and its return highest, inequality is also highest. This suggests that policies that reduce the returns to schooling, or decrease education inequality (increasing school attainment should do both) will reduce inequality. We return to this point shortly. Third, the institutional variables affect the impact of schooling on inequality. The further a village is from a county seat, the smaller is the disequalizing effect of human capital. In other words, the distribution of education matters less in more remote villages. A similar interpretation applies to the commercialization variable. Unequal educational outcomes adversely

affect income inequality most when accompanied by commercial development. On the other hand, employment opportunities in TVE's reduce the adverse effect of education inequality, presumably by providing income opportunities for the less educated.

The institutional variables and their interactions with education can also be viewed on their own (as opposed to partial effects of education). The total institutional effects include the interaction terms, evaluated at some level of schooling inequality. For example, the median Gini of education inequality is 0.16 (it ranges from 0.08 to 0.28, with most lying between 0.15 and 0.21). The total effect of an increase in the rate of return to schooling on the Gini of income inequality is $-2.486 + (17.038 \times 0.16) = 0.24$, so that an increase in the rate of return to schooling by 1 percentage point (0.01) adds 0.0024 points to the Gini coefficient in a village with a median level of education inequality.[29] We obtain several interesting and statistically significant results. First, inequality is higher the further away a village is from the county seat. Economic opportunity, as represented by distance to markets, has an equalizing effect on income. However, inequality of education erodes the positive effect of market integration, and inequality worsens when opportunities exist in villages with unequal distributions of schooling. Second, the more commercialized a village, the lower its Gini. Again, this indicator of market transition worsens inequality when combined with education inequality. Viewed differently, commercialization reinforces the positive effects of a more equal distribution of human capital. Finally, TVE employment worsens inequality if education inequality is zero, but it improves the income distribution for any Gini above 0.17, which means any Gini above the median.

In summary, there are important interactions between human capital and market development in determining the level of inequality at the village level. Generally, commercialization and economic opportunity equalize incomes. However, the effects are less benign when accompanied by an unequal

[29] In fact, 0.01 is a small change in the rate of return to schooling given the variation in our sample. While the sampling variation is also large (see Table 5), we estimate village rates of return to education at an average of 0.07 for income, with a standard deviation of 0.08 (the 25-75 percentile split is 0.03 and 0.10); and an average of 0.04 for expenditure, with a standard deviation of 0.04 (the 25-75 percentile split is 0.02 to 0.07).

distribution of education, whereby these opportunities will presumably be taken advantage of most by the more educated, at the expense of a higher income inequality.

V. Conclusions

Our empirical explorations yield a number of suggestive results linking economic transition to income inequality in China. Two important themes emerge. First, is the apparent role that "economic opportunity" plays in determining the relative position of winners and losers in transition. Second, is the role that human capital plays in allowing households to access these opportunities. Inequality of economic development interacts with the unequal distribution of human capital, leading to more inequality within villages than differences of income across villages. One relatively pessimistic implication of our results is that rural inequality is likely to worsen before it improves. The distribution of human capital likely changes quite slowly, while market institutions and opportunities may change rapidly. Given the current distribution of education, many of these institutional developments will disproportionately benefit the higher educated. Compounding this, current patterns of capital accumulation suggest that the rich will be better positioned to increase their incomes, and thus their future wealth.

Of course, many questions remain unanswered, and our results provoke more questions than they answer. For example, we ignore the role of households as an economic institution. We also ignore cohort dimensions to inequality. For example, is education primarily unequally distributed among generations (cohorts), or is there considerable inequality of education within all age groups? The answers to these questions have important implications for the likely speed with which the education distribution, and thus the income distribution, will evolve. Nevertheless, our explorations suggest a number of important avenues for future empirical work, and highlight the importance of collecting new household survey data.

VI. References

An, Mark and Dennis Yang (1998): "Human Capital, Entrepreneurship, and Farm Household Earnings," mimeo, Duke University.

Benjamin, Dwayne and Loren Brandt (1997): "Land, Factor Markets, and Inequality in Rural China: Historical Evidence," Explorations in Economic History, 34(4), October, pages 460-94.

Benjamin, Dwayne and Loren Brandt (1999): "Markets and Inequality in Rural China: Parallels with the Past," American Economic Review, 89(2), May, pages 292-295.

Byrd, William and Alan Gelb (1990): "Township, Village and Private Enterprise in China's Economic Reform," World Bank, Policy Research and External Affairs Working Papers Series, no. 406.

Cheng, Yuk-Shing (1996): "A Decomposition Analysis of Income Inequality of Chinese Rural Households," China Economic Review, 7(2), pp. 155-167.

Deaton, Angus (1997): The Analysis of Household Surveys, The World Bank, Washington D.C.

Deaton, Angus and Christina Paxson (1994): "Intertemporal Choice and Inequality," Journal of Political Economy,102(3), June, pages 437-67.

Glewwe, Paul and Margaret Grosh (1998): "The World Bank's Living Standards Measurement Study Household Surveys," Journal of Economic Perspectives, 12(1), Winter, pp. 187-196.

Gregory, Robert and Xin Meng (1995): "Wage Determination and Occupational Attainment in the Rural Industrial Sector of China," Journal of Comparative Economics, 21, pages 353-374.

Guojia Jiwei Yanjiu Suo Ketizu (Chinese State Planning Commission, Economic Research Institution Research Group) (1995): "Woguo dongzhongxi san da diqu nongcun jingji fazhan chaju de fenxi,"(Analysis of Uneven Rural Economic Development of the Eastern, Central and Western Regions of China), Jingji Yanjiu Cankao (References on Economic Research),30, pp 2-42.

Hannum, Emily (1998): "Educational Inequality: Hidden Consequences of the Reform Era in Rural China," Ph.D. Dissertation, University of Michigan.

Hare, Denise (1994): "Rural Nonagricultural Activities and Their Impact on the Distribution of Income: Evidence from Farm Households in Southern China," China Economic Review, 41, pp 59-82.

Ho, Samuel (1994): Rural China in Transition, New York: Oxford University Press.

Howes, Stephen and Athar Hussain (1994): "Regional Growth and Inequality in Rural China," mimeograph, London School of Economics.

Hussain, Athar, Peter Lanjouw, and Nicholas Stern (1994): "Income Inequalities in China: Evidence from Household Survey Data," World Development, 22(12), pp. 1947-57.

Khan, Azizur and Carl Riskin (1998): "Income and Inequality in China: Composition, Distribution and Growth of Household Income, 1988 to 1995," China Quarterly, 154, pp. 221-251.

Knight, John and Li Shi (1997): "Cumulative Causation and Inequality Among Villages in China," Oxford Development Studies, 25.2, pages. 149-172.

Knight, John and Lina Song (1993), "The Spatial Contribution to Income in Inequality in Rural China," Cambridge Journal of Economics, 17, pages 195-213.

Lardy, Nicholas (1994): China in the World Economy, Institute for International Economics: Washington, D.C.

Lavely, William, Zhenyu Xiao, Bohua Li, and Ronald Freedman (1990): "The Rise in Female Education in China: National and Regional Patterns," China Quarterly, 121, pp. 61-93.

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

Lin, Justin (1992): "Rural Reforms and Agricultural Growth in China, American Economic Review, 82.1, pages 34-51.

Ministry of Agriculture (1995): Zhongguo Nongye Fazhan Baogao '95 (China's Agricultural Development Report, 1995). Beijing: Zhongguo Nongye Chubanshe.

Morduch, Jonathan and Terry Sicular (1999): "Politics, Growth, and Inequality in Rural China: Does it Pay to Join the Party?," Mimeo, The University of Western Ontario.

Naughton, Barry (1996): Growing out of Plan, New York: Cambridge University Press.

Orazem, Peter and Milan Vodopivec (1995): "Winners and Losers in Transition: Returns to Education, Experience, and Gender in Slovenia," World Bank Economic Review, 9(2), May, pages 201-230.

Park, Albert and Scott Rozelle (1998): "Reforming State-Market Relations in Rural China," mimeo, University of Michigan.

Psacharapoulos, George (1985): "Returns to Education: Further International Update and Implications," Journal of Human Resources, 20(4), December, pages 583-604.

Putterman, Louis (1993): Continuity and Change in China's Rural Development : Collective and Reform Eras in Perspective, New York : Oxford University Press.

Ravallion, Martin and Shaohua Chen (1996): "Data in Transition: Assessing Rural Living Standards in Southern China," China Economic Review; 7(1), Spring, pages 23-56.

Riskin, Carl (1994): "Chinese Rural Poverty: Marginalized or Dispersed?" American Economic Review, 84(2), May, pages 281-284.

Rosenzweig, Mark (1995): "Why Are There Returns to Schooling?" American Economic Review, 85(2), May, Pages 153-158.

Rozelle, Scott (1994), "Rural Industrialization and Increasing Inequality: Emerging Patterns in China's Reforming Economy," Journal of Comparative Economics, 19, pp. 362-391.

Schultz, Theodore W. (1975): "The Ability to Deal with Disequilibria," Journal of Economic Literature, 13(3), September, pages 827-46.

Shorrocks, Anthony (1982): "Inequality Decomposition by Factor Components," Econometrica, 50(1), Jan. 1982, pages 193-211.

Shorrocks, Anthony (1983): "The Impact of Income Components on the Distribution of Family Incomes," Quarterly Journal of Economics, 98(2), May, pages 311-26.

Sicular, Terry (1995): "Redefining State, Plan, and Market: China's Reforms in Agricultural Commerce," China Quarterly, vol. 144, pages 1020-1046.

State Statistical Bureau: Zhongguo Tongji Nianjian, (Statistical Yearbook of China), Beijing, selected years.

Tsui, Kai-Yuen (1998): "Factor Decomposition of Chinese Rural Income Inequality: New Methodology, Empirical Findings, and Policy Implications," Journal of Comparative Economics, 26, pages. 502- 528.

Tsui, Kai-Yuen (1997): "Economic Reform and Attainment in Basic Education in China," China Quarterly, 149, pp. 104-12.

Tsui, Kai-Yuen (1991): "China's Regional Inequality, 1952-1985,"Journal of Comparative Economics, 15(1), March, pp. 1-21.

Turner, Matthew, Loren Brandt and Scott Rozelle (1999): "Property Rights Formation and the Organization of Exchange in Rural China", mimeo, University of Toronto.

Van Luong, Hy and Jonathan Unger (1998) "Wealth, Power, and Poverty in the Transition to Market Economies: The Process of Socio-Economic Differentiation in Rural China and Northern Vietnam," The China Journal, No. 40, July, pages 61-93.

Vermeer, E.B. (1982): "Income Differentials in Rural China," China Quarterly, 111, pp. 1-33.

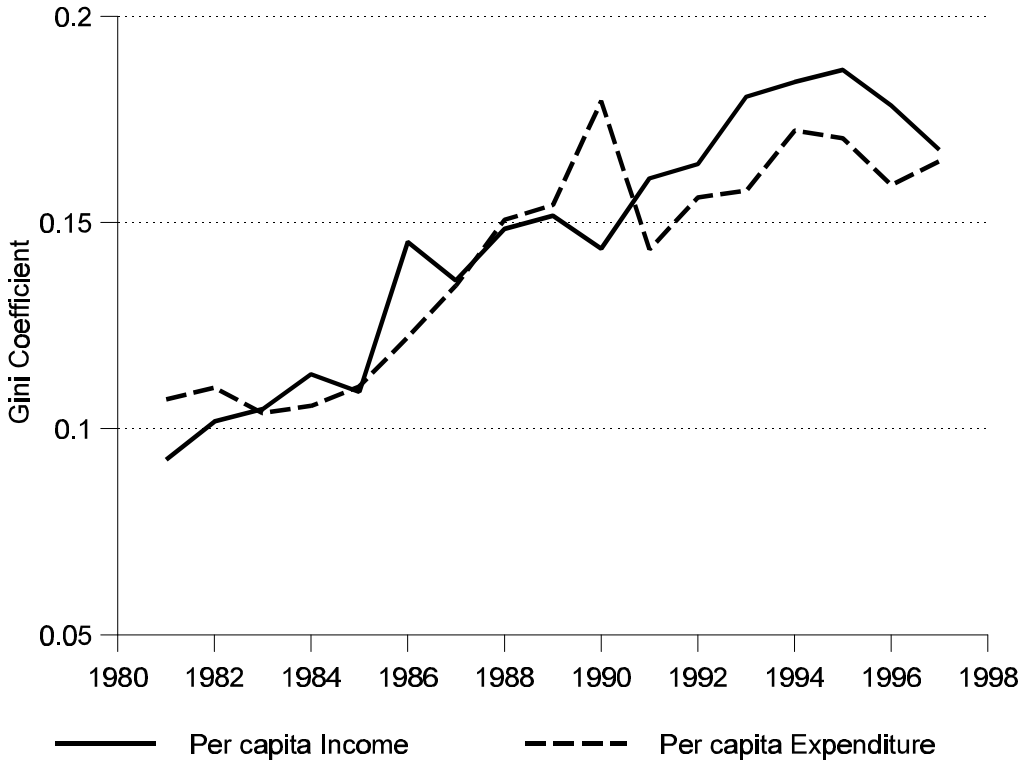
West, Lorraine and Yaohui Zhao, Editors (1999): The Flow of Rural Labor in China. Berkeley: University of California Press.

Yang, Dennis (1997), "Education and Off-farm Work," Economic Development and Cultural Change, 45, pp. 613-32.

Yang, Dennis (1999): "Urban-Biased Policies and Rising Income Inequality in China," American Economic Review, 89(2), pages 306-310.

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

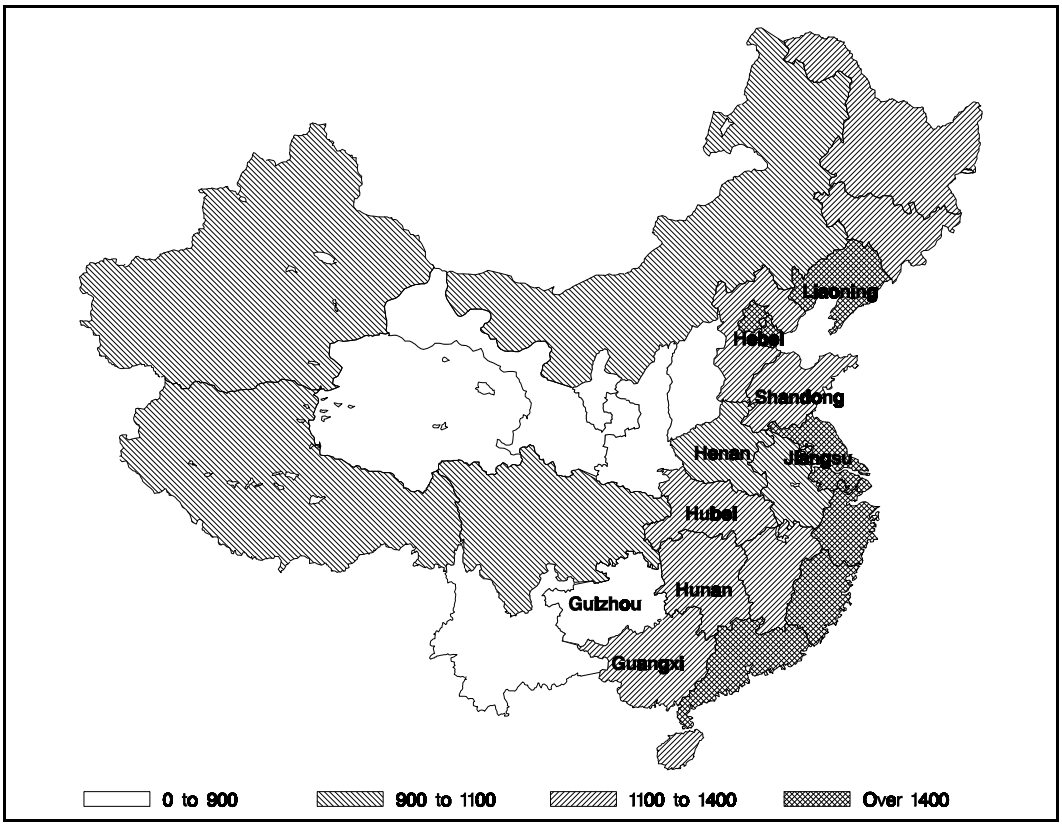
FIGURE 1: SPATIAL (RURAL) INEQUALITY ACROSS PROVINCES OF CHINA
GINI COEFFICIENT



Notes: This figure shows the simulated level of inequality (Gini coefficient), assuming that everyone in a province has the same living standards (per capita income or consumption).

Source: Author's calculations based on rural household survey data of the State Statistical Bureau. Provincial means are reported on an annual basis in Zhongguo Tongji Nianjian (Statistical Yearbook of China), various years.

FIGURE 2: SAMPLE PROVINCES AND RURAL PER CAPITA INCOME, 1994



Notes: Map shows average provincial per capita income in 1994. Provinces sampled by CHNS and NCLSS are also labelled.

Source: Author's calculations based on rural household survey data of the State Statistical Bureau. Provincial means are reported on an annual basis in Zhongguo Tongji Nianjian (Statistical Yearbook of China), various years.

FIGURE 3: KERNEL DENSITY ESTIMATES OF LOG PER CAPITA INCOME AND EXPENDITURE

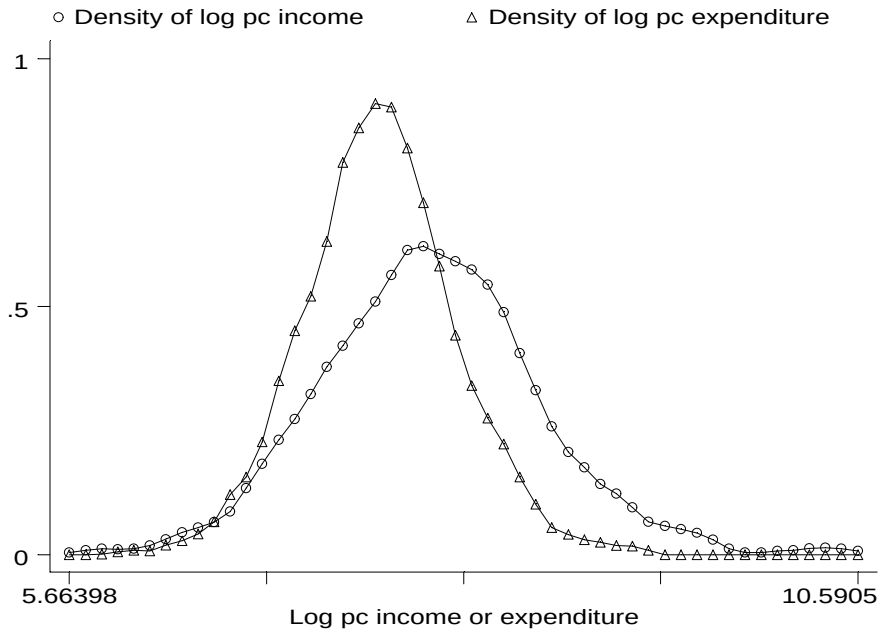


TABLE 1
SOURCES AND DISTRIBUTION OF HOUSEHOLD INCOME

Source of Income:	Mean	Percentage Not Zero	Gini Coefficient
Men's Wages	1578	41	.41
Women's Wages	412	16	.40
Farm income	5714	98	.48
Animal Husbandry	1506	77	.75
Non-agricultural Family Enterprises	1837	30	.64
Other income	959	65	.66
Imputed Capital Services (House + durables)	1065	100	.39
Total Income	13071	100	.40
Total Expenditure	8378	100	.28
Per Capita Income	3510	100	.39
Per Capita Expenditure	2250	100	.26

Notes: 1) Means are presented over the whole sample (778 households) in the first two columns. 2) The Gini Coefficient is calculated only over those households with positive income from that particular source. 3) Total Income is the sum of income from all sources including the imputed value of owner occupied housing and durables (Imputed Capital Services). Similarly, Total Expenditure includes the value of imputed capital services. 4) Farm Income includes crop income, plus income from agricultural sidelines (vegetables, fruit, wood, and greenhouses) and income from the sale of crop by-products. 5) Animal husbandry includes income from animal husbandry plus aqua-culture and agricultural businesses reported in the "family run business" section of the survey that are mostly poultry farms. 6) Non-agricultural Family enterprises are family run businesses that are not directly "agricultural." 7) Other income includes rental income, remittances, transfer income, and interest income.

TABLE 2
PROPORTION OF INCOME VARIATION / INEQUALITY EXPLAINED BY LOCATION

	TOTAL	Province	County	Village
Per Capita Expenditure, 1995 Survey				
Gini	.261			
Variance of log(pcx)	.219	.090	.167	.249
Theil L	.101	.097	.169	.304
Theil T	.128	.077	.130	.248
Per Capita Income, 1995 Survey				
Gini	.385			
Variance of log(pcy)	.464	.057	.117	.245
Theil L	.238	.098	.167	.329
Theil T	.291	.080	.134	.305
Per Capita Income, CHNS 1993 Survey				
Gini	.495			
Variance of log(pcy)	1.175	.036	.151	.267
Theil L	.496	.073	.166	.290
Theil T	.469	.076	.173	.303

Notes: 1) Sample size for the 1995 data is 773 for income, and 778 for expenditure. 2) Sample size for the 1993 CHNS is 1653 (positive income observations). 3) TOTAL indicates the total sample inequality for each measure of inequality; 4) The remaining three columns show the proportion of each inequality index that is explained by differences in income across different geographic units (province, county, and village). 5) For the 1995 data, there are 5 villages in each of 6 counties (30 villages), split evenly between two provinces; 6) For the CHNS data, there are 3 villages in 4 counties in each of 8 provinces (for a total of 32 counties and 96 villages).

TABLE 3
DECOMPOSITIONS OF INCOME INEQUALITY BY SOURCE OF INCOME

	Full 1995 Sample			1995 Sample Excluding Highest Income		
	Share	**OLS**	**2SLS**	**Share**	**OLS**	**2SLS**
Source:						
Other Income	0.073	0.061	0.090	0.072	0.047	0.080
Durables and Housing	0.081	0.029	0.059	0.082	0.035	0.077
Men's Wages	0.120	0.023	0.054	0.123	0.031	0.075
Women's wages	0.031	-0.001	0.015	0.032	0.000	0.020
Crop income	0.435	0.236	0.178	0.444	0.302	0.247
Non-Ag Businesses	0.141	0.433	0.478	0.125	0.312	0.333
Animal Husbandry	0.119	0.219	0.125	0.122	0.273	0.168

CHNS 1993 Sample

	Mean	**Share**	**OLS**
Other Income (per capita)	82	0.096	0.029
Wage Income (per capita)	206	0.240	0.147
Crop Income (per capita)	294	0.344	0.132
Non-Ag Businesses (per capita)	187	0.219	0.379
Animal Husbandry (per capita)	51	0.059	0.287
Total Per Capita Income	856		

Notes: 1) The shares are the percentages of household income accounted for by each source of income. 2) OLS is the OLS estimate of the correlation between total per capita income and the per capita income of a given type. 3) 2SLS is the 2SLS estimate of this correlation, using total per capita expenditure (less capital services) as an instrument. This instrument is not available for the CHNS sample. 4) All means and correlations are reported weighted by the number of individuals in the household. 5) The income sources for the 1995 sample are described in Table 1. 6) For the CHNS, analogous income categories are created. Sample size in the CHNS is 1653 (the number of households with positive income).

TABLE 4
DECOMPOSITIONS OF INCOME INEQUALITY BY SOURCE OF INCOME
INTERACTIONS WITH SPATIAL INEQUALITY

		Full 1995 Sample			
		Coefficients on PCY With Village Fixed Effects		F-Tests for Different Coefficients by Village	
	Share	OLS	2SLS	OLS	2SLS
Source:					
Other Income	0.073	0.064	.101	4.71 (.000)	1.22 (.201)
Durables and Housing	0.081	0.024	.059	3.51 (.000)	1.65 (.018)
Men's Wages	0.120	0.029	.070	6.20 (.000)	1.21 (.204)
Women's wages	0.031	-0.001	.013	0.72 (.862)	1.02 (.438)
Crop income	0.435	0.179	.073	10.95 (.000)	1.81 (.006)
Non-Ag Businesses	0.141	0.468	.560	10.09 (.000)	1.28 (.151)
Animal Husbandry	0.119	0.237	.122	5.69 (.000)	0.38 (.999)
		1995 Sample Excluding Highest Income			
Other Income	0.072	.048	.089	4.73 (.000)	1.04 (.416)
Durables and Housing	0.082	.029	.080	2.75 (.000)	0.91 (.601)
Men's Wages	0.123	.038	.099	5.49 (.000)	1.04 (.416)
Women's wages	0.032	-.001	.018	0.72 (.869)	1.00 (.470)
Crop income	0.444	.249	.149	7.93 (.000)	1.33 (.119)
Non-Ag Businesses	0.125	.333	.376	12.97 (.000)	1.48 (.051)
Animal Husbandry	0.122	.304	.189	14.13 (.000)	0.82 (.074)

Notes: 1) All specifications are the same as in table 3, except village fixed effects have been added. 2) The F-tests (p-values in parentheses) are tests of the joint significance of the village × income interaction terms in a separate analysis of covariance specification. 3) The shares are the percentages of household income accounted for by each source of income. 4) OLS is the OLS estimate of the correlation between total per capita income and the per capita income of a given type (with village fixed effects). 5) 2SLS is the 2SLS estimate of this correlation, using total per capita expenditure (less capital services) and village interactions with this variable as instruments. 6) All means and correlations are reported weighted by the number of individuals in the household. 5) The income sources for the 1995 sample are described in Table 1.

TABLE 5
DETERMINANTS OF HOUSEHOLD PER CAPITA INCOME AND CONSUMPTION
(standard errors in parentheses)

	ln PCY		ln PCX	
	(I)	(II)	(III)	(IV)
Average Years of Schooling of Adults in the Household	0.081 (0.012)	0.066 (0.012)	0.052 (0.008)	0.038 (0.008)
Number of members with Technical Training	0.171 (0.043)	0.109 (0.042)	0.135 (0.030)	0.103 (0.028)
Number of members with apprentice training	0.006 (0.045)	-0.016 (0.043)	0.040 (0.031)	0.016 (0.028)
Indicator for having allocated land	-1.389 (0.199)	-1.093 (0.218)	-0.547 (0.136)	-0.441 (0.145)
Log of allocated land	0.307 (0.032)	0.309 (0.036)	0.077 (0.022)	0.097 (0.024)
Share of land suffering adverse shock	-0.558 (0.143)	-0.613 (0.167)	-0.073 (0.098)	-0.208 (0.111)
Log Household Size	-0.510 (0.086)	-0.487 (0.085)	-0.370 (0.059)	-0.399 (0.057)
F-Test for Village Fixed Effects	NA	5.53 (0.000)	NA	7.23 (0.000)
F-Test for Village × Education	NA	1.51 (0.04)	NA	1.22 (0.20)
R-Squared	0.24	.381	.26	.43

Notes: 1) The coefficients are from an OLS regression of ln PCY (per capita income) or ln PCX (per capita expenditure) on the variables shown in the table, PLUS controls for land quality and household age-sex structure (the coefficients of which are not shown). 2) In columns II and IV, village fixed effects are added, with the f-test and p-value of the joint significance of the village fixed effects also reported. 3) In a separate specification, interaction effects between village and education (years of schooling) are added to the same based village-fixed effects specification. The F-tests for these interaction terms are reported in columns II and IV. 4) The sample sizes are 773 for income, and 778 for expenditures.

TABLE 6
PROBITS FOR HOUSEHOLD ENGAGEMENT IN VARIOUS ACTIVITIES
(standard errors in parentheses)

	Animal Husbandry	Male Wage Labour	Female Wage Labour	Non-Ag Business
Mean Probability:	0.78	0.43	0.19	0.33
Average Years of Schooling of Adults in the Household	-0.072 (0.038)	0.076 (0.030)	0.151 (0.042)	-0.006 (0.031)
Number of members with Technical Training	0.150 (0.128)	0.241 (0.107)	-0.092 (0.121)	0.183 (0.106)
Number of members with apprentice training	0.032 (0.127)	0.182 (0.108)	0.017 (0.123)	0.299 (0.106)
Indicator for having allocated land	-0.348 (0.582)	0.838 (0.560)	1.760 (0.882)	0.17 (0.573)
Log of allocated land	0.129 (0.102)	-0.174 (0.094)	-0.104 (0.123)	-0.198 (0.099)
Share of land suffering adverse shock	0.043 (0.492)	0.059 (0.428)	0.063 (0.546)	-0.818 (0.461)
Log Household Size	0.312 (0.253)	0.564 (0.221)	2.189 (0.326)	0.534 (0.230)

Notes: 1) The coefficients are the estimated coefficients from Probit estimation of whether a household has income from a particular source. 2) The fraction of households engaged in each activity is in the Mean Probability Row. 3) All equations include the same regressors as reported in table 5, including village fixed effects. Some of the village effects predict an activity perfectly. 4) All estimation is conducted with sample weights equal to household size.

TABLE 7
DETERMINANTS OF LEVELS OF HOUSEHOLD PER CAPITA INCOME IN VARIOUS ACTIVITIES
(standard errors in parentheses)

	Crop Income	Animal Husbandry	Male Wage Labour	Female Wage Labour	Non-Ag Business	All Non-Crop
Average Years of Schooling of Adults in the Household	0.020 (0.013)	0.021 (0.045)	0.088 (0.026)	0.177 (0.062)	0.084 (0.061)	0.098 (0.026)
Number of members with Technical Training	-0.030 (0.046)	-0.080 (0.146)	0.178 (0.084)	0.204 (0.139)	0.133 (0.193)	0.237 (0.088)
Number of members with apprentice training	-0.055 (0.046)	-0.244 (0.163)	0.050 (0.086)	-0.032 (0.167)	-0.050 (0.179)	0.101 (0.091)
Indicator for having allocated land	-1.774 (0.257)	-1.952 (0.981)	0.794 (0.528)	2.660 (1.347)	-0.233 (1.071)	0.810 (0.489)
Log of allocated land	0.775 (0.039)	0.181 (0.145)	-0.211 (0.090)	-0.176 (0.187)	0.123 (0.202)	-0.208 (0.083)
Share of land suffering adverse shock	-0.942 (0.180)	-0.321 (0.651)	-0.168 (0.394)	-0.050 (0.771)	0.539 (0.926)	-0.073 (0.369)
Log Household Size	-0.903 (0.092)	-0.239 (0.329)	-0.489 (0.202)	-0.948 (0.424)	-0.746 (0.464)	0.155 (0.189)
Sample Size	741	446	320	125	213	659
R-Squared	0.62	0.233	0.42	0.63	0.28	0.23

Notes: 1) The coefficients are estimated from regressions of log per capita income of each type on the same set of explanatory variables as described in table 5 (including village fixed effects). 2) Regressions are weighted by household size. 3) "All Non-Crop" is income from all income categories except crops and "non-labor" income, i.e., the sum of animal husbandry, wage labour, and non-agricultural businesses.

TABLE 8
SAVINGS AND ASSET ACCUMULATION
(standard errors in parentheses)

	I	II	III	IV	V	VI
Dependent Variable:	**Savings "Rate"**	**ln per capita Assets**	**ln per capita Assets**	**ln per capita Assets**	**ln per capita Assets**	**ln per capita Assets**
Explanatory Variable:	**OLS**	**OLS**	**2SLS**	**OLS**	**OLS**	**2SLS**
Savings "Rate"					0.113	0.821
					(0.058)	(0.256)
Ln Per Capita Income		0.602	1.493			
		(0.049)	(0.118)			
Average Years of Schooling of the adults in the household	0.028			0.081	0.101	0.091
	(0.011)			(0.018)	(0.015)	(0.017)
Number of members with Technical Training	0.006			0.246	0.221	0.237
	(0.039)			(0.063)	(0.062)	(0.068)
Number of members with apprentice training	-0.032			0.117	0.100	0.120
	(0.040)			(0.064)	(0.063)	(0.069)
Fraction of Land suffering adverse shock	-0.408			-0.563		
	(0.158)			(0.252)		
R-Squared	0.17	0.28	-	0.27	0.23	0.07

Notes: 1) Savings is measured as ln(income) - ln (expenditure); 2) Assets are measured as ln (per capita assets); 3) All specifications include village fixed effects. 4) For specification I and IV, additional controls are the same as in Tables 5-7. 5) For specification II, only the village fixed effects are included, as well as lnpcy; 6) For specification III, lnpcy is instrumented with lnpcx, as in Table 3. 7) For specification V, only the education, log household size, and land indicators, as well as the village fixed effects are included as controls. 8) In specification VI, the savings rate is instrumented by the shock variable, farm size, and the ratio of kids and teens. The overidentification test is 1.96 (distributed $\chi^2(5)$ under the null hypothesis of no correlation between identifying instruments and error term) which is insignificant at conventional levels.

TABLE 9
SUMMARIZING PATTERNS OF VILLAGE INEQUALITY
(Standard errors in parentheses)

	Gini Per Capita Income		Gini of Per Capita Expenditure	
	OLS	**OLS**	**OLS**	**OLS**
	(I)	(II)	(III)	(IV)
Gini of Education	0.105	-1.667	0.251	0.200
	(0.260)	(0.733)	(0.220)	(0.882)
Gini of Per Capita land	0.162	0.871	0.061	0.748
	(0.101)	(0.325)	(0.071)	(0.337)
Gini of Land × Gini of Education		-5.885		-4.819
		(1.933)		(1.983)
Rate of Return to Schooling (income)	0.301	-2.486		
	(0.161)	(0.607)		
Rate of Return to Schooling (expend)			0.006	1.430
			(0.197)	(0.998)
Gini of Ed × Rate of Return		17.038		-8.983
		(3.565)		(6.031)
Share of income from Non-ag business	0.167	-0.140	-0.065	-0.147
	(0.082)	(0.090)	(0.065)	(0.118)
Share of non-ag × Gini of Education		0.000		0.000
		(0.000)		(0.000)
Distance from County Seat	0.000	0.004	0.000	0.001
	(0.000)	(0.001)	(0.000)	(0.001)
Distance × Gini of Education		-0.015		-0.008
		(0.004)		(0.005)
Fraction of Ag output sold	-0.001	-0.008	-0.001	-0.007
	(0.000)	(0.002)	(0.000)	(0.002)
Fraction Sold × Gini of Education		0.039		0.034
		(0.009)		(0.011)
Fraction Employed in Rural Enterprises	-0.060	0.603	-0.012	1.110
	(0.083)	(0.230)	(0.070)	(0.342)
TVE × Gini of Education		-3.673		-6.628
		(1.327)		(1.968)
R-squared	0.610	0.920	0.340	0.700

Notes: 1) The coefficients are estimated from a regression of the village-level Gini-coefficient for per capita income and expenditure on the variables listed above. 2) All regressions are weighted by village population. 3) The rate of return to schooling at the village level is estimated from the interaction effects in the income and expenditure regressions reported in Table 5.