The Evolution of Income Inequality in Rural China

DWAYNE BENJAMIN and LOREN BRANDT
University of Toronto

JOHN GILES
Michigan State University

I. Introduction

China has recorded impressive growth over the past 20 years, with a commensurate increase in average living standards.[1] However, there is mounting concern that increases in inequality indicate that many are being left behind and not sharing in the fruits of development.[2] Indeed, most well-behaved social welfare functions rank unequal distributions below equal ones with the same mean income levels. An important question is whether recent increases in inequality are sufficiently high to offset general rises in average income: is rising inequality an uncomfortable but otherwise innocuous price to pay when the rising tide is raising all boats? Not necessarily, as there is additional concern that high levels of inequality may slow economic transition and hinder future growth.[3]

One striking feature of the current discussion of inequality in China is the absence of well-documented facts about inequality and temporal changes in

We are grateful for helpful comments and suggestions from Carsten Holz, Martin Ravallion, Sangui Wang, and two anonymous referees. Dwayne Benjamin and Loren Brandt thank the Social Sciences and Humanities Research Council of Canada for financial support. John Giles gratefully acknowledges financial support from the National Science Foundation (SES-0214702) and the Intramural Research Grants Program at Michigan State University.

[1] Annual series from the National Bureau of Statistics (NBS) suggest an average GDP per capita growth rate of 8.2% from 1980 and 2000 (NBS 2003). Rawski (2001), among others, has criticized China's recent GDP statistics, but few dispute the considerable growth that China has experienced in recent years since the onset of reforms.

[2] The World Bank (2003), Beijing office, notes concerns about the consequences of increasing inequality for support for continued economic reform, and the new leadership in Beijing openly voices concerns about potential adverse consequences of rising inequality (Hutzler 2003).

[3] See Aghion, Caroli, and Garcia-Penalosa (1999) for a review of the growth-inequality relationship from the perspective of new growth theory; Banerjee and Duflo (2004) for a cross-country growth regression analysis; and Dayton-Johnson and Bardhan (2002) for an analysis of the impact of inequality on common property management in village settings.

 0013-0079/2005/5304-0001$10.00

the patterns and structure of the income distribution. To some extent, this reflects lack of access to nationally representative data, so that China has been cited as an exception to global trends in our understanding of inequality in the developing world.[4] Certainly, the level of basic knowledge about China's income distribution pales in comparison to attention paid to developed countries like the United States or other developing countries like Taiwan, Thailand, or India. Furthermore, China is simply bigger and more complicated so that summaries of inequality may be less meaningful than those of other economies. Still, simple impressions of inequality have emerged from the existing literature: first, inequality has gone up during the transition; second, this is largely driven by widening interprovincial income differences; and third, in rural areas, the development of nonfarm opportunities has provided uneven rewards for households and is an important underlying source of inequality.[5]

Our objective in this article is to fill in gaps in our understanding of inequality in rural China. The centerpiece of our work is a nationally representative household survey that has been collected by the Research Centre for Rural Economy (RCRE) under the Ministry of Agriculture in China continuously from 1986 to the present and that covers most of the reform period.[6] By using a common household survey across years, we are able to address a number of important methodological and measurement issues associated with describing inequality at a point in time and comparing inequality across time periods. By employing additional, different data sets, we are also able to explore the sensitivity of our conclusions to the use of our primary data set. We further provide simple decompositions of inequality by space (village, province, and region) and source of income (e.g., farm and nonfarm) that yield important insights about the evolution of inequality.

Ideally, we would like to attribute changes in inequality to various factors associated with economic transition (moving to a market economy) and development (e.g., growth of a nonfarm sector).[7] This is difficult, however, as both processes are potentially confounded in a common trend and almost certainly intertwined anyway. Moreover, some of the recent rise in inequality appears to be a by-product of collapsing agricultural prices and not the consequence of a "Kuznets-like" structural process.[8] While far from perfectly

[4] See Deaton (2005).

[5] See, e.g., Rozelle (1994), Gustafsson and Li (2002), Morduch and Sicular (2002), Xin Meng (2004).

[6] There were gaps in the panel in 1992 and 1994 when the survey was not conducted.

[7] This simple (and optimistic) two-factor attribution of the sources of inequality during Chinese transition is outlined in Benjamin and Brandt (1999).

[8] For the original discussion, see Kuznets (1955).

integrated with world markets, crop prices in China fell by more than a third between 1996 and 2000, mirroring large drops in world prices.[9] Since many households in rural China are still dependent on crop production for a substantial portion of their incomes, they have experienced absolute, not just relative, declines in their standard of living. While we cannot explain the drop in crop prices, we show that it is clear that if prices remain low, without offsetting rapid development of the more equalizing sources of nonfarm income, many in the countryside will remain poor, with commensurate political and migration pressures. In fact, we may now be observing China in a state of transition, with low returns to agriculture (as currently structured) a catalyst for a more rapid secular shift away from farming. Whether the Chinese economy is flexible enough to facilitate this transformation or public policy sufficiently nimble and focused on the human cost of this potentially massive adjustment remain important open questions.

We first briefly review the existing literature on rural inequality, highlighting several data and conceptual issues that we focus on in our work. We then describe our main data set, based on a panel of villages surveyed by RCRE. Included is a discussion of issues arising in the consistent measurement and definition of income and consumption from 1987 to 1999. We summarize various features of the income distribution for selected years in this time span and show that, while average incomes have undeniably risen, so has inequality. Of particular concern, we show that, after initially rising, the absolute living standards of the poor declined considerably from 1995 to 1999, so that they approach income levels of 1987. Moreover, as much as half of all households were not unambiguously better off in 1999 than in 1987: the rising tide did not lift all boats. After describing the overall distribution, we then present a spatial decomposition, where we challenge the popular perception that inequality in rural China is primarily a geographic phenomenon. In fact, most inequality is local. Our final exercise breaks down total income by source, where we see that the increase in inequality is driven by the combination of falling farm incomes with rising local nonfarm incomes. Wage incomes from temporary migrant employment, by contrast, are actually correlated with reductions in inequality. In the final section, we offer some interpretation of these results and outline questions for future research.

II. Previous Studies

There is an extensive literature concerned with the evolution of inequality in rural China since the onset of reform in the late 1970s. The primary focus

[9] As an example, the price of rice in the United States fell from $463.97 per metric ton in 1996 to $367.36 per ton in 2000, a nominal decline of about 20%.

has been on (i) estimating the level of inequality and its changes over time and (ii) identifying underlying sources of inequality and its changes. Our purpose here is not to offer a comprehensive literature review of this work. We will be very selective, with an eye to major findings and some of the limitations that have informed our analyses with the RCRE data. At the risk of some simplification, however, the general consensus is that inequality has increased significantly over time. Motivated by the contrasting economic performance of rich and poor provinces, much of this increase is attributed to spatial differences linked to regional factors and the highly uneven rate of growth of the nonagriculture sector in the countryside.

Data collected by the National Bureau of Statistics (NBS) have been the sole source of estimates of long-run trends in rural income inequality using household-level data. These data are not in the public domain, and details of construction of the NBS's income estimates and Gini's is sketchy. Nonetheless, Bramall's (2001) summary of NBS trends shows an increase in the national Gini coefficient for rural China of almost 50%, from 0.24 in 1980, to 0.31 in 1990, to 0.34 in 1995, and finally to 0.35 in 1999. The only estimates of rural inequality that are remotely comparable to those of the NBS are for 1988 and 1995 using data collected as part of the China Income Project (CIP; see Khan and Riskin 1998). Covering a smaller sample of provinces and households and based on a modified NBS household survey instrument and definition of income, these data suggest both higher levels of overall inequality and a more rapid increase over a subperiod, with the Gini coefficient rising from 0.34 in 1988 to 0.42 in 1995.

Numerous studies stress the spatial aspects of this rise in rural inequality. A majority of these, however, use provincial-level or subprovincial per capita averages, as opposed to household level survey data. Only four out of 16 recent papers summarized by Gustafsson and Li (2002), for example, used household-level data. Papers using regionally aggregated data have been interested in looking at trends in interprovincial inequality or have used provincial-level data to analyze interregional trends. There are two obvious limitations of these analyses. First, they underestimate inequality because they ignore any differences arising from household differences within administrative units. And second, conclusions about trends and the role of contributing factors to inequality such as township and village enterprises are valid only insofar as most rural inequality arises from differences in mean incomes across these units.

A few studies have employed household-level data to decompose rural inequality into spatial components. Benjamin et al. (2002) showed (for a single point in time) that within-village inequality dominated cross-region inequality, although their study had limited geographic coverage. Gustafsson and Li

(2002), using the CIP data for 1988 and 1995, provide the most comprehensive study exploring inequality across space and time. Their analysis suggests that the contribution of spatial differences at the county or provincial level was significant and rising between 1988 and 1995. Differences in the counties (but not provinces) from which households were surveyed in the two years, however, pose some problems for interpretation.

Finally, a number of other studies (Hare 1994; Khan and Riskin 1998; Tsui 1998; Kung and Lee 2001) have used household data to look at the role of the emerging nonagricultural sector in explaining inequality. These studies emphasize the potential role played by the changing structure or composition of income in generating higher inequality. Several authors have also stressed the role played by political power and connections in facilitating access to new opportunities (Nee 1992; Cook 1998; Morduch and Sicular 2002). With the exception of Benjamin et al. (2002), these studies do not separate or net out the spatial dimensions of income composition, but they do confirm the significant contribution of nonagricultural income sources to income inequality. Interpretation is handicapped, however, in the way that alternative income sources are often lumped together. For example, wage income from local and nonlocal sources is usually aggregated, while total wage income is sometimes combined with income from family-run businesses. Insofar as these sources of income are less than perfectly correlated with each other, grouping them together hides important aspects of emerging inequality and their links to household attributes and the external economic environment with which these households interact.

III. Data

The data used for our analyses come from annual household surveys conducted by the Survey Department of the Research Center on the Rural Economy in Beijing. Household-level surveys from over 100 villages in nine provinces (Anhui, Gansu, Guangdong, Henan, Hunan, Jiangsu, Jilin, Shanxi, and Sichuan) are matched with corresponding village-level data.[10] In each province, counties in the upper, middle, and lower income terciles were selected, from which a village was then randomly chosen. Subject to the limits of this stratification, the RCRE sample should reasonably capture both inter- and intraprovincial income variations. Depending on village size, between 40 and

[10] The complete RCRE survey covers over 22,000 households in 300 villages in 31 provinces and administrative regions. RCRE's complete national survey is 31% of the annual size of the NBS rural household survey. By agreement, we have obtained access to data from nine provinces, or roughly one-third of the RCRE survey.

120 households were randomly surveyed in each village. The survey spans the period 1986–99 and includes between 7,000 and 8,000 households per year.

RCRE originally intended a longitudinal survey, following the same households over time. While there is a significant panel dimension to the household sample, nearly one-third of households were lost to attrition during the period 1986–99, much of which is a product of village attrition that occurred during two 2-year gaps when RCRE was unable to conduct the survey in 1992 and 1994 because of funding difficulties. RCRE replaced lost villages by "comparable" villages in the same counties. Households lost through attrition were replaced (at least in principle) on the basis of random sampling.

The survey collected detailed household-level information on incomes and expenditures, education, labor supply, asset ownership, land holdings, savings, formal and informal access to credit, and remittances.[11] In common with the NBS *Rural Household Survey*, respondent households keep daily diaries of income and expenditure, and a resident administrator living in the county seat visits with households once a month to collect information from the diaries. The large number of households surveyed from each village and the lengthy span of the survey enable us to track the evolution of consumption, incomes, and inequality during a time of changing market access and development in rural China. Of particular importance for our purposes, we are able to track a panel of villages, even where there has been household attrition. This will allow us to maintain geographic comparability over the entire period.[12]

A variety of definitions are worth clarifying, and further details related to attrition issues are provided in appendix A. First, household membership is defined on the basis of residency and registration.[13] Second, income is calculated as the sum of net income (gross revenue less current expenditures) from agriculture, farming sidelines (e.g., animal husbandry and livestock), and family-run businesses plus wage income and transfers. We calculate the value of farm output that is not sold, and thus largely consumed (or stored) by the household,

[11] One shortcoming of the survey is the lack of individual-level information. However, we know the number of dependents and individuals working as well as the gender composition of household members.

[12] Eighty-two of the original 110 villages surveyed in 1987 were among the 103 villages surveyed in 1999.

[13] It includes individuals in the household with rural registration (*hukou*) plus a small number of individuals with nonrural registration but who live in the village full-time. This definition of household membership differs slightly from that of the World Bank's Living Standard Measurement Survey, which for other than the household head, bases membership on the actual number of months of residency in the house.

at market prices.[14] Household income is also gross of taxes and fees. Third, our measure of consumption includes nondurable goods expenditure plus an imputed flow of services from household durable goods and housing.[15]

We deflate all income and expenditure data into 1986 prices using the NBS rural consumer price index for each province. For some key results we explore the sensitivity to geographic differences in price levels. In those cases we spatially deflate using a cross-province consumer price index (CPI) deflator constructed by Brandt and Holz (2004), based on expenditure weights from the NBS rural household survey. The spatial CPI adjusts for systematic differences in price levels across provinces (at a point in time), because price levels and incomes are positively correlated, possibly exaggerating differences in living standards across regions. Finally, RCRE's sampling is not proportional to provincial population. For example, the number of households surveyed in Sichuan is nearly the same as that surveyed in Gansu, despite the fact that Sichuan has a rural population that is nearly five times larger. Therefore, we use provincial rural population (by year) to weight all calculations.[16]

In order to establish the robustness of our conclusions to various permutations of sample selection, we carried out our analyses on three different data sets. The first, or "full," sample includes every household (panel, attrited, and replacement) in each survey year. The second accounts for the fact that inequality measures may be sensitive to outliers (at both ends of the distribution), and this "trimmed-full" sample drops extreme outliers among households.[17] The third sample is a "balanced-panel," comprising the 4,352 households for which we have data for every year of the survey. As it turns out, our results are consistent across all three data samples. To minimize tables, we restrict our reported results to those from the "trimmed-full" data set.

[14] RCRE's surveys follow pre-1990 NBS conventions and value nonmarketed grain at quota prices. We follow an approach used by Chen and Ravallion (1996) to recalculate the value of nonmarketed grain at market prices. This is discussed in more detail in appendixes A and B.

[15] In order to convert the stock of durables into a flow of consumption services, we assume that current and past investments in housing are "consumed" over a 20-year period and that investments in durable goods are consumed over a period of 7 years. We also annually "inflate" the value of the stock of durables to reflect the increase in durable goods' prices over the period.

[16] Specifically, weight = [(Province Rural Population)/(Number of Households Sampled in Province)] × (Household Size).

[17] The lowest income households actually have negative incomes. These are typically households that have high gross incomes but also high business-related expenses. The problem of measurement error for these households is especially severe. We discuss our procedure for eliminating outliers in app. A, Sec. III. In each year, less than one-tenth of 1% of households were dropped from the panel.

TABLE 1
PER CAPITA INCOME AND CONSUMPTION: LEVELS AND GROWTH SELECTED YEARS, RCRE

	Spatial Deflator?	1987	1991	1995	1999	Implied Growth Rate
Income	No	578	551	772	712	.019
	Yes	567	538	760	699	.019
Consumption	No	410	402	548	508	.019
	Yes	402	392	541	497	.019
Observations		7,983	7,903	6,738	6,987	

Note. This table shows mean real per capita household income and consumption (in constant 1986 RMB yuan) for selected years. The implied growth rate is defined as the average annual compound growth rate that would turn 1987 incomes to 1999 levels. The spatially deflated rows adjust for regional price differences using the price deflator in Brandt and Holz (2004), described in app. A.

IV. Results

A. Income Distribution over Time

We begin by summarizing the evolution of average income and consumption per capita over the span of our sample in table 1. To keep the table manageable, we report results for four evenly spaced years—1987, 1991, 1995, and 1999—that reflect the patterns in the more complete sample.[18] Mean household per capita incomes were 578 RMB in 1987, the beginning of our selected years.[19] Average incomes dropped slightly through 1991 and rose sharply to 772 in 1995. The average annual growth rate over the 1991–95 period was an impressive 5.3%, reinforcing the optimism of the expansion that followed the retrenchment between 1989 and 1991. This growth, which was broad-based in the rural economy, proved to be short-lived, as average per capita incomes actually fell to 712 by 1999. A sharp reversal in farm prices and cropping incomes following the increase between 1993 and 1995 underlies much of this decline. Setting aside the cyclical variation in growth, the average rate of growth from the beginning to the end of the sample was 1.9%. An important question is whether the decline in average incomes was disproportionately borne by the poor, as this would certainly worsen the distributional consequences of rising inequality.

Our results for consumption closely mirror those for income, both in terms of the cyclical patterns and the implied growth rate over the entire period. Given that the only overlap in the two series is home-produced consumption, it is reassuring that these two otherwise independent measures of welfare track each other so closely (though this may be less assuring for those who believe that consumption should be much smoother than income). In levels, con-

[18] Results for tables 1 and 2 (our main results) for the complete sample of annual observations and for the different samples (panel and nonpanel) are available from the authors by request. Some of the information for other years is contained presented in figs. 1 and 2 below.

[19] RMB is the abbreviation for "Renminbi," the name of the Chinese currency. The unit of account is the yuan (equal to 1 RMB). In 1986, the exchange rate was 3.45 yuan per U.S. dollar.

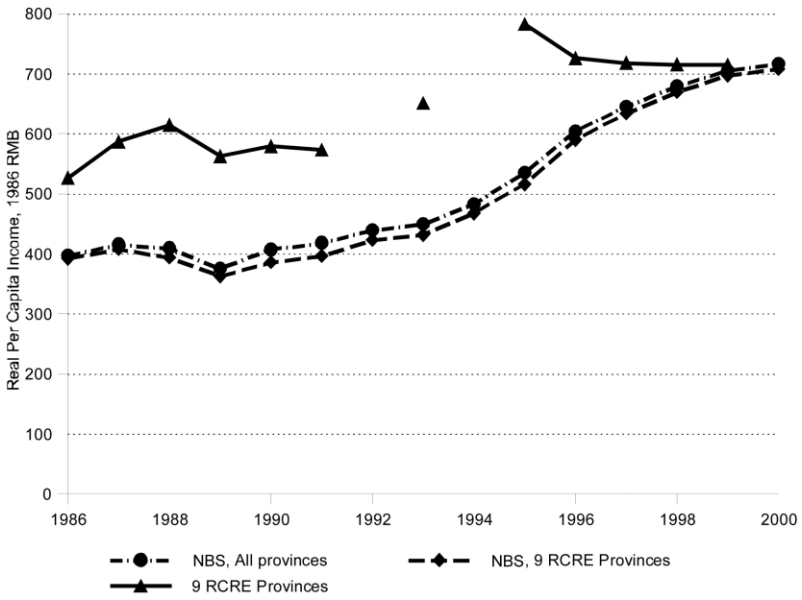

Figure 1. Comparing income trends in the RCRE and NBS surveys. This figure compares real per capita household income levels and trends in the RCRE surveys with corresponding numbers reported in the NBS yearbooks. The NBS results are shown for all provinces, as well as the same nine provinces in the RCRE sample. All figures are deflated to 1986 RMB yuan.

sumption is approximately three-quarters of income. Some of this gap probably reflects measurement error, but it also reflects genuinely high rates of savings and the fact that incomes are measured before deduction of taxes and other fees.

In table 1 we also show results for spatially deflated mean income and consumption. Spatial deflation makes no difference for this exercise. This is not surprising, as the aggregate numbers do not provide much scope for differential provincial price levels to affect the evolution of the average. The deflator is more likely to matter when we compare incomes across regions, as when we examine inequality.

How do the RCRE numbers compare to other data sources from China? The only other data set that spans this period is the nationally representative—and publicly unavailable—NBS rural household survey. In figures 1 and 2, we plot average incomes and consumption for each year of the RCRE survey along with the corresponding NBS rural averages reported in statistical yearbooks. We do not make this comparison assuming that the NBS data are the gold standard (nor are the RCRE data for that matter), but they are the only nationally representative data collected on an annual basis that could be used for this type of analysis.

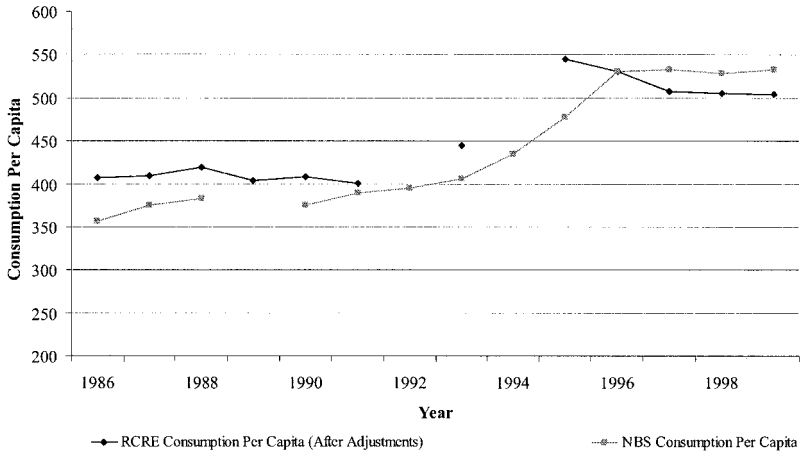

Figure 2. Comparing consumption trends from the NBS and RCRE surveys. This figure compares real per capita household consumption levels and trends in the RCRE surveys with corresponding numbers reported in the NBS yearbooks. The NBS results are shown for all provinces and reflect household expenditures on durable and nondurable consumer items. The RCRE consumption series values grain at market prices and uses a flow value durable and housing consumption. All figures are deflated to 1986 RMB yuan.

Up through 1995, the basic patterns in the RCRE and NBS income series are very similar, with the notable difference that the RCRE incomes are consistently much higher than those of the NBS. The higher relative incomes of the RCRE erode by 1999, when the mean incomes in the two surveys actually converge. Thus, the RCRE data show a flatter time series and a correspondingly lower growth rate. Despite the differences in the magnitude of the trend between 1987 and 1999, both series suggest that growth was fastest in the period 1991–95, with a significant attenuation (or decline) from 1997 to 1999. The consumption estimates, on the other hand, line up better in both levels and growth. The difference in the rate of growth of consumption between the two series over the period from 1987 to 1999 is less than half of that for incomes.

One possible source of difference is that the RCRE provinces are not nationally representative. Including only the RCRE subset of provinces in the NBS data (which come at the provincial level), means from the RCRE subset of provinces almost perfectly track national means. Alternatively, the two surveys may differ in terms of the kinds of households that are sampled. We do not have access to the NBS data set, but we can line up common indicators calculated from the 1996 agricultural census and RCRE surveys as a check. We find that 92.8% of RCRE households and 90.8% of households in the agricultural census had positive income from agriculture but that a slightly higher number of individuals per household were primarily engaged in non-

agricultural activities in the RCRE survey. Using the information from the 1990 census and 1996 agricultural census as benchmarks, it is difficult to argue that the RCRE survey oversampled agricultural households, but RCRE households were slightly larger than the nine-province agricultural census average by 0.45 individuals.[20]

In appendix B, we consider the consequences of significant differences in the calculation of income and consumption. Differences in the treatment of taxes and fees, the valuation of income from home production, and depreciation on fixed assets loom important. For years prior to 1990, these three factors are the source of slightly more than 60% of the difference, with most of this due to taxes and fees and the valuation of home production. It has been suggested (Ravallion and Chen 2004) that after 1990 NBS began to value in-kind components of income at market prices as we do over the entire period in our RCRE estimates. Still, significant differences persist. The gap in per capita income actually widens between the two series after 1990, peaking in 1995, before eventually disappearing altogether. In the case of the consumption series, valuing the in-kind component consistently between the two series also helps narrow the gap prior to 1990, but we observe similar differences in behavior to that for the income series after 1990.

We believe that continued differences in how in-kind income and consumption are actually being valued between the two surveys provide an explanation for some of the gap between 1990 and 1996. As detailed more carefully in appendix B, NBS adjusted in-kind valuation procedures three times during the 1990s, but only in a single year, 1997, did they use market prices. From 1991 to 1996 nonmarketed grain was valued at the average contract price (this was the average of two administratively determined prices, namely, the quota and above-quota price) if this price existed, otherwise enumerators were instructed to use market prices. The weighted average contract price in use between 1990 and 1996 was well below the market price, and this explains some of the difference between the two series. In 1997 agricultural commodities consumed in kind were valued at market prices; and then from 1998 onward, commodities (grain, meats, and other crops) not marketed by the households and consumed in kind were valued at 85%–90% of the market price, depending on the commodity.[21]

[20] Appendix A discusses this sampling issue in more detail, and a comparison between RCRE and agricultural census indicators is shown in app. A, table A2.

[21] Appendix B provides a more detailed discussion of in-kind valuation procedures and implications for trends in the NBS rural household survey. A translation of the relevant part of the document can be found on Giles's Web site (www.msu.edu/~gilesj) as well as excerpts from the original Chinese-language document.

TABLE 2
PER CAPITA INCOME AND CONSUMPTION INEQUALITY: VARIOUS MEASURES AND SELECTED YEARS, RCRE

	1987	1991	1995	1999
Income:				
Gini (not spatially deflated)	.32	.33	.33	.37
Gini (spatially deflated)	.29	.30	.30	.35
Other measures of inequality:				
Variance of logs	.57	.60	.59	.73
Atkinson (sensitivity = 2)	.28	.32	.33	.97
Atkinson (sensitivity = 1)	.16	.17	.17	.21
Atkinson (sensitivity = .5)	.08	.09	.09	.12
Percent below half cont. mean	.14	.17	.17	.22
Percent below 1987 half cont. mean	.16	.19	.06	.14
90th/10th split	4.06	3.98	3.93	5.24
Consumption:				
Gini (not spatially deflated)	.25	.27	.27	.31
Gini (spatially deflated)	.22	.24	.25	.29
Other measures of inequality:				
Variance of logs	.44	.47	.47	.55
Atkinson (sensitivity = 2)	.18	.20	.21	.26
Atkinson (sensitivity = 1)	.09	.11	.11	.14
Atkinson (sensitivity = .5)	.05	.06	.06	.08
Percent below half cont. mean	.08	.09	.10	.16
Percent below 1987 half cont. mean	.08	.10	.03	.08
90th/10th split	3.09	3.21	3.29	4.07

Note. This table provides various distributional summary statistics corresponding to the mean per capita income and consumption levels reported in table 1. We show (1) the Gini coefficient, repeated for spatially undeflated and spatially deflated levels; (2) the variance of log per capita income and consumption; (3) the Atkinson Index, calculated with three inequality aversion parameters (decreasing in aversion for 2.0, 1.0, and 0.5); (4) the proportion of households with incomes below one-half the contemporaneous mean income (i.e., the 50% of mean income that year); (5) the proportion of households below one-half the mean income level for 1987 (an approximation to a constant "poverty line"); and (6) the ratio of the 90th to the 10th percentiles.

As emphasized originally by Chen and Ravallion (1996), differences in the valuation of home-consumed grain can substantially affect estimates of income and consumption. Changes in the market versus quota price, or the formula used to apply them, will further bias any estimated trends. As we compare trends in the RCRE and NBS during the 1990s, we know from NBS documentation that in-kind valuation was not done at market prices, probably underestimating incomes in the early 1990s and exaggerating income growth. That said, there may yet remain differences in the RCRE and NBS sampling frames that we cannot detect on the basis of comparable observables. Since the NBS data are not publicly available, it is not possible to calibrate further the residual differences between average incomes in the two surveys at the beginning of our sample.

Changes in Inequality over Time

Table 2 provides measures of income and consumption inequality that highlight a variety of distributional characteristics. We begin with the Gini co-

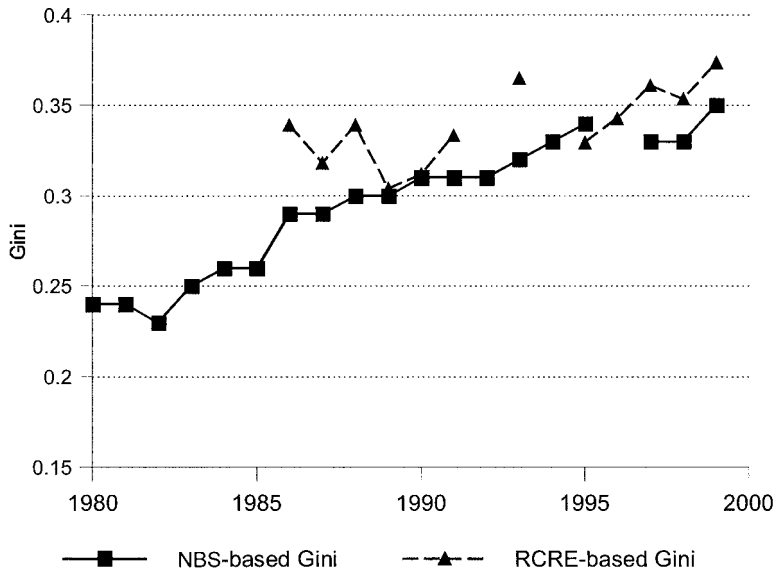

Figure 3. A comparison of inequality trends based on RCRE and NBS households surveys. This figure compares the Gini coefficients for household per capita income that we calculated using the RCRE, with Gini coefficients based on the national sample NBS data. The NBS results come from Bramall (2001).

efficient for income, arrayed in the first row. The Gini increased from 0.32 in 1987 to 0.37 in 1999, an increase of 0.05 or 16%. Is this increase economically significant? There are few benchmarks for comparison, though it is worth noting that inequality measures evolve slowly over time, and a 16% increase is large.[22] Of particular note, almost all of this increase was over the short period between 1995 and 1999. Combined with the decline in average incomes in the late 1990s, it should come as no surprise that concerns over inequality have intensified. In the second panel, we show the results for consumption, where the Gini rises from 0.25 to 0.31 over the complete sample. As is usually the case, the Gini for consumption is lower than income, but the trend and overall time pattern are basically the same. If anything, the increase in consumption inequality was slightly larger in percentage terms. Measuring welfare either way, it appears that inequality went up over this period, especially since 1995.

In figure 3 we place the RCRE results beside those from the NBS, where the NBS Gini's are drawn from Bramall (2001). The NBS numbers show inequality rising from 0.24 in 1980 to around 0.3 in 1990, slightly lower

[22] See Deininger and Squire (1996) for a discussion of the evolution of inequality over a broad range of countries.

than in the RCRE sample, and ending at 0.35 in 2000. As with the RCRE data, the NBS data thus show a Gini rising by about 0.05 points over the 1990s, though the increase is much smoother. Some of these differences, especially the sharp drop in inequality exhibited by the RCRE data between 1993 and 1995, can also be linked to differences in the valuation of the in-kind component of income. Other slight differences in the magnitudes of the Gini in particular years can potentially be attributed to differences in sampling, sample sizes, stratification, or other differences in the survey instrument; however, both the trends and magnitude of the increase in inequality are similar in the two surveys. Thus, while the NBS and RCRE data differ in detail, they paint a similar overall picture for the evolution of rural household welfare over the 1990s.

Returning to table 2, in the second row we apply the spatial price deflator to household income. As expected, the magnitude of the Gini drops. Higher income areas appear less well off once account is taken of the higher prices faced by consumers in these provinces. The magnitude of the drop is 0.02–0.03, but spatial deflation does not materially affect our conclusions about the overall trend in inequality: in other words, we cannot deflate away the increase in inequality and attribute it to widening gaps in regional nominal (versus real) incomes.

While the Gini shows an increase in inequality, is there any sensitivity of our conclusions to our particular choice of inequality measure? The next four rows of the table present inequality measures for the variance of log income and the Atkinson index evaluated at three different inequality aversion parameters (ranging from 2.0 to 0.5 in decreasing magnitude of inequality aversion). While these measures can be used for comparison with other studies, the main purpose they serve for us is to confirm the rise in inequality, especially between 1995 and 1999. Most worrying, the Atkinson index with high inequality aversion increases from 0.28 to 0.97 over this 4-year period! This suggests that the bottom part of the income distribution did especially badly.

A direct comparison of Lorenz curves yields a nonparametric comparison of changes in inequality, possibly avoiding the need to choose any of the inequality measures in table 2. We show the Lorenz curves for 1987, 1995, and 1999 in figure 4. Panel A shows the curves for the entire sample, while panel B provides a close-up view of the bottom part of the distribution (from panel A). The 1999 Lorenz curve lies outside that for 1987 over the entire distribution, suggesting that for any inequality measure we choose, 1999 will look more unequal. Unambiguous comparisons for other pairs of years are not possible since the Lorenz curves cross. However, for much of the distribution the 1999 curve also lies outside that for 1995, while the curves for 1987 and

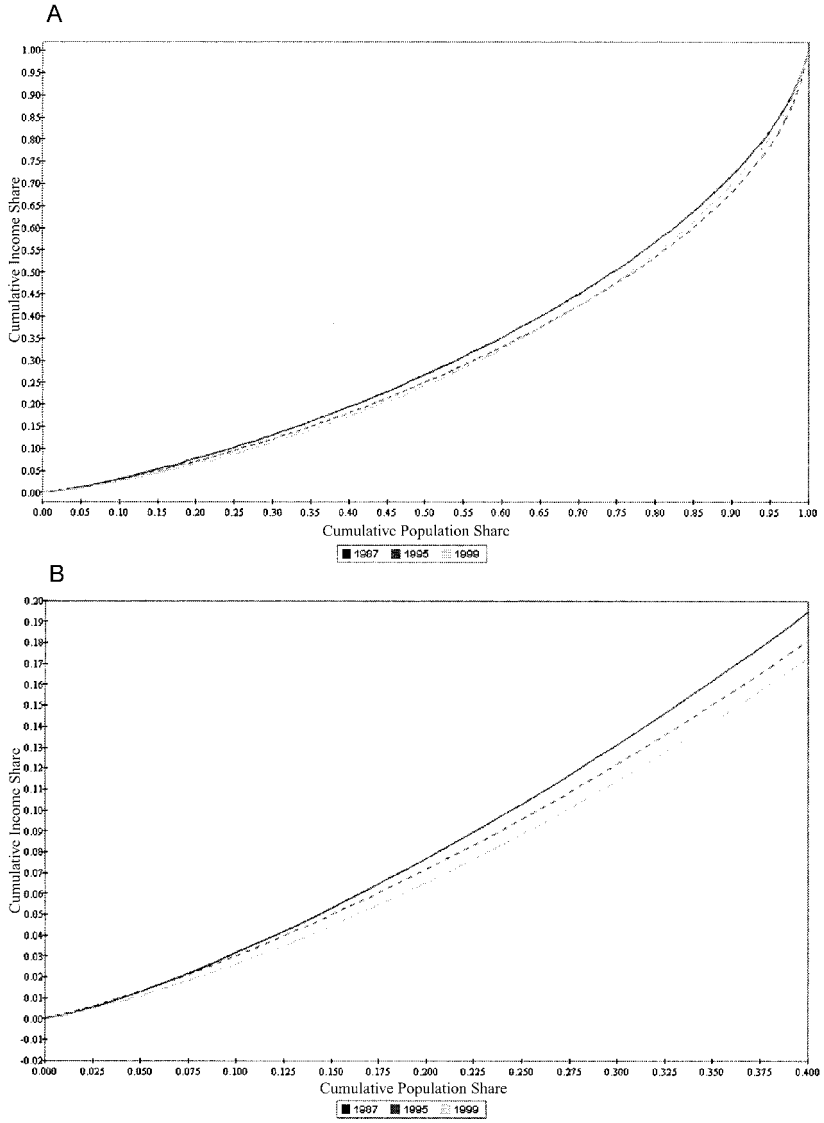

Figure 4. Lorenz curves for per capita income, selected years. *A*, The full picture. *B*, The bottom 40% (from panel *A*). This figure shows the Lorenz curves for the distribution of real household per capita income for selected years, using the RCRE survey data. Panel *A* shows the entire curves (with the reference 45-degree line), while panel *B* "magnifies" the curves, focusing on the poorest 40% of households. The 1999 curve lies beneath the 1987 and 1995 curves for much of the distribution, and the difference between the curves is statistically significant (see app. C, table C1).

1995 are almost indistinguishable, with a slight advantage at lower incomes for households in 1987.[23]

Changes in Relative and Absolute Poverty

Increases in inequality are less worrisome if incomes are increasing for the entire income distribution and the welfare of the poor is improving even as inequality is increasing. We next investigate changes in both relative and absolute inequality to understand how changes have occurred at the lower end of the income distribution. In the next two rows of table 2 we switch from overall measures of inequality to considerations of relative inequality between rich and poor, and poverty. First we present the proportion of households with incomes below half the contemporaneous mean.[24] Essentially, this is a "relative poverty line" set at 50% of the mean income for the year: this poverty line moves with average incomes, and while it does not tell us much about absolute poverty, it provides another useful way of characterizing the distribution of income.

The proportion of the sample below half the contemporary mean income increases from 16% in 1987 to 22% by 1999, with most of the jump occurring after 1995. These results suggest that the increase in inequality reflects a worsening of the relative position of low-income households: using this relative measure, the rich are getting richer and the poor are getting relatively poorer. A stronger point about absolute changes in poverty can be made when we keep the "poverty line" constant at half 1987 mean levels (in real terms), thus allowing us to track progress on the elimination of poverty with a constant benchmark. The poverty rate worsens between 1987 and 1991, which comes across as the worst year for the poor. The best year was 1995, with half as many people "poor" as in 1987. By 1999, however, the poverty rate (so measured) has returned to essentially the same level as 1987, doubling in just 4 years from the level in 1995. Despite an increase of average incomes by 25% between 1987 and 1999, the fraction of people below the 1987 "poverty line" is nearly the same.

Use of half the contemporary mean or any poverty line is somewhat arbitrary, and for this reason we pursue two additional exercises using the full distribution

[23] We use the Davidson and Duclos (2000) procedure implemented in the DAD software package (Duclos, Araar, and Fortin 2004) and test for significance of the difference in Lorenz curves at different points. We find that the 1999 Lorenz curve is significantly below the 1987 Lorenz curve along the entire distribution and that it is significantly below the 1995 curve for the lowest 90% of the distribution. A summary of results of the tests of significance is presented in app. C, table C1.

[24] See Jenkins (2000) for an example of using the "half-contemporaneous mean" in summarizing the evolution of the distribution of income in the United Kingdom.

of income per capita: we examine cumulative distribution functions and generalized Lorenz curves. Cumulative distribution functions (CDFs), shown for income per capita in 1987, 1995, and 1999 in figure 5, permit using any common poverty line applied to the three years to compare poverty rates. For example, if we chose the 1987 half-contemporaneous mean of 290 RMB per capita as our poverty line, we can recover the poverty rates from table 2: approximately 6% of people have incomes below 290 in 1995, compared to 14% in 1999 and 16% in 1987. The CDF for 1995 lies everywhere below the one for 1987, and so the distribution from 1995 first-order stochastically dominates the 1987 distribution. This implies that for any poverty line we might select, there will be less poverty in 1995 than in 1987. For other pairs of years, however, the CDFs cross, thereby complicating comparisons.

We use Davidson and Duclos's (2000) approach to estimate the critical poverty lines where CDFs cross and the point at which the stochastic dominance of one distribution over the other will switch. Comparing the 1999 distribution to the 1987 distribution, we find that for poverty lines above 226 RMB per capita the 1999 distribution dominates the 1987 distribution, and we would judge there to have been a fall in head count measures of poverty. For lower poverty lines, however, the 1987 distribution dominates, and we would find an actual increase in poverty. A critical value of 226 RMB lies well below the rural poverty line of 302 yuan RMB (in 1986 RMB) estimated by Ravallion and Chen (2004), and we may conclude that the number of individuals living in poverty has declined even as the living standards of the very poor may have worsened.[25]

Welfare Comparisons Using Generalized Lorenz Curves

A second representation of welfare that incorporates inequality is captured by the generalized Lorenz curves shown in figure 6. The Lorenz curves from figure 4 are multiplied by average annual per capita income. A Lorenz curve illustrates the share of the pie going to lower income percentiles, while the generalized Lorenz curve adjusts for the size of the pie.[26] In panel A we see that the curve for 1995 lies strictly above the other years, indicating that the distribution from 1995 generalized Lorenz (second-order stochastic) curves dominates those from 1987 and 1999. At the top end of the distribution, the 1999 distribution

[25] Ravallion and Chen (2004) estimate a poverty line of 850 RMB yuan per capita in 2002, which is equal to 302 RMB yuan per capita in 1986 values. This is actually quite close to the half 1987 mean income per capita value of 290 RMB yuan per capita. For this poverty line, the 1999 distribution dominates the 1987 distribution.

[26] The generalized Lorenz curve was introduced by Shorrocks (1983b). See also Foster and Shorrocks (1988), Atkinson (1987), and Deaton (1997) for more general discusions relating generalized Lorenz curves to stochastic dominance, welfare rankings, and poverty.

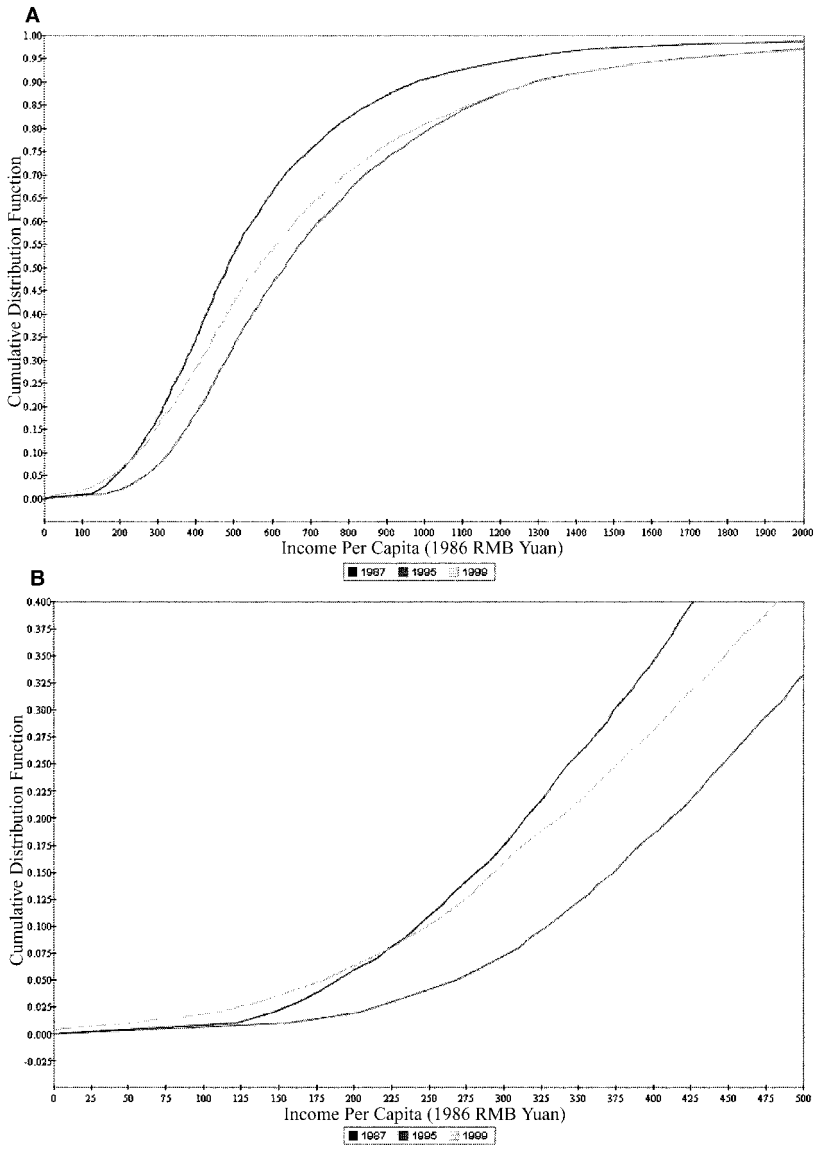

Figure 5. The cumulative distribution function (CDF) of per capita household income, selected years. *A*, The full picture. *B*, The bottom 40% (from panel *A*). This figure shows the CDF for the distribution of real household per capita income for selected years, using the RCRE survey data. Panel *A* shows the entire curves, while panel *B* "magnifies" the curves, focusing on the poorest 40% of households. The critical poverty line when comparing the 1999 and 1987 distributions is 226 RMB per capita with a standard error of 36.3. These figures and significance of critical poverty lines were calculated using DAD (Duclos et al. 2004).

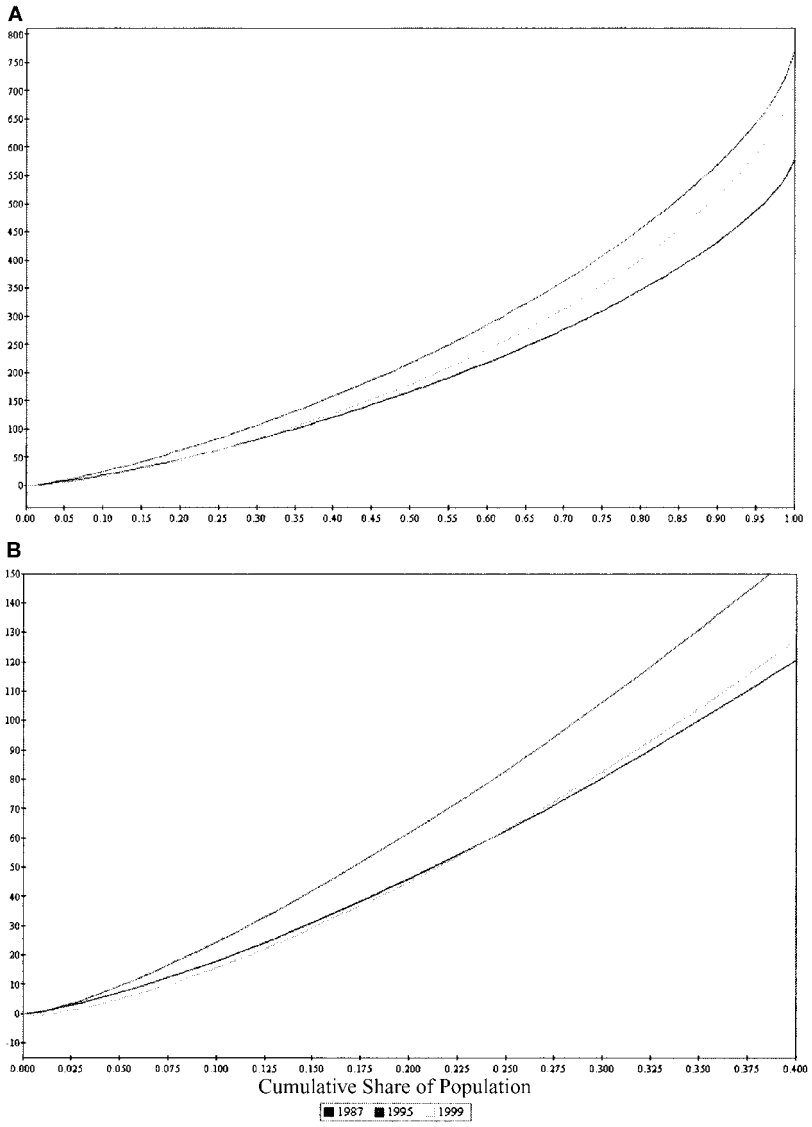

Figure 6. Generalized Lorenz curves of per capita household income, selected years. *A*, The full picture. *B*, The bottom 40% (based on panel *A*). This figure shows generalized Lorenz curves for distribution of real household per capita income for selected years, using the RCRE survey data. Panel *A* shows the entire curves, while panel *B* "magnifies" the curves, focusing on the poorest 40% of households. Tests of statistical difference between the curves are shown in app. C, table C2.

is second best, as richer households are better off than those from 1987. What is most striking, however, is the plight of the bottom quarter of the income distribution. The generalized Lorenz curves cross around 0.23, suggesting that the lower 23% were actually worse off in 1999 than in 1987. The difference in generalized Lorenz curves is statistically significant only for the bottom 5% of the population, however, suggesting that deterioration in living standards is only significant for the very poor.[27] On the other hand, the 1999 generalized Lorenz curve is unambiguously above the 1987 curve for only the upper 50% of the distribution. These results then are consistent with the bottom half being no better off, and the poorest 5% of households being worse off, after 12 years of economic growth.

Finally, the bottom panel of table 2 shows the matching results for consumption. As with the Ginis and mean incomes, the picture based on household per capita consumption is essentially the same as the one based on income.

Robustness of Results to Sampling and Data Issues: Evidence from Other Data Sources

Taken on their own, these results from the RCRE surveys may raise concerns about the evolution of the income distribution in recent years. But in the broader context of world income inequality, a Gini rising from 0.32 to 0.37 over such a dynamic period may not seem so dramatic, at least in proportion to the concern expressed. Can the 0.37 Gini for 1999 be taken as a reliable upper bound for inequality in rural China? As we already saw in figure 3, the RCRE results line up with the NBS's in 1999. However, as shown in table 3, this conclusion is premature, as there are other data sets that show higher levels of inequality and simultaneously point to potential weaknesses in both the NBS and RCRE survey designs.

For select years, we compare mean incomes, the composition of income, and inequality using the RCRE and two other household surveys. In table 3 we report a comparison of tabulations of data from the fourth wave of the China Health and Nutrition Survey (CHNS) that covers the calendar year 1997, the RCRE survey for 1997, similar RCRE data for 1999, and a collaborative household-level survey carried out in 2000 covering 1,200 households in six provinces (the CCAP 2000 survey).[28] We break income down into that from farming, farm sidelines (forestry, livestock, and fisheries), wage income, family-run businesses, and a residual category, "other," which is largely composed of private and public transfers. We report total per capita income,

[27] See app. C, table C2, for tests of the difference between generalized Lorenz curves.

[28] The 2000 survey was a collaborative effort involving Bai Nansheng (formerly of the RCRE), Loren Brandt, Scott Rozelle (University of California, Davis), and Zhang Linxiu (Chinese Center for Agricultural Policy, or CCAP). See app. A for further details.

TABLE 3

COMPARISON OF SURVEYS: LEVELS OF INCOME AND INEQUALITY

	CHNS		RCRE				CCAP			
	1997		1997		1999		2000a		2000b	
	Mean	% not 0	Mean	% not 0	Mean	% not 0	Mean	% not 0	Mean	% not 0
Income, by source:										
Agriculture	816	78.7	624	96.1	507	94.8	606	90.4	607	90.9
Livestock	92	50.0	212	77.8	182	75.2	211	75.2	211	76.4
Wages	764	37.3	748	64.9	851	68.0	892	62.4	893	62.5
Other	386	51.8	148	86.0	150	80.3	155	52.6	155	34.0
Family business	418	21.9	522	54.5	494	50.1	796	29.7	534	28.2
Total income	2,477	97.9	2,255	100.0	2,184	99.9	2,667		2,370	
Inequality: Gini	.43		.36		.37		.50		.44	

Note. This table compares levels, composition, and inequality of per capita income for three data sets (surveys) at similar (though not identical) points in time. All reported values are in nominal (undeflated) terms. For each data set, we report mean values of per capita income by source as well as the percentage of households with nonzero observations ("% not 0"). The three primary data sets are China Health and Nutrition Survey (CHNS) for 1997, the RCRE for 1997 and 1999, and the Chinese Centre for Agricultural Policy (CCAP) survey for 2000. Results from the CCAP survey are shown for 2000a, a data set that includes all households with positive income, and for 2000b, based on the same sample but excluding the top 1% of per capita income households. Further details are available in app. A.

per capita income by source, and the percentage of households in each of the surveys that report nonzero income from each source. For the CCAP 2000 survey we report two tabulations: (1) based on the full sample and (2) based on the full sample, but dropping the top 1% of households (in terms of per capita income). The second tabulation is performed in order to address the possibility that the NBS and RCRE surveys may be undersampling higher income households in rural areas. One caveat to our comparisons is that there is only limited overlap across the surveys in terms of the provinces sampled. As we saw in figure 1, the RCRE provinces are not peculiar compared to national averages, but there will still be limits to comparability of the other surveys. First note the comparison of the RCRE with the CHNS for 1997. The structure of income (mean incomes by source) is similar in the two surveys, with slightly higher (by less than 10%) income in the CHNS. To some extent, this reflects a slightly higher fraction of suburban households in the CHNS rural sample (notice the slightly smaller proportion of farmers in the CHNS). Most notable, however, is the higher level of inequality reflected in the CHNS, with a Gini of 0.43.

The most striking gap is between the 1999 RCRE and the CCAP for 2000, with average incomes higher by 22% in the CCAP survey. Rural income growth was relatively flat between 1999 and 2000, so the difference cannot be attributed to economic growth.[29] Moreover, the rural CPI was falling over this period, and so the differences in real income are even slightly more pronounced than the nominal figures that we report. Much higher reported income from family-run businesses in the CCAP 2000 data appears to be the source of most of the difference. Mean per capita income from family businesses was 796 in the 2000 survey, but only 494 for the 1999 RCRE. This difference represents 62.5% of the gap in mean incomes between the two surveys.

A comparison of inequality measures based on these surveys reveals an even more substantial difference in the two surveys, with the CCAP survey suggesting a Gini of 0.50, which is much higher than that for the RCRE (or other surveys). But a comparison with a slightly "trimmed" version of the CCAP 2000 survey identifies the likely source of the problem. Official surveys often exclude the richest households that often earn substantial incomes from family-run businesses.[30]

[29] The NBS rural household survey data show an increase in nominal per capita net incomes from 2,210 to 2,254, or an increase of less than 2%.

[30] Refusal rates are likely to be higher among households with a high opportunity cost of time, and neither RCRE nor NBS report refusal rates. By design, both surveys are also less successful in accurately estimating household incomes from family-run businesses. For example, both the RCRE and CHNS surveys simply ask respondents for total revenue and expenditures from family

The most revealing comparison is made between the trimmed and untrimmed versions of the CCAP sample. Dropping the top 1% of households lowers mean household per capita income from 2,667 to 2,370, or 11%, and bringing it closer in line with the RCRE and CHNS estimates. Incomes by source also line up very well between the RCRE and CCAP trimmed samples. Almost all of the drop in mean incomes, and resulting improvement in correspondence of the surveys, comes from the decline in average incomes from family-run businesses. The Gini coefficient also falls considerably, from 0.50 to 0.44, more in line with the inequality reflected in the CHNS. If we drop the top 3%, the Gini falls slightly more, to 0.42.

This exercise highlights several important points. First, measured inequality in rural China is sensitive to the top tail of the income distribution. To the extent that the RCRE (or NBS) surveys miss the very richest households (possibly because they are not "representative") overall inequality will be understated.[31] Second, poor measurement of family-run business income alone can lead to a significant misrepresentation of the level of inequality. The CHNS and CCAP surveys thus suggest that NBS and RCRE-based estimates of the level of inequality are too low (possibly by as much as 0.10 Gini points). To the extent that family-run businesses have been increasing in importance over the reform period, the RCRE and NBS likely understate the upward trend in inequality. The results in table 2 and figure 3 probably provide a lower bound of the extent to which inequality has risen. Combined with the results from RCRE for roughly the lower 90% of the income distribution, we are confident in concluding that there has been some stagnation of rural welfare, and perhaps erosion, over the latter half of the 1990s.

B. Decompositions by Geography

The role of widening regional income differences and their contribution to increasing inequality is a common theme in the literature on inequality in

businesses, and while enumerated, we believe that the distinction between fixed and variable costs is often lost in the enumeration process. In general, this is among the most difficult sources of income to enumerate accurately and, thus, often the noisiest (Vijverberg and Mead 2000). It is also believed to drive rural inequality, so underestimation of this income source may significantly lead to understatement of the level of inequality. Considerable effort was placed in the CCAP 2000 survey on minimizing problems stemming from inaccurate reporting of nonfarm business income, both in terms of random sampling of households within villages and a more careful enumeration of the balance sheets of family businesses.

[31] The record-keeping requirements of the RCRE and NBS surveys also suggest that they likely underenumerate low-income, illiterate households. The Gini is slightly less sensitive to the lower tail. In the context of the CCAP survey, dropping the lower 10% of households reduces the Gini by only several points.

China.[32] Rising disparities between localities, especially provinces (inland versus coastal, for example), are often seen as the most important source of the rising income differences, as some provinces are better situated to take advantage of market liberalization and new off-farm opportunities. At the outset of the reforms, spatial differences may have also been present due to differences in per capita land endowments, access to urban markets, and the level of development of commune and brigade-run enterprises. With decline in importance of restrictions on migration created by China's residential registration system (the *hukou* system) and opening up of markets for migrant labor, however, we expect a decline in the importance of region in overall inequality. Our sample, which includes the rapidly growing coastal provinces of Guangdong and Jiangsu and slower growth interior provinces of Sichuan and Gansu, seems reasonably well suited to look for these differences and their trends.

There are a number of approaches one can take in decomposing inequality across regions. Unfortunately, the Gini coefficient is not readily (or neatly) decomposed. Gustafsson and Li (2002) report spatial decompositions for the decomposable mean log difference and Theil inequality indices. We adopt a simpler strategy, decomposing the variance of log income inequality index. This entails estimating the following regression:

$$\ln y_i = D_L'\gamma + u_i, \tag{1}$$

where D_L is a vector of dummy variables indicating the location of individual i.[33] The R-squared from this regression indicates the proportion of the variation (or variance) of $\ln y_i$ that is explained by the location dummies. The remainder is the (within-location) residual variance of log income and a measure of the degree to which household income cannot be explained by the average income of its neighbors.

Table 4 reports the results of this exercise for income and consumption per capita with location defined at three levels of aggregation. We estimate the equation above separately using region-, province-, and village-level dummies and also distinguish between spatially deflated and undeflated household income and consumption per capita.[34] For comparability, we also report the results from decomposing the Theil by region.

In the first row of table 4, we see the proportion of inequality as measured

[32] Kanbur and Zhang (1999) provide an excellent overview of the literature on regional inequality, highlighting inland versus coastal, and urban versus rural dimensions. See also Gustafsson and Li (2002).

[33] For example, for a provincial-level decomposition, this will be a set of provincial dummies.

[34] We define three regions as West (Gansu, Shanxi, and Sichuan), Central (Anhui, Henan, and Hunan), and East (Jilin, Jiangsu, and Guangdong).

TABLE 4

CONTRIBUTION OF LOCATION TO INCOME AND CONSUMPTION INEQUALITY: RCRE, SELECTED YEARS

	1987	1991	1995	1999
	Contribution to Variance			
Dependent variable ln (income per capita):				
Without spatial deflator:				
Contribution of region	.186	.162	.154	.120
Contribution of province	.237	.218	.183	.153
Contribution of village	.500	.466	.413	.424
With spatial deflator:				
Contribution of region	.069	.063	.062	.047
Contribution of province	.133	.105	.085	.077
Contribution of village	.431	.389	.344	.373
Dependent variable ln (consumption per capita):				
Without spatial deflator:				
Contribution of region	.190	.184	.162	.181
Contribution of province	.278	.246	.189	.231
Contribution of village	.560	.529	.507	.525
With spatial deflator:				
Contribution of region	.051	.063	.064	.085
Contribution of province	.137	.102	.083	.117
Contribution of village	.474	.439	.442	.454
	Contribution to Theil-T Index			
Dependent variable income per capita:				
Without spatial deflator:				
Contribution of region	.043	.054	.048	.069
Contribution of province	.163	.122	.085	.122
Contribution of village	.490	.452	.441	.456
With spatial deflator:				
Contribution of region	.080	.090	.063	.065
Contribution of province	.114	.126	.078	.092
Contribution of village	.398	.402	.374	.401
Dependent variable consumption per capita:				
Without spatial deflator:				
Contribution of region	.218	.220	.182	.186
Contribution of province	.283	.273	.213	.241
Contribution of village	.539	.539	.507	.508
With spatial deflator:				
Contribution of region	.047	.068	.055	.062
Contribution of province	.137	.114	.080	.106
Contribution of village	.446	.433	.424	.419

Note. This table shows the fraction of variation of real log per capita income (and consumption) attributed to location. This is simply the R^2 from a regression of log per capita income on a set of location dummies. The decompositions are reported with or without the income variable spatially deflated. The effect of location is reported at three levels of aggregation: (1) the village (103 villages); (2) province (nine provinces, as described in app. A); and (3) region, defined as West (Gansu, Shanxi, and Sichuan), Central (Anhui, Henan, and Hunan), and East (Guangdong, Jiangsu, and Jilin).

by the variance of log income explained by "region" declined from 0.19 in 1987 to 0.15 in 1995, and to 0.12 in 1999. Using the Theil, the results are very similar. There are limits to comparability, but we also discuss our results in light of Gustafsson and Li's (2002). Their results for the proportion of mean log difference (MLD) (which is similar to the variance of logs) explained by region are 0.12 for 1987 and 0.27 in 1995. Our results thus differ in both level and trend. Some of this difference may be driven by differences in measurement of income.[35] Sampling might also be an issue. The provinces used in the two studies are not the same; however, the regions are similarly defined. More significantly, the RCRE subsample comprises a panel of counties (82 of the 103 villages in the subsample in 1999 had been in the sample since 1987), while the data used by Gustafsson and Li are drawn from a different set of counties in the two periods. This introduces a dimension of noncomparability across time periods for studies based on the Khan and Riskin (1998) subsample of NBS data, the potential bias of which is difficult to assess ex ante.

Turning to province-level results, the difference in patterns between studies is also apparent, though less pronounced. Here, we find that the proportion of income inequality explained by province declined from 0.24 to 0.18 between 1987 and 1995, and further to 0.15 by 1999. By contrast, Gustafsson and Li (2002) start out with a similar proportion explained by province in 1987 (0.24), but their proportion rises to 0.32 by 1995.

In figure 7, we explore possible differences between the RCRE and NBS household surveys (the basis of the samples used by Gustafsson and Li [2002]). Our objective is to compare the amount of province-level inequality present in the two data sets. For the RCRE, we calculate overall Gini coefficients assuming everyone in a province earns the same income (provincial mean income). The national Gini is constructed on the basis of these interprovincial income differences (weighted by rural population). For the NBS, we use provincial mean incomes from Statistical Yearbooks of China to calculate similar national Gini coefficients, weighting mean incomes by provincial rural population and effectively attributing everyone in a province the same income. This procedure sets within-province inequality to zero and calculates the implied Gini arising from differences in provincial mean incomes alone. The basic pattern in figure 7 is similar to the one we saw in figure 2: the level of inequality is almost the same in the two surveys, but the slope is slightly steeper in the NBS. The NBS shows an especially steep increase in spatial

[35] For example, the NBS-based estimates may not fully reflect the effect of the increase in agricultural prices in the mid-1990s on incomes, inequality, and spatial differences.

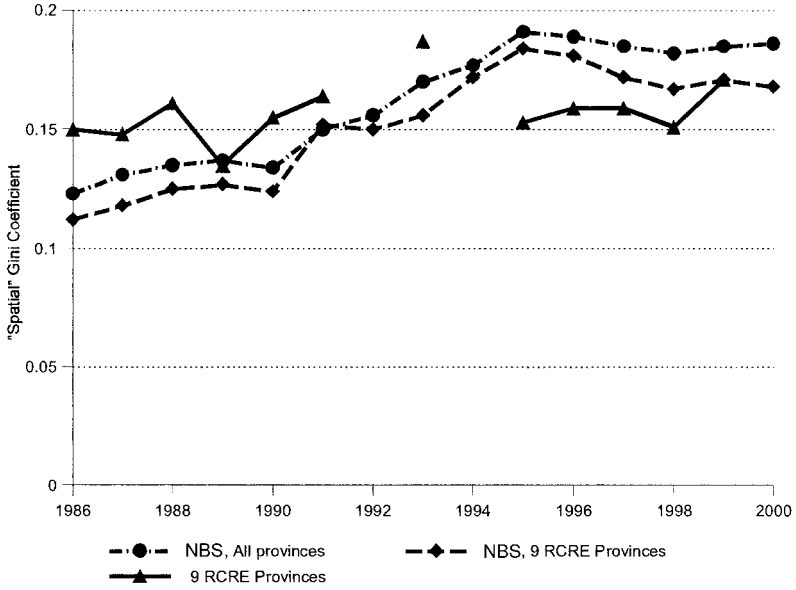

Figure 7. Comparing interprovincial inequality in the RCRE and NBS surveys. This figure compares the amount of implied interprovincial inequality in the RCRE and NBS surveys. Interprovincial inequality is calculated on the basis of attributing to everyone in a given province the mean per capita income and calculating the implied Gini with provincial population weights. This "simulation" exercise is conducted for all provinces in China using NBS-reported mean provincial incomes; for the subset of provinces covered by the RCRE sample (but using NBS mean incomes); and for the RCRE sample.

inequality between 1988 and 1995 (Gustafsson and Li's sample years), compared to the RCRE. Both series show a significant flattening of this trend since 1995, with the NBS showing actual declines in spatial inequality from the mid-1990s to 2000. In figure 7, we also note that the RCRE subset of provinces had a slightly lower level of spatial inequality than the full national sample but not enough to render the RCRE provinces unrepresentative. In summary, while there is disagreement about the initial level of spatial inequality (as we saw in fig. 2) and the size of the increase up through 1995, the two data sources essentially agree on the magnitude of spatial inequality and, especially, the relative decline since the mid-1990s.

Returning to the RCRE data source, our results suggest that the role of provincial rural income differences has declined over time. Another way to illustrate this point is to look at the behavior of provincial mean incomes and within-province inequality. In figure 8 we plot average provincial growth rates against the initial (1987) level of income. There is no obvious pattern here, and results are sensitive to a single observation, notably Guangdong. Excluding Guangdong, this figure suggests some degree of income convergence, with poorer provinces growing more rapidly over the 1987–99 period. This is

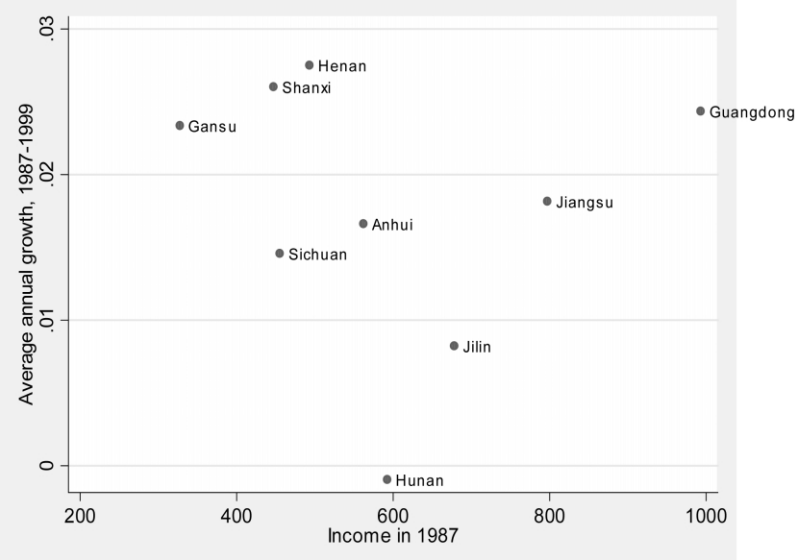

Figure 8. Growth in per capita income by initial per capita income, RCRE provinces. This figure arrays average annual growth rates for incomes by province (based on RCRE data) by "initial income," that is, mean provincial income in 1987.

consistent with a narrowing of interprovincial inequality. But of course, we cannot just exclude Guangdong in painting the complete picture. Nonetheless, figure 8 provides no evidence that provincial income levels were diverging.

In figure 9 we plot within-province Ginis from 1999 against the provincial Gini for 1987. We also show a 45-degree line in order to benchmark the inequality levels in the two years. Here we see that inequality rose in all provinces except Gansu, and in Jilin it rose to over 0.40. Again, this is consistent with the decompositions that show that within-province inequality became more important between 1987 and 1999. Furthermore, excluding Jilin, we see that the plot suggests that provinces with lower inequality in 1987 had higher increases of inequality to 1999, implying convergence of Ginis across provinces.

Next, in the third row of table 4 we show the fraction of inequality explained by village. Here, we see the proportion fall from approximately 0.50 in 1987 to 0.40 in 1999: most of the inequality in our sample occurs within, as opposed to across, villages.[36] An obvious question is whether 0.50 represents a half full

[36] Gustaffson and Li (2002) do not report results for village, as their finest unit of location is the county. With this caveat, they find that 40% of inequality was across counties in 1988, rising to 50% by 1995. While trends are opposite ours, the basic magnitude is similar. Also note that the NBS data show a decline in interprovincial inequality from 1995 to 1999 (fig. 6), so more recent numbers from the NBS might line up more closely.

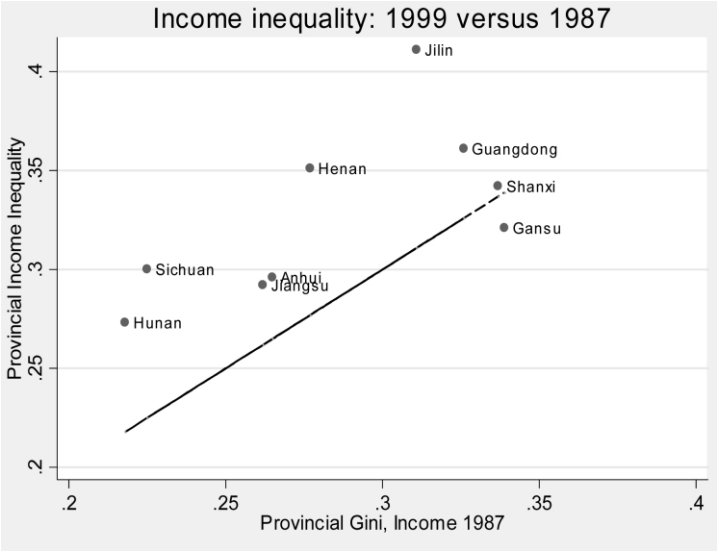

Figure 9. Changes in provincial inequality, 1987–99. This figure plots the provincial Gini for per capita income in 1999 (calculated with the RCRE data) versus the provincial Gini for 1987. The 45-degree line serves as a reference, whereby points lying above the 45-degree line correspond to increases in provincial inequality between 1987 and 1999.

or half empty glass, in terms of the role played by geography. Nothing in our conclusions diminishes the fact that location is an important (perhaps the single most important) determinant of household income. Furthermore, even a diminution of the role played by village is consistent with persistence of low incomes within and across villages, such as one would expect with geographic poverty traps as identified by Jalan and Ravallion (2002). Instead, we view our results as pointing to the significant role played by within-village differences in incomes as a contributing factor in overall inequality and correspondingly draw attention to those factors that generate inequality within villages.

Finally, in table 4 we evaluate the impact of spatial deflation on the decompositions. Accounting for interprovincial price differences cuts the share explained by province or region in half. This suggests that the interprovincial income gaps overstate the differences in the standards of living across provinces; however, the absolute bias is declining over time. Spatial deflation has a much smaller effect on the role of village in the decompositions, but this is to be expected since the spatial deflator uses provincial level prices. Nonetheless, by 1999, between-village differences accounting for spatial price differences are the source of only about a third of overall inequality.

In figures 10, 11, and 12 we explore the evolution of village-level inequality.

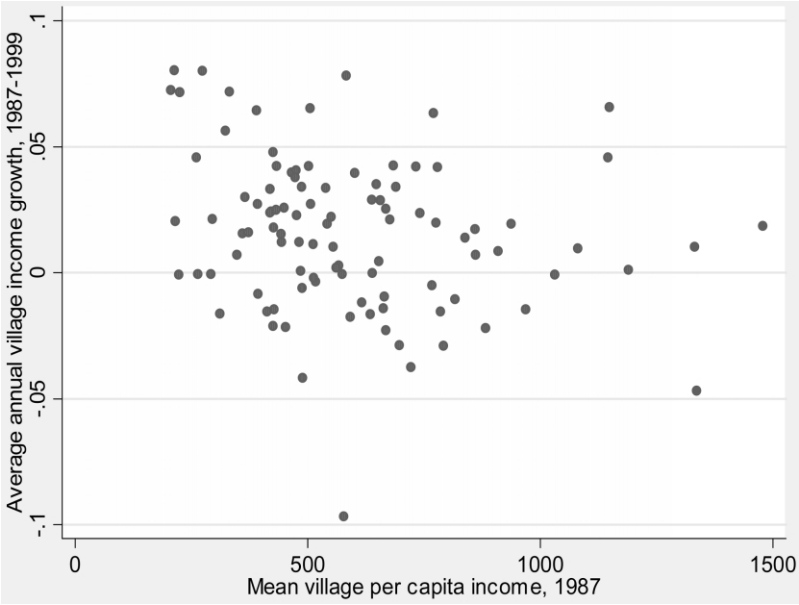

Figure 10. Within-village growth versus initial village income. This figure arrays average annual growth rates for incomes by village (based on RCRE data) by "initial income," that is, mean village income in 1987. The points in this figure can be used to estimate a "convergence" regression. Such a regression yields: Growth = 0.14 − 0.02lnY87 (t = 2.9).

First, in figure 10 we show that there is evidence of convergence of income levels across villages (poorer villages tended to grow more rapidly between 1987 and 1999). All else equal, this convergence reduces the role of village in explaining inequality (as we saw in table 4). Figure 11 shows histograms for the village-level Ginis, clearly showing the shift upward of within-village inequality. This shift is also readily apparent in figure 12, where we see that a majority of villages experienced increases of their income Ginis, with a considerable fraction experiencing increases over 0.10, though village sample sizes are small enough to warrant a caution on placing too much stock on a single Gini. As with the provincial-level inequality measures, it appears that there is convergence of inequality levels, whereby low-inequality villages experienced greater increases in inequality.

The broad conclusion from the spatial income-inequality decompositions is that (1) no more than half (and probably less) of total inequality is driven by income differences between villages and (2) the role of geography has remained relatively constant and likely declined in recent years. There is little evidence to suggest that widening spatial differentials account for a disproportionate share of the increase in rural inequality. In terms of understanding the sources of inequality, this should serve to turn more attention toward local village

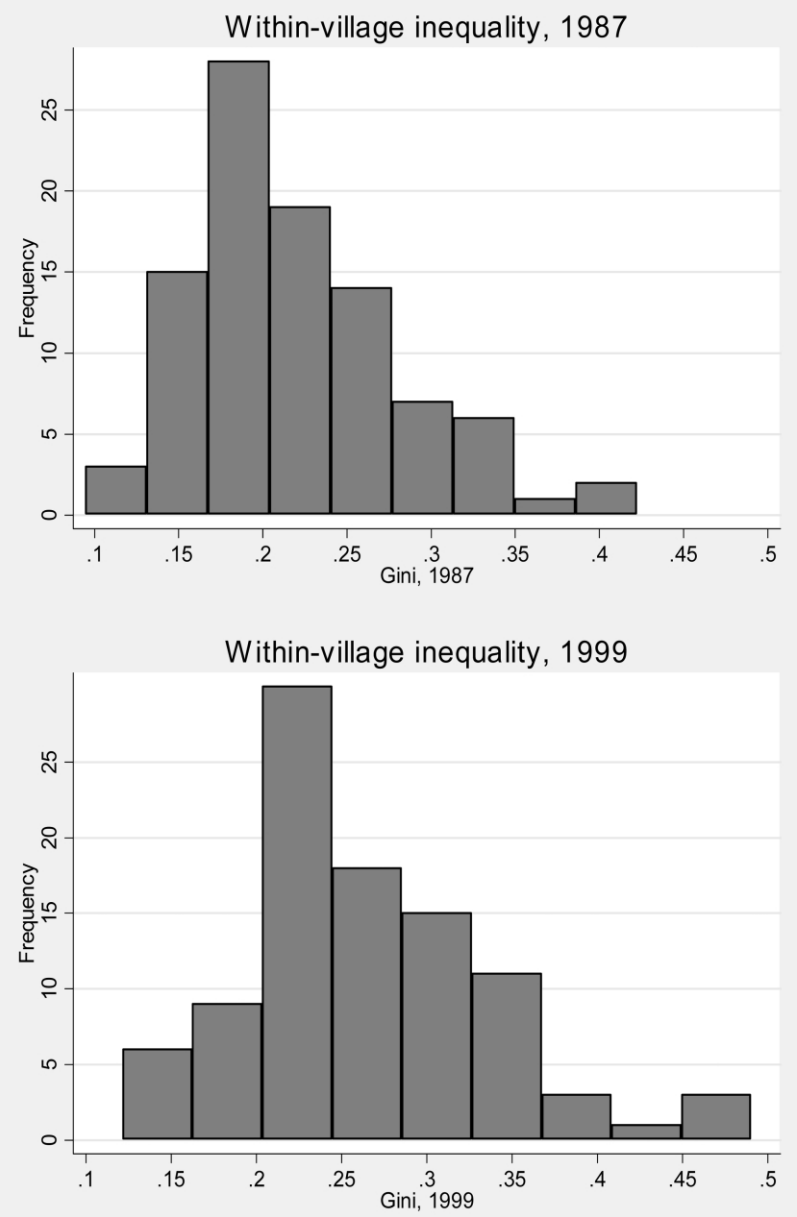

Figure 11. Evolution of within-village income inequality. These histograms report the frequency of various magnitudes of village-level Gini coefficients for 1987 and 1999.

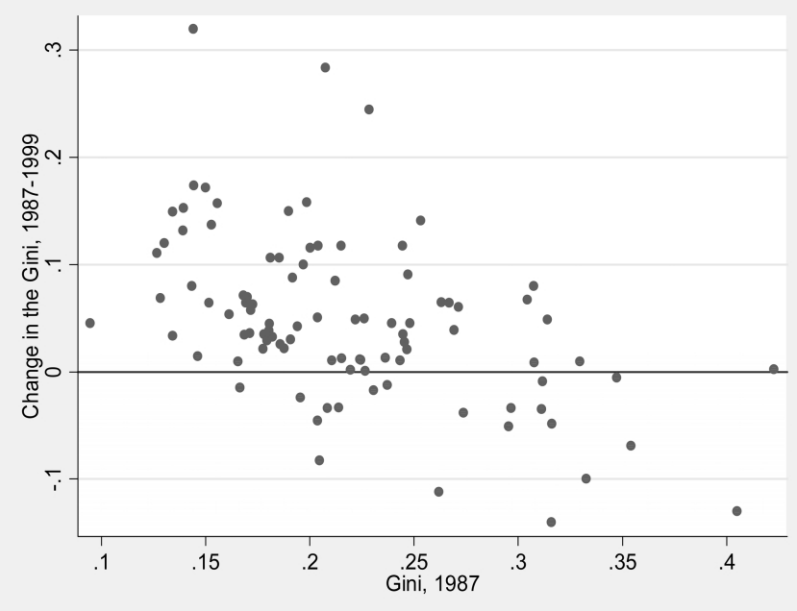

Figure 12. Changes in village inequality versus initial inequality. This figure plots the changes in village-level inequality between 1987 and 1999 (based on the village-level Gini) versus the initial level of inequality (the Gini in 1987). This figure illustrates "convergence" of inequality levels across villages.

and township institutions related to governance and investment in public goods and local variation in the distribution of endowments (like skills, education, and land) and their returns.

C. Decompositions by Source

Why has inequality gone up within villages? Answering this question requires an understanding of the evolution of institutions across villages that map household endowments into family income and is a significant research enterprise in itself. Our more limited objective here is to sketch some of the correlates of within-village inequality, particularly those related to the composition of household income. Previous studies have emphasized the role of nonfarm income in contributing to rising inequality.[37] We can use the RCRE data to confirm the role of nonfarm income, to evaluate finer details of the composition of subcomponents of nonfarm income, and most importantly, to gauge trends in the role played by income composition in explaining increases in overall inequality.

The key tools in our analysis are descriptive statistics of the structure of income and Shorrocks (1982, 1983a) decompositions. The Shorrocks decom-

[37] See Benjamin et al. (2002) for a survey of these studies.

position tells us the proportion of total inequality that can be attributed to inequality of income source k. It is a purely descriptive tool, and there are limits to the extent that one can attribute a causal interpretation to the coefficients. For example, it is difficult to use the results to simulate the impact of an increase in the inequality of a particular income source on total inequality, without further specifying the nature of the increase in inequality of that income source. However, as we shall see, even within the limits of interpretation the decompositions are illuminating.

As an outline of the procedure, consider a decomposition of the mean of household income, based on household i's income, y_i:

$$y_i = \sum_{k=1}^{K} y_{ik}, \tag{2}$$

which is the sum of K subcomponents of income y_{ik}. Clearly, mean household income can be written:

$$\bar{Y} = \bar{Y}_1 + \bar{Y}_2 + \ldots + \bar{Y}_k. \tag{3}$$

A 1% increase in mean income from source k will lead to a W_k proportionate increase in $\bar{Y}$, where W_k is the share of income from source k. Decomposition of the sources of mean income is thus straightforward, and decomposition of inequality is designed analogously. We wish to estimate S_k, the proportion of inequality attributable to the inequality of income source k:

$$I(Y) = \sum_{k=1}^{K} S_k I(Y_k), \tag{4}$$

where $I(Y)$ is the index of inequality for total income Y, and $I(Y_k)$ is the index of inequality for income source k. Shorrocks showed that for any additively decomposable measure of inequality, S_k is estimated by:

$$\hat{S}_k = \frac{\mathrm{Cov}(y_{ik}, y_i)}{\mathrm{Var}(y_i)}. \tag{5}$$

So S_k captures the degree to which income source k is correlated with total income. In this sense, it measures the degree to which particular income sources are earned by the rich or poor. If an income source is earned primarily by the rich, then the decomposition will attribute a larger share of total income inequality to inequality of income earned from that source. How can we interpret these S_k? One benchmark is zero: if an income source is negatively correlated with total income, then it is earned disproportionately by the poor, and no inequality (indeed a negative share) of total income is correlated with

that income source. Presumably, marginal increases of inequality of that source of income (maintaining the same correlation with total income) would further reduce overall inequality. Very few sources of income will have negative S_k. Another helpful benchmark is the mean share of income from that source, or W_k. If $S_k > W_k$, then inequality of income source k contributes more to inequality than it does to mean income, $\bar{Y}$, which we denote as a disproportionate effect on inequality. In other words, if income from family businesses composes 10% of average income, but 20% of inequality, we will conclude that family business income has a disproportionate effect on inequality.

As a matter of computation, $\hat{S}_k$ can be estimated by the following regression:

$$y_{ik} = \beta_{0k} + \beta_{1k} y_i + u_{ik}, \tag{6}$$

as $\beta_{1k} = \hat{S}_k$. This regression presentation also aids in interpretation: all we are estimating is the correlation of a particular source of income y_{ik} with total income, y_i. Once we broaden our objective to the estimation of this correlation, we can also recognize the possible impact of measurement error: overestimates of income from a particular source will lead to an overstatement of the correlation with total income, and β_{1k} will be overstated. This overstatement for income source k will spill over to the other β_{1k}, leading to an underestimate of their contribution. One simple way to address this possibility is to estimate the regression by two-stage least squares, using another indicator for total income as an instrument for y_i. An obvious candidate is total household consumption, which should not suffer from the same type of measurement error as the y_{ik}. Aside from the instrumental variables interpretation, this procedure can be viewed as exploring the sensitivity of our conclusions concerning the correlation of income from a particular source with whether a household is rich or poor, to alternative definitions of rich and poor, based on income or consumption. As a final refinement on the Shorrocks procedure, we compare decompositions with and without village dummies. Inclusion of village dummies allows us to decompose within-village inequality and to net out the possible effect of variation of income sources across villages, and this links to cross-village inequality.

We begin with a description of mean incomes by source, reported in table 5. In 1987, agricultural income (crop income) composed 40% of total income. The largest subcomponent was grain income, at 30% of total household income. Adding income from agricultural sidelines (fish, forestry, and especially livestock) raises the broadly defined share of agriculture to 53%, over half of family income. Family business, mostly in commerce and services, composed

TABLE 5
THE COMPOSITION OF INCOME IN 1987 COMPARED TO 1999 (IN 1986 RMB)

	1987			1999			
	Mean	Share	%>0	Mean	Share	%>0	Growth
Total income	578	1.000	1.000	714	1.000	.999	.018
Agricultural income	229	.397	.981	158	.222	.942	−.031
Grain income	175	.303	.978	113	.158	.926	−.037
Cash crop income	46	.080	.812	30	.042	.564	−.036
Fruits, tea, and dates	8	.014	.248	15	.022	.257	.056
Agricultural sidelines	74	.129	.955	68	.095	.764	−.007
Forest products	17	.029	.333	10	.014	.185	−.043
Livestock	54	.093	.950	50	.069	.746	−.007
Aquaculture	4	.007	.125	8	.012	.058	.062
Family businesses	91	.157	.616	162	.227	.501	.048
Household industry	27	.048	.135	44	.061	.073	.039
Construction	6	.010	.072	11	.016	.046	.056
Transportation	17	.030	.065	26	.037	.076	.036
Commerce, service, and trade	25	.042	.126	57	.079	.172	.070
Other family business income	16	.027	.395	24	.034	.251	.036
Wage income	145	.251	.711	276	.387	.680	.054
Local wage income	85	.147	.452	79	.111	.257	−.006
Temporary migrant	46	.080	.390	175	.245	.505	.111
Local government employment	14	.024	.069	22	.031	.049	.038
Family transfers	29	.050	.525	34	.048	.495	.016
Government transfers	4	.008	.651	6	.008	.708	.022
Other income	5	.009	.138	8	.012	.110	.040

Note. This table compares the composition of income in 1987 to that in 1999. Real per capita income is shown for detailed subcategories of income, along with the share of income ("Share") accounted for and the proportion of households with nonzero income for that source. The last column reports the implied annual growth rate of income for that source. Note that wage income is divided between "local" wage income and "temporary migrant." Temporary migrant employment includes both commuters returning home on weekends and longer-term temporary migrants and in most cases involves employment outside of the township.

16% of income, while wage income was the second largest overall component, at 25%. Most wage income was earned locally, within the village.

This structure of income changed dramatically by 1999. Most notable is the absolute decline in the amount of income from agriculture. Grain income alone dropped from 175 yuan in 1987 to 113 in 1999.[38] This 35% decline can be attributed primarily to the collapse in grain prices described earlier. Other sources of agricultural income—with the minor exception of income from fruit—declined to the extent that the overall share of farm income declined to 32% of total income, a drop of 20 percentage points from 1987. What is especially important to note is that the decline in this share is not due merely to increasing relative importance of nonfarm income but to an

[38] While the 1987–99 trend suggests a straight decline of grain income from 175 to 113 RMB per capita, grain income peaked at 257 RMB per capita in 1995 and declined rapidly between 1995 and 1998.

absolute decline in levels of agricultural income. Moving down the column, we see improvements in income from family businesses, in absolute terms from 91 to 162 yuan, and from 16% to 23% as a share of total income. But the largest improvements in family income came from wage earnings, especially wages earned by temporary migrants. The wage earnings of temporary migrants include household members still resident in the village, but who commute outside the village to work and return on weekends, as well as wage earnings brought home by locally registered household members who work outside the village for a substantial portion of the year. The RCRE survey does not permit a further disaggregation. Clearly, however, locally earned wages have become less important in both relative and absolute terms, while employment opportunities outside the village and accessed through migration have become a more important source of labor earnings.

The Shorrocks decompositions are presented in table 6. As a general summary, controls for location rarely matter, indicating that composition of income matters within villages much the same way as across villages. Furthermore, the ordinary least squares (OLS) and two-stage least squares (2SLS) estimates generally agree, at least in terms of broad conclusions, and we focus on 2SLS results in our discussion.

For 1987, we find that agricultural income, while disequalizing, contributed less to overall inequality than its share of total income (19% vs. approximately 40%). The same applied to agricultural sidelines, so that only 21% of total inequality was attributed to inequality of agricultural income, even while this source accounted for 53% of total income. Nonfarm family businesses contributed most to inequality compared to their share of income (27% compared to about 16%), followed by wage income (31.6% compared to 25.1%). Within the wage category, local wages were relatively disequalizing, while wages from employment outside the village were relatively equalizing.

The results for 1999 are significantly different and even more different than the change in average composition would suggest. First note that inequality of agricultural income contributed only 3.5% of overall income inequality. Even adding livestock and other sidelines, the overall contribution of farming income to inequality was 6.3%. It would seem that to the extent that the machinery of redistribution (village-controlled land reallocations, for example) is directed toward minimizing inequality of farm income, it is misdirected. Inequality of nonfarm family business income contributes more to inequality in 1999 than 1987, though this is not surprising given its increased importance as a source of income. Perhaps the most striking result of the decompositions is the large share—47.5%, or almost half—of total inequality attributed to wage earnings. Local wage earnings, while they have declined in magnitude,

TABLE 6
SHORROCKS DECOMPOSITIONS RCRE, 1987 AND 1999

	1987				1999			
	Share (1)	OLS (2)	2SLS (3)	2SLS (4)	Share (5)	OLS (6)	2SLS (7)	2SLS (8)
Village dummies?	NA	No	No	Yes	NA	No	No	Yes
Agricultural income	.397	.126	.190*	.171	.222	.045	.035*	.040
Grain income	.303	.050	.090*	.099	.158	.002	.004	.021
Cash crop income	.080	.050	.075*	.061	.042	.026	.022	.008
Fruits, tea, and dates	.014	.026	.026	.011	.022	.017	.009*	.011
Agricultural sidelines	.129	.074	.106*	.104	.095	.061	.028*	.038
Forest products	.029	.009	.025*	.033	.014	−.002	−.006*	.007
Livestock	.093	.047	.062*	.064	.069	.036	.013*	.019
Aquaculture	.007	.018	.019	.007	.012	.027	.021	.012
Family businesses	.157	.345	.232	.268	.227	.438	.391*	.409
Household industry	.048	.140	.092*	.115	.061	.213	.169*	.175
Construction	.010	−.004	.000	−.002	.016	.015	.011	.002
Transportation	.030	.077	.027	.020	.037	.062	.031*	.017
Commerce, service, and trade	.042	.109	.093	.121	.079	.122	.154*	.203
Other family business income	.027	.023	.020	.015	.034	.026	.026	.012
Wage income	.251	.373	.376	.316	.387	.400	.475*	.401
Local wage income	.147	.298	.310	.270	.111	.150	.204	.170
Temporary migrant	.080	.062	.044	.009	.245	.206	.214	.133
Local government employment	.024	.013	.023*	.037	.031	.045	.057*	.098
Family transfers	.050	.076	.090	.122	.048	.043	.056*	.094
Government transfers	.008	.000	.000	.008	.008	.002	.004	.009
Other income	.009	.007	.005*	.010	.012	.011	.010	.009

Note. The table shows Shorrocks decompositions, described in th text. Household per capita income by source is regressed on total per capita income. Columns 2 and 6 show OLS coefficients of income per capita, and cols. 3 and 7 show the same coefficients, but with income per capita instrumented by consumption per capita, as a "control" for measurement error in income. Columns 4 and 8 show 2SLS results with village dummy variables added to control for the possible geographic differences of income composition. For reference, the share of income by source is reported (and is the same as in table 5).
* Asterisks indicate where the OLS and 2SLS coefficients are significantly different (using a standard Hausman test) and thus that the 2SLS coefficients are to be preferred.

are relatively unequally distributed and disproportionately earned by higher income households. Inequality of wage earnings from temporary migrants outside the village explains 21.4% of overall inequality, but this is actually lower than its share of total income. To this extent, access to these wage opportunities is relatively equalizing. Note also that this is one example where controls for village dummies make some difference, as the within-village contribution (13.3%) is less than the total contribution, reflecting spatially uneven development of labor markets for temporary migrants.

Taken together, these decompositions highlight two important sources of inequality, especially when we compare 1999 to 1987. First is the sharp decline of the relatively equalizing source of income from farming. Second is the relative increase in disequalizing income from nonfarm family businesses and the failure of nonfarm labor markets to provide income opportunities for low income

households that offset the collapse of agricultural income. Past emphasis on the role of nonfarm income as a source of inequality was only partially correct: these results suggest that given the recent trajectory of farm income, efforts to improve the rural distribution of income should be placed on improving access to nonagricultural employment for low income households. Increasing agricultural incomes—at least in an equalizing way—is unlikely to improve overall income distribution, if for no other reason than that agricultural incomes are only weakly associated with overall income, and they are also very low.

V. Discussion and Conclusions

There is certainly a risk of oversimplification in attempting to summarize our key findings. After all, the core underlying data are based on household surveys with about 8,000 observations per year for 12 years, and such measures as the Gini coefficient are summaries themselves that obscure the complexities of income distributions. That said, our analysis points to an uneven, but long-run increase in inequality in rural China, and our estimates may actually underestimate the magnitude of the increase. Especially worrisome is the deterioration in performance in the last half of the 1990s, which left as much as half of the population not much better off in 1999 than 12 years earlier, and the bottom 5% worse off.

The most obvious next question is why this deterioration has occurred: have economic reforms failed? Have market reforms created an economy that disproportionately rewards winners and heavily penalizes losers? Our results provide preliminary answers to these questions. First, we rule out geography as the most important factor for understanding the dispersion of incomes: at any point in time, more than half (and as much as two-thirds) of inequality is due to inequality between neighbors within a village, not differences in income between rich and poor villages. Furthermore, we find that the importance of spatial income differences at the regional, provincial, and village level is declining over time. If most inequality is within villages, then this should turn our attention to determinants of within-village inequality, such as village-level institutions, market development, and the distribution of household endowments. An important avenue for future research is to document the joint evolution of village incomes and the distribution of village income, including a careful assessment of causal linkages between village growth and inequality, as well as other correlates of village-level growth and inequality. For example, very little is known about the role of education and the potential interaction of human capital with market development and access to nonfarm opportu-

nities.[39] Efforts to design appropriate social safety nets and to improve local tax policy in rural areas need to be informed by a better understanding of the ways in which local institutions and markets influence prospects for reductions in poverty and inequality and improve the growth prospects of the local economy.

Second (like previous researchers), we confirm that nonagricultural incomes are an important source of inequality. Indeed, to the extent that studies use NBS-like data (including the RCRE), both the level and trend of this source of inequality may be understated. But it would be a mistake to conclude that runaway income growth in nonfarm income drives the winner-loser divergence in rural areas. Certainly, inequality driven by households at the very top of the income distribution is associated with lucrative family businesses. However, access to nonagricultural employment—possibly in other people's family businesses and, in particular, employment outside the home county and accessed through migration—seems to be relatively equalizing. Rising inequality and falling incomes at the lower end of the distribution are driven by inability of poorer households to earn income from nonagricultural sources.

This conclusion is emphasized by our third key finding that the failure for living standards to improve since 1995 for as much as half of the rural population in our sample is driven by falling agricultural incomes. Given that output has not generally fallen, most of the decline in incomes can be attributed to sharply falling crop prices. An important area of future research thus concerns the determinants of farm prices. Are the low prices in the last half of the 1990s a transitory shock, reflecting temporary global market conditions? This may be the case, as crop prices have shown some recovery beginning in 2003. Or are they low more permanently, because improvements in farm productivity and entry into the World Trade Organization have changed the terms of trade between agricultural and nonagricultural goods within China? If crop prices are likely to be low (though possibly fluctuating) in the near future, then this raises a number of difficult policy questions. Almost all rural income-support policies are based on guaranteeing households access to land on an approximately per capita basis, through village land allocation. While this provides households a means to feed themselves, when crop prices are low (absolutely and relative to nonagricultural prices), the real value of this income support is quite low. Whatever the possible merits of this in-kind transfer for mini-

[39] Benjamin et al. (2002) show exploratory results that suggest that the combination of rising education levels and the development of nonfarm employment opportunities can reduce inequality and, furthermore, that the effect of the distribution of education on overall income distribution depends on local development of markets.

mizing poverty, it has obvious limitations as a redistribution mechanism. With low returns in agriculture, a land policy that attempts to equalize farm incomes will have only a weak impact on overall inequality, given the small and declining share of income earned in agriculture.

Finally, our results show that before the big picture can be fully understood, there are a number of critical data and measurement issues to be confronted. An important starting point would be improved access to NBS household survey data, so that richer cross-time and cross-space comparisons can be made. Although our results on the stagnation of income growth and poverty reduction in the late 1990s are broadly consistent with the NBS-based findings of Ravallion and Chen (2004), significant differences in results from the two studies remain. Unfortunately, lack of access to the NBS data makes it difficult to understand all of the factors underlying these differences (e.g., sampling issues or in-kind income valuation). As good as the RCRE data are, it would be helpful to broaden participation in the evaluation of poverty and inequality policy by opening up the NBS to more users. In addition, our comparison with other surveys shows the importance of measuring income from family businesses in understanding overall inequality, especially at the top end of the distribution. Combined with insights from other (more comprehensive) studies, a richer set of questions should be included in the NBS surveys to track this important source of income.[40] Even with their current limitations, however, results based on the RCRE survey should raise some alarm at the fate of the poor in rural China, both for their own sake and for the sustainability of future reforms.

Appendix A

Data

1. RCRE Village Locations: Province and Region

The data for the analyses of this study come from nine provinces of the Research Center for Rural Economy village and household surveys. Basic information on sampling within province and region is provided in table A1. We follow the literature in grouping provinces into regions. On average, RCRE surveyed households in 30 villages in both the western and the eastern region and in 45 villages of the central region. Management of the survey was delegated to provincial offices, which made decisions regarding within-province sampling rates. In each province, equal numbers of poor, medium, and rich counties were selected, from each of which a village of average socioeconomic status

[40] See Vijverberg and Mead (2000), for example.

TABLE A1
BASIC STATISTICS ON SAMPLE BY PROVINCE AND REGION

	Year						
Region/Province	1987	1989	1991	1993	1995	1997	1999
Western region:							
Number of villages	32	34	34	34	31	30	30
Average households/village	424.2	452.6	525.0	471.2	493.5	503.4	508.9
Average sampled households/village	56.1	56.4	55.7	53.0	55.7	55.3	55.1
Central region:							
Number of villages	48	48	49	44	44	44	44
Average households/village	313.4	332.0	353.2	367.0	405.3	408.2	403.6
Average sampled households/village	72.2	72.1	72.5	63.3	63.7	63.9	63.4
Eastern region:							
Number of villages	31	30	30	29	26	31	29
Average households/village	455.7	451.9	475.7	508.0	457.3	461.1	503.4
Average sampled households/village	87.6	80.6	82.1	72.7	74.8	70.8	75.2
Overall total:							
Number of villages	111	112	113	107	101	105	103
Average households/village	385.1	400.7	437.4	438.3	445.7	451.0	462.0
Average sampled households/village	71.9	69.6	70.0	62.5	64.1	63.5	64.2

Note. Province-by-province annual information on attrition is available by request.

was surveyed. At the village level, between 10% and 20% of households, or roughly 30–130 households, were then randomly selected.[41]

II. Attrition of Households and Villages from the RCRE Surveys

RCRE first fully implemented the national survey in 1986, doubling the number of villages in 1987. Since then, there has been relatively little change in the number of sampled villages. Attrition of villages from the survey has occurred, however, primarily for two reasons. First, RCRE's mandate is to use the survey to study agricultural production and factors influencing changes in agricultural productivity. Over the period from 1986 to 1999 four villages in Jiangsu and two in Guangdong were dropped and replaced because they were no longer engaged in agricultural activities. Second, attrition has also occurred as a result of disagreements between county or village leaders and provincial administrators of the survey. Of the 103 villages in the survey at the end of 1999, 82 have been in the survey since 1987. A significant share of village

[41] Our sample originally included Zhejiang province. An examination of county gross value of output and mean rural per capita income revealed that a disproportionate number of the surveyed counties in the province were from the upper third of the distribution. Thus, they were unlikely to represent well the distribution of income within Zhejiang, leading to biased estimates of inequality across regions. Sampling of counties in other provinces appears to be consistent with RCRE's guidelines.

attrition occurred during gaps when the survey was not conducted in 1992 and 1994. Much less change in villages (and households) occurred during the periods without gaps from 1987 to 1991 and 1995 to 1999, for which 98 and 97 villages, respectively, were in the sample for all 5 years. In principle, dropped villages were to be replaced by a representative village in the same county.

Attrition has also occurred at the household level (a detailed table is available upon request) and averages roughly 5% per year. Considerably more attrition came during the 2-year gaps, and it is largely associated with the loss of entire villages. Our estimate of attrition is also conservatively high. Households with the same household identifier in two successive years, but with significant differences in demographic structure, characteristics of housing, or economic activities, were treated as separate households. In these cases, we treat the year t household as a new observation and consider the year $t - 1$ household to have dropped from the sample. For the entire period between 1987 and 1999, we have a full panel of 4,352 households. For the two subperiods, namely, 1987–91 and 1995–99, panel size is 6,691 and 5,796 households, respectively.

III. The Sample Used in Our Analyses

In the sample used in our analyses, we trim extreme outliers from the data set because we suspect coding errors or errors in which fixed investment is inappropriately coded as an operating cost in the household budget. To identify potential outliers, we first calculate median income and consumption per capita in each village for each year. We then drop households if the absolute value of the difference between household reported income per capita and village median income per capita for the year is greater than five times village median income. We apply the same criteria to household consumption per capita. In each year, less than one-tenth of 1% of households were dropped under these criteria; altogether, 382 observations were dropped over the 1987–99 period.

Our "full trimmed" sample is thus this trimmed household sample, with all available observations, including panel, attrited, and replacement households. We also performed the analysis with the panel households only ("trimmed panel"), which would be most sensitive to possible biases introduced by attrition, and our results did not differ significantly.

IV. How Representative Is the RCRE Household Survey?

The subsample of the RCRE household survey covers between 7,000 and 8,000 households across 100 villages in nine provinces. Given that the Ministry of Agriculture, which is presumably interested in agricultural production, carries out the survey, one might worry that agricultural households

are oversampled and that this might lead to considerable bias. In order to consider this potential source of bias, we compare publicly available information from abstracts of the 1990 population census and the 1996 agricultural census with RCRE data in table A2. With respect to comparisons with the population census in 1990, there are no indices that are directly comparable for 1990 with either the RCRE or NBS rural household surveys. RCRE was not yet asking individual laborers to identify their primary activity. The population census for 1990 reports that of the population living in rural counties and holding rural registration, 92% reported agriculture as their primary activity. Yet this is not directly comparable to the RCRE and NBS surveys, because both of these surveys sample households based on residence and registration, not registration alone. Average household size in the population census is smaller by nearly one individual, but from the publicly available abstracts, we are grouping urban households of rural counties together with rural households. Since these urban households were subject to much tighter family-planning restrictions, their inclusion will reduce average family size.

A better sense of sampling can be achieved using the 1996 agricultural census abstract (NBS 1997). The NBS rural household survey samples from the same population covered in the 1996 agricultural census.[42] Both include all households that live and work in rural areas, including households with official rural registration and households living in the rural area for more than a year but lacking rural registration. Both include productive and nonproductive households and all households in rural areas regardless of whether productive activity is in agriculture or nonagricultural activities. We present information on seven indices that can be calculated in a comparable fashion from the 1996 agricultural census and the 1996 RCRE household survey. In addition, we also include average household size and the number of individuals of working age per household from summary statistics of the NBS rural household survey for 1996.

From these rough comparisons, it is difficult to argue that the RCRE surveys systematically oversample agricultural households, but it does appear that average household size is slightly higher for the RCRE survey. At the same time, however, average household sizes in the RCRE and NBS household surveys are nearly identical, suggesting either differences in definition of household membership between these surveys and the census or the possibility that both of these surveys have a slight bias toward surveying larger households with more working-age laborers per household.

[42] See NBS (1997) for documentation of who is surveyed in the agricultural census, and NBS (2000) for documentation of the NBS rural household survey sample frame.

TABLE A2

COMPARISON OF RCRE NINE-PROVINCE SAMPLE TO CENSUS INFORMATION AND NBS RURAL HOUSEHOLD SURVEY

	A. 1990 Population Census, 1990 RCRE Household Survey, 1990 NBS Rural Household Survey			
	1990 Population Census (Rural Counties)	1990 Population Census (Nine RCRE Provinces)	1990 Nine-Province RCRE Sample	1990 Rural Household Survey
Share of households that:				
Specialize in agriculture*	…	…	.640	…
Specialize in nonagricultural activities†	…	…	.135	…
Have positive income from agriculture	…	…	.960	…
Share of rural registered laborers with:				
Agriculture as main activity‡	.920	…	…	…
Nonagriculture as main activity‡	.080	…	…	…
Average household size§	3.99	4.00	4.93	4.84
Average laborers per household	…	…	2.85	2.92
	B. 1996 Agricultural Census, 1996 RCRE Household Survey, 1996 NBS Rural Household Survey			
	1996 National Agricultural Census	1996 Agricultural Census (Nine RCRE Provinces)	1996 RCRE (Nine-Province Survey)	1996 NBS Rural Household Survey
Share of households that:				
Are purely agricultural	.593	.574	.538	…
Are purely nonagricultural	.097	.092	.072	…
Have positive income from agriculture	.903	.908	.928	…
Average household size	4.086	4.051	4.495	4.420
Average laborers/household with:				
Agriculture as main activity	1.988	1.966	1.969	…
Nonagriculture as main activity	.638	.662	.756	…
Average laborers/household	2.626	2.629	2.725	2.840

Note. The 1996 NBS rural household survey samples are from the same population covered in the 1996 agricultural census; see NBS (1997) for discussion of who is surveyed in the agricultural census and NBS (2000) for discussion of the NBS rural household survey sample frame. Both include all households that live and work in rural areas, including both households with official rural registration, and households living in the rural area for more than a year but lacking rural registration. Both include productive and nonproductive households, and all households in rural areas regardless of whether productive activity is in agriculture or nonagricultural activities.

* Defined here as more than 50% of working days in agriculture.

† Defined here as less than 20% of household working days in agriculture.

‡ Information on main activities comes from 1990 population census (NBS 1993) summary tables on activities of rural registered residents from rural counties.

§ Note that in the census household size and labor force information can be broken down by rural counties, but individuals living in the (urban) county seat and county towns are grouped together with rural households, so these averages are not perfectly comparable with the RCRE household survey and NBS rural household survey.

While we note that 2% more households from the RCRE nine provinces report positive income from agriculture, 3.6% fewer households report that they are purely agricultural than in the agricultural census. Further, the average number of laborers per household reporting agriculture as their primary activity was nearly identical across the RCRE surveys and the census.

Finally, while we are unable to break down primary activity for the NBS rural household survey we do know that the share of income from agriculture was consistently higher in the NBS survey than in the RCRE survey. In 1997, for example, the share of net per capita rural income from cropping in the NBS data is 45.1% compared to 27.7% in the RCRE data. This was the one year that NBS valued nonmarketed commodities at market prices, and this income result is inconsistent with more pronounced oversampling of agricultural households by RCRE.

Recall that when comparing the RCRE survey with the CHNS and CCAP surveys above in Section IV.A we suggested that it is likely that both the RCRE and NBS surveys tend to undersample high income nonagricultural households. We believe that this likely source of potential bias affects both the NBS and RCRE surveys. Neither NBS nor RCRE report refusal rates for participation in their surveys, and given the onerous amount of work necessary to maintain daily diaries, we believe it reasonable that households with a higher opportunity cost of time opt out of the survey. The 2000 CCAP survey was conducted with great attention to keep refusal rates down and, not surprisingly, seems to include the upper tail of a distribution that is missing from the NBS and RCRE household surveys. When the wealthiest 1% of households from the CCAP is dropped, the estimated Gini coefficient falls considerably, and when the top 3% is dropped, the Gini approaches the lower values found in the RCRE and NBS surveys.

Given that the RCRE survey has a significant panel component of households that have been in the survey since 1986, one may also expect that the aging of the panel households may make RCRE somewhat less representative over time. The extent of this bias, however, is very difficult to gauge without full access to both datasets.

V. Issues in the Calculation of Household Income and Consumption per Capita

Grain crops remain an important component of household production, yet in the RCRE survey grain produced for own consumption or stored is valued at prices reflecting the quota price rather than the market price. Up through the mid-1990s, quota prices were well below market prices. For this reason, income from grain production and consumption out of home production are both likely biased downward. To deal with this problem in our analyses, we revalue

the household's nonmarketed grain (and grain consumption out of own production) at average village market prices.

A. Definition of Income and Consumption per Capita

Household income is the sum of income from all household-managed activities (farming, agricultural sidelines, and nonagricultural activities), local wage employment, migrant remittances, formal transfers from the village, subsidies from higher levels of government, and informal transfers from friends or family (but excluding borrowing). Consumption is calculated as the sum of expenditures on food and nondurable goods purchased during the year, the value of home-produced goods consumed, the value of the flow of services from the household's stock of durable goods and housing, and the value of services (education, health care, and other) purchased by the household during the period. Nominal values are converted into 1986 RMB using the provincial rural CPI from National Statistical Bureau yearbooks.

B. Consumption (Durables and Housing)

Our measure of household consumption per capita includes the value of the flow of services from the stock of consumer durables and housing. The RCRE surveys provide estimates of the original value of housing and durable goods and report current expenditures on durables and new investment in expanding houses. To value the flow of services from housing and durables, we must first use this information to come up with a reasonable estimate of the current value of housing and durables, and then estimate the flow value of consumption.

Durable goods. We assume that durable goods (and production assets) were accumulated in equal portions over the years between 1978 and the first year that the household appeared in the survey. We assume further that durable goods and production assets have a useful life of 7 years (we checked robustness using 5- and 10-year lives) and that the nondepreciated portion of the durable good maintains its "real" value. For each year we depreciated one-seventh of the current value of the good, appreciated the remaining value of the good using a rural provincial capital goods price index, and added the new durables accumulated during that year. From this annual value of the stock of durables, we assume that the household consumes one-seventh of the existing stock of durable goods during the current year.

Housing. We assume that housing is consumed over a 20-year period. For the initial year of the survey and for the first year that a new household appears in the survey, we value the housing stock using information from other households in the village on the real cost per square meter of living area in new housing constructed in the village in year $t - 1$, t, and $t + 1$, and the livable

floor space of the household. For each succeeding year we subtract one-twentieth of the estimated value of the house (as depreciation), appreciate the remaining nineteen/twentieths of the value of the house using the rural capital goods price index, and add on the real value of new additions to the house made during the year. The current flow consumption of housing is one-twentieth of this current year value of housing.

VI. Discussion of CHNS and CCAP Surveys and Sample Sizes in Table 4

The CCAP rural household survey was carried out in the six provinces of Liaoning, Hebei, Shanxi, Sichuan, Hubei, and Zhejiang and covers the year 2000. Altogether, 1,200 households in 60 villages were surveyed, or 20 households per village. In each province, counties were stratified on the basis of the gross value of agricultural and industrial output, and one county was selected from each of the five quintiles. Within each county, townships were similarly stratified, and a township selected from both the upper and bottom half of the distribution. A village was then randomly selected from each of the townships. Within each village, households were randomly selected on the basis of the most recent village household registry. Household membership was defined in a manner analogous to that used by the RCRE survey. Details on the CHNS survey can be found on the CHNS Web site, http://www.cpc.unc.edu/projects/china/.

VII. Spatial Deflator

In order to control for absolute differences in price levels facing households across provinces, we deflate rural incomes and consumption using a spatial deflator constructed by Brandt and Holz (2004) for 1990. Using the NBS rural household survey expenditure data, they construct a single nationwide consumption basket that includes food, clothing, articles in daily use, energy, services, housing, and durable goods. The basket is then priced using provincial-level price data. The range across the nine provinces in 1990 in the absolute price level is more than 50% (Guangdong 1.37 and Sichuan 0.87).

Appendix B

Understanding Differences between the RCRE and NBS Surveys

In principle, there is a range of plausible explanations for observed differences between the NBS and RCRE series, including sampling strategy (both at the village and household levels), survey design, and differences in how income and consumption are defined and measured. Without actual access to the NBS data, information on survey implementation, and criteria used in the data "cleaning" and organization process, there are many issues that we cannot

confront. Still, we are able to consider potential differences that remain in the way that income and consumption are defined. In appendix A, Section IV, we considered the possibility that biases could be caused by oversampling of agricultural households in the RCRE surveys. If we compare the few indicators that can be calculated in a comparable manner across surveys, it is not obvious that an explicit focus on agricultural households by the Ministry of Agriculture is driving the bias. Other types of sampling bias may be likely as a result of the panel nature of the RCRE data set or the initial sampling design, but these are more difficult to assess without full access to both data sources.

One of the important differences between the two series shown in figure 1 lies in the valuation of in-kind income (and consumption). In constructing consumption and income from the RCRE surveys, we valued both in-kind income and consumption in all years at market prices. In contrast, prior to 1990, the NBS valued in-kind components at prices that were at or near the quota price; subsequently, they used a higher set of prices. While Ravallion and Chen (2004) suggest that NBS started valuing in-kind income and consumption at market prices in 1990, official documentation of the rural household survey appears to indicate otherwise (NBS 2000).[43] From 1991 to 1996 NBS instructed survey teams to use an "average contract price" to value nonmarketed grain from own production. This average contract price was the weighted average of the quota price and the above quota price for sales over quota to the local grain bureau. Both of these prices were administratively determined. Survey teams were further instructed to use weighted average market prices only if there were no local administrative prices (meaning no local crop procurement).[44] It was not until 1997 that NBS instructed survey teams to value all nonmarketed agricultural commodities at market prices. However, this was revised in 1998 with instructions to value grain and meat products at 90% of the market price and other nonmarketed commodities at 85% of market price.

An NBS official involved with the rural household survey since 1989 verified that the official documentation cited above (and translated in app. C) accurately reflected the survey protocol for the rural household survey during the 1990s. This official also volunteered an additional source of downward bias in income and consumption growth between 1991 and 1995: some localities used the

[43] A translation of these sections of the authoritative NBS document and a scan of Chinese language originals are available from Giles's Web site (www.msu.edu/~gilesj/).

[44] Wang, Xia, and Liu (1996) also point out that up through 1995, NBS policy was to use a weighted average of the quota price (called *tonggou jia* during the 1990s) and above-quota/negotiated price (*chaoguo jia*).

same weighted average quota and above quota sales price calculated in 1990 for purposes of valuing in-kind income for years between 1991 and 1995.[45]

Relative to the RCRE data, NBS treatment of the in-kind component likely results in an underestimate of incomes and consumption prior to 1990 and an overestimate of growth in incomes or consumption if one calculates average annual growth rates for the 1987–99 period. Given that grain income was valued at a price still systematically related to the quota price after 1990, we expect the difference in average income to be correlated with the gap between the market and the quota price.

Two other differences in the income series lie in the treatment of taxes and fees and depreciation on fixed assets, both of which are subtracted from the NBS series. We concentrate on earned income and do not subtract off taxes because we are focusing on earnings ability of households, and taxes would confound the issue. Second, recorded depreciation expenses are notoriously arbitrary, so we have depreciated assets ourselves as discussed in appendix A above.

To examine the extent to which our adjustments to the RCRE series may be driving differences, we do two things. First, we subtract off taxes and fees, as well as depreciation on fixed assets for all years. Second, we use the quota price in the valuation of in-kind grain income in the RCRE data for 1986–90. For all years after 1990, the lack of a consistent method for valuing nonmarketed commodities makes it impossible for us to mimic the method of valuation used by the NBS, so we use the market price to value the in-kind components. Figure B1 presents RCRE income and expenditure series calculated in a manner consistent with the calculations purportedly followed in the NBS series prior to 1990.

1. Implications for Income

From 1986 to 1990 differences between the two surveys in the treatment of taxes and fees and the valuation of in-kind income are the source of approximately half of the differences in per capita income between the two series. Depreciation explains an additional 10% of the gap. For years after 1990, tax, fees, and depreciation are the source of roughly 25% of the gap in the years just after 1990 and nearly the entire gap by 1999. Significant differences persist, however, between the two series, which is not too surprising given that NBS was valuing in-kind income at the weighted average of two administrative prices until 1997. The gap in per capita income widens between the two series, peaking in 1995, before eventually disappearing after 1997.

[45] Loren Brandt interview with an NBS division chief conducted on December 18, 2004.

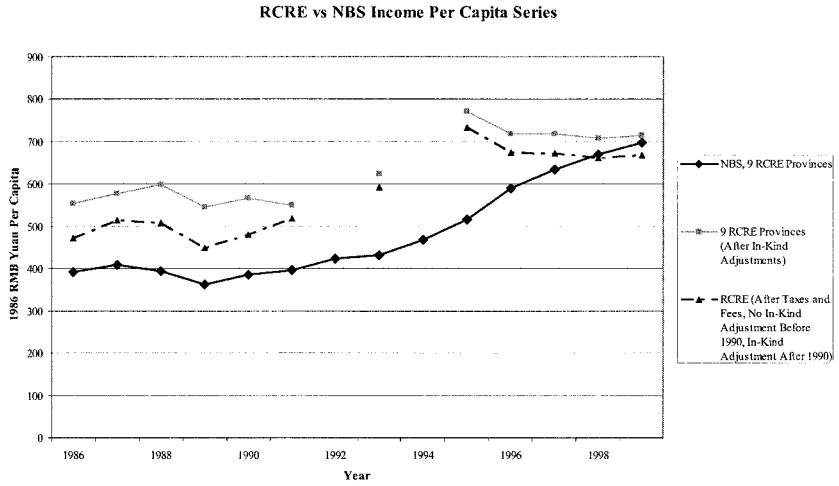

Figure B1. Differences between RCRE and NBS income per capita series

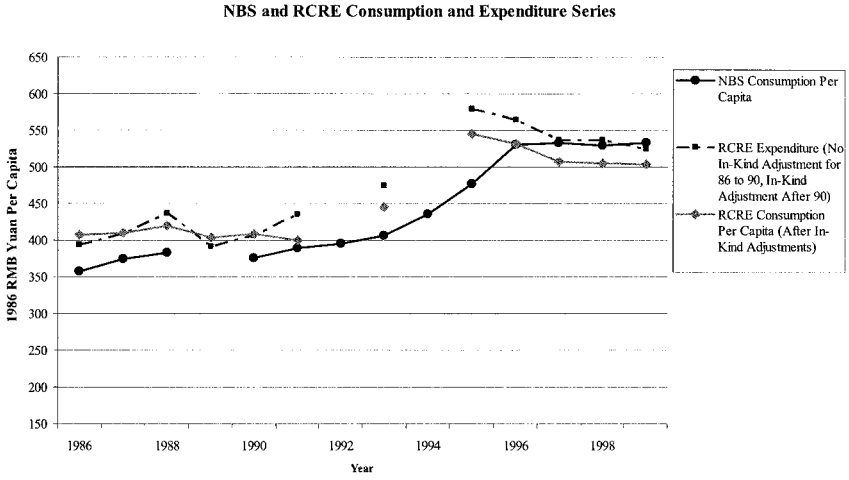

Figure B2. Differences between RCRE and NBS consumption and expenditure per capita series

Adjustments to valuation of grain during the early period help to narrow some of the differences in growth implied by the two series, but a significant difference remains. With the adjustment, the difference between the two series in annual growth in per capita incomes declines from 2.8% (4.6% − 1.8%) to 2.2% (4.6% − 2.4%).

II. Implications for Consumption

In general, the consumption series line up better than do the income series both in terms of levels and growth; however, there are differences. First, the

NBS consumption series is really an expenditure series. Valuing consumption from in-kind income consistently and looking at expenditures eliminates much of the difference in per capita consumption differences for years prior to 1990. This also helps to narrow the difference in the growth rate in expenditure between the two series from 1.3% (2.9% − 1.6%) to 0.8% (2.9% − 2.2%). As we saw in the case of the income data, however, there is a widening gap in per capita expenditure between the two series from 1991 to 1995, followed by elimination of the gap. We believe that for the purposes of our analyses, however, consumption is a more appropriate measure than expenditure.

III. Can Documented Differences in Valuation of In-Kind Income Explain the Gap between RCRE and NBS Averages after 1990?

We believe that some of the gap in both the consumption and income series after 1990 reflects the fact that the NBS valued in-kind grain at a price that remained well below the market price. Brandt and Holz (2004) also discuss the issue of grain pricing by the NBS in some detail. With the change in NBS method of valuing nonmarketed grain in 1990, they find that the new average implicit price at which the NBS valued grain increased from 0.378 to 0.513 yuan per kilogram. This is still only modestly higher than the average state procurement price of 0.5 yuan per kilogram and considerably below the average market and state-guidance prices of nearly 1.0 yuan per kilogram. Given that we know from NBS documentation that the rural household survey teams are using average unit sales prices calculated from the quota and above-quota administrative prices, this implies a weighted average price that will be well below the market price if the above-quota price is close to the expected market price.

Indirect confirmation that imputed prices were well below market is suggested by differences in the behavior of farm incomes between 1993 and 1995 in the NBS and RCRE samples. Over this period, crop prices increased by slightly more than 90%. Farm input prices increased by approximately 50%, while the rural CPI increased by 45%. Assuming that value added in farming is 60%, ceteris paribus, this should result in an increase in real farm incomes from cropping of 50%.[46] This is more or less what we see in the RCRE data; however, the increase in the NBS data is less than half of this. Much of the difference between the two series in the growth of per capita incomes from 1993 to 1995 can be linked to the more rapid growth in crop income in the RCRE data.

[46] Sixty percent value added is a rough conservative estimate based on information from the 2000 CCAP survey discussed in the paper.

Given NBS treatment of in-kind incomes after 1990, we expect the bias in income estimates to depend on the gap between the quota and market price, which tended to be highly cyclical. In fact, we find a strong positive correlation (0.92) between the size of the gap between the two income series, and our estimate of the ratio of the market-to-quota price, which is as high as 1.46 in 1995, then falls below 1 by 1997. After 1997, the NBS series is directly related to the market price, and we observe some convergence in the income and consumption series.

Finally, biases introduced by NBS approaches to valuation of in-kind income may also help to explain some of the slight differences in the behavior of the Gini coefficient over time between the two series. The NBS income Gini coefficients rise monotonically. The RCRE data suggest modest growth prior to 1995 and then a significant increase over the next 4 years. One interpretation for this behavior is that the much higher valuation of in-kind income in the RCRE data for the years 1993–95 is helping to dampen the effect of dis-equalizing growth in wage and business incomes. The effect is temporary, however, and once farm prices begin to fall, we see the sharp increase in inequality associated with falling farm incomes.

Appendix C

Testing for Statistical Significance in Differences of Lorenz Curves, CDFs, and Generalized Lorenz Curves

We test for the statistical significance of the difference between Lorenz curves and generalized Lorenz curves using DAD, a distribution analysis software developed by Duclos et al. (2004). We present estimated differences between Lorenz curves and generalized Lorenz curves at different cumulative population shares in tables C1 and C2, respectively. In the discussion of Section IV.A we consider differences between curves to be statistically significant when the difference between curves is greater than twice the standard deviation of the differences.

In figure 5, we also use DAD to find a critical poverty line of 226 RMB per capita. This is where the 1999 CDF of income per capita crosses the 1987 CDF from above. The standard deviation of this critical poverty line is 36.3.

TABLE C1

DO LORENZ CURVES FOR INCOME PER CAPITA DIFFER SIGNIFICANTLY?

Cumulative Population Share (%)	Difference 1999LC − 1987LC	Cumulative Population Share (%)	Difference 1999LC − 1995LC
1	−.0018	1	−.0012
	(.0005)		(.0006)
2	−.0030	2	−.0024
	(.0007)		(.0009)
3	−.0039	3	−.0034
	(.0009)		(.0011)
4	−.0048	4	−.0044
	(.0012)		(.0013)
5	−.0056	5	−.0053
	(.0015)		(.0015)
10	−.0095	10	−.0097
	(.0028)		(.0024)
25	−.0206	25	−.0194
	(.0066)		(.0049)
50	−.0374	50	−.0290
	(.0125)		(.0082)
75	−.0427	75	−.0296
	(.0167)		(.0097)
90	−.0331	90	−.0175
	(.0162)		(.0086)
95	−.0257	95	−.0092
	(.0142)		(.0071)
99	−.0108	99	.0001
	(.0069)		(.0041)
100	.0000	100	.0000
	(.0000)		(.0000)

Note. Standard errors of difference are in parentheses.

TABLE C2

DO GENERALIZED LORENZ CURVES FOR INCOME PER CAPITA DIFFER SIGNIFICANTLY?

Cumulative Population Share (%)	Difference 1999GLC − 1987GLC	Cumulative Population Share (%)	Difference 1999GLC − 1995GLC
1	−1.0571	1	−.8666
	(.3471)		(.4722)
2	−1.5522	2	−1.8624
	(.5445)		(.6417)
3	−1.8577	3	−2.7566
	(.7120)		(.7784)
4	−2.0672	4	−3.6421
	(.8837)		(.9120)
5	−2.2014	5	−4.5068
	(1.1067)		(1.0757)
10	−2.3117	10	−8.7575
	(2.1502)		(1.8195)

TABLE C2 (*Continued*)

Cumulative Population Share (%)	Difference 1999GLC − 1987GLC	Cumulative Population Share (%)	Difference 1999GLC − 1995GLC
25	.3612	25	−20.1981
	(5.2284)		(3.9434)
50	13.1336	50	−37.1695
	(10.5384)		(8.2987)
75	44.7752	75	−52.0422
	(16.2268)		(14.7880)
90	81.1552	90	−55.7593
	(20.8501)		(20.1826)
95	99.3007	95	−55.4198
	(23.5031)		(22.3429)
99	122.7926	99	−55.2562
	(28.5704)		(24.2397)
100	135.5869	100	−58.7977
	(32.5647)		(25.4232)

Note. Standard errors of difference are in parentheses.

References

Aghion, Philippe, Eve Caroli, and Cecilia Garcia-Penalosa. 1999. "Inequality and Economic Growth: The Perspective of the New Growth Theories." *Journal of Economic Literature* 37, no. 4 (December): 1615–60.

Atkinson, A. B. 1987. "On the Measurement of Poverty." *Econometrica* 55, no. 4: 749–64.

Banerjee, Abhijit V., and Esther Duflo. 2004. "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth* 8, no. 3 (September): 267–99.

Benjamin, Dwayne, and Loren Brandt. 1999. "Markets and Inequality in Rural Contemporary China: Parallels with the Past." *American Economic Review Papers and Proceedings* 89, no. 2 (May): 292–95.

Benjamin, Dwayne, Loren Brandt, Paul Glewwe, and Guo Li. 2002. "Markets, Human Capital and Income Inequality in an Economy in Transition: The Case of Rural China." In *Inequality around the World*, ed. Richard Freeman. Houndsmills, Basingstoke: Palgrave-Macmillan.

Bramall, Chris. 2001. "The Quality of China's Household Income Surveys." *China Quarterly*, no. 167 (September): 689–705.

Brandt, Loren, and Carsten Holz. 2004. "Spatial Price Differences in China." Unpublished manuscript, Department of Economics, University of Toronto.

Chen, Shaohua, and Martin Ravallion. 1996. "Data in Transition: Assessing Rural Living Standards in Southern China." *China Economic Review* 7, no. 1:23–56.

Cook, Sarah. 1998. "Work, Wealth and Power in Agriculture: Do Political Connections Affect Returns to Household Labor?" In *Zouping in Transition: The Process of Reform in Rural North China*, ed. Andrew G. Walder. Cambridge, MA: Harvard University Press.

Davidson, Russell, and Jean-Yves Duclos. 2000. "Statistical Inference for Stochastic

Dominance and for the Measurement of Poverty and Inequality." *Econometrica* 68, no. 6 (November): 1435–64.

Dayton-Johnson, Jeff, and Pranab Bardhan. 2002. "Inequality and Conservation of the Local Commons: A Theoretical Exercise." *Economic Journal* 112, no. 481 (July): 577–602.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore: Johns Hopkins University Press for the World Bank.

———. 2005. "Measuring Poverty in a Growing World (or Measuring Growth in a Poor World)." *Review of Economics and Statistics* 87, no. 1 (February): 1–19.

Deininger, Klaus, and Lyn Squire. 1996. "New Data Set Measuring Income Inequality." *World Bank Economic Review* 10, no. 3 (September): 565–91.

Duclos, Jean-Yves, Abdelkrim Araar, and Carl Fortin. 2004. "DAD 4.3: A Software for Distributive Analysis/Analyse Distributive." MIMAP Programme, International Development Research Centre, Government of Canada and CREFA, Université Laval.

Foster, J. E., and A. F. Shorrocks. 1988. "Poverty Orderings." *Econometrica* 56:173–77.

Gustafsson, Bjorn, and Li Shi. 2002. "Income Inequality within and across Counties in Rural China, 1988 and 1995." *Journal of Development Economics* 69, no. 1 (October): 179–204.

Hare, Denise. 1994. "Rural Nonagricultural Activities and Their Impact on the Distribution of Income: Evidence from Farm Households in Southern China." *China Economic Review* 5, no. 1 (Spring): 59–82.

Hutzler, Charles. 2003. "For China's Poor Regions, a Kinder, Gentler Beijing—Incoming Leadership Makes a Priority of Those Ignored by Headlong Liberalization." *Wall Street Journal*, February 25, 2003, A13.

Jalan, Jyotsna, and Martin Ravallion. 2002. "Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China." *Journal of Applied Econometrics* 17, no. 4 (July–August): 329–46.

Jenkins, Stephen P. 2000. "Modelling Household Income Dynamics." *Journal of Population Economics* 13, no. 4 (December): 529–67.

Kanbur, Ravi, and Xiaobo Zhang. 1999. "Which Regional Inequality? The Evolution of Rural-Urban and Inland-Coastal Inequality in China from 1983 to 1995." *Journal of Comparative Economics* 27, no. 4 (December): 686–701.

Khan, Azizur Rahman, and Carl Riskin. 1998. "Income and Inequality in China: Composition, Distribution and Growth of Household Income, 1988 to 1995." *China Quarterly*, no. 154 (June): 221–53.

Kung, James, and Yiu-fai Lee. 2001. "So What If There Is Income Inequality? The Distributive Consequence of Nonfarm Employment in Rural China." *Economic Development and Cultural Change* 50, no. 1 (October): 19–46.

Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *American Economic Review* 45:1–28.

Meng, Xin. 2004. "Economic Restructuring and Income Inequality in Urban China." *Review of Income and Wealth* 50, no. 3:357–79.

Morduch, Jonathan, and Terry Sicular. 2002. "Rethinking Inequality Decomposition,

with Evidence from Rural China." *Economic Journal* 112, no. 476 (January): 93–106.

NBS (National Bureau of Statistics). 1993. *Tabulation on the 1990 Population Census of the People's Republic of China* [*Zhongguo 1990 nian renkou pucha ziliao*]. Beijing: National Bureau of Statistics Publishing House.

———. 1997. *Abstract of the 1996 Agricultural Census of China* (English version). Beijing: National Bureau of Statistics Publishing House.

———. 2000. *How to Use the Statistical Yearbook* [*Ruhe shiyong tongji nianjian*], ed. Chengxiang Liu, Ke Liu, and Zhaofeng Jin. Beijing: National Bureau of Statistics Publishing House [Zhongguo tongji chuban she], September. The ISBN number for this volume is 7-5037-3380-2/C 1834.

———. 2003. *China Statistical Yearbook, 2003* [*Zhongguo tong ji nian jian 2003*]. Beijing: National Bureau of Statistics Publishing House.

Nee, Victor. 1992. "Organizational Dynamics of Market Transition: Hybrid Forms, Property Rights, and Mixed Economy in China." *Administrative Science Quarterly* 37:1–127.

Ravallion, Martin, and Shaohua Chen. 2004. "China's (Uneven) Progress against Poverty." Unpublished manuscript, World Bank, Washington, DC.

Rawski, Thomas G. 2001. "What Is Happening to China's GDP Statistics?" *China Economic Review* 12, no. 4:347–54.

Rozelle, Scott. 1994. "Rural Industrialization and Increasing Inequality: Emerging Patterns in China's Reforming Economy." *Journal of Comparative Economics* 19, no. 3 (December): 362–91.

Shorrocks, Anthony F. 1982. "Inequality Decomposition by Factor Components." *Econometrica* 50, no. 1:193–211.

———. 1983a. "The Impact of Income Components on the Distribution of Family Incomes." *Quarterly Journal of Economics* 98, no. 2:311–26.

———. 1983b. "Ranking Income Distributions." *Economica* 50:3–17.

Tsui, Kai-yuen. 1998. "Trends and Inequalities of Rural Welfare in China: Evidence from Rural Households in Guangdong and Sichuan." *Journal of Comparative Economics* 26, no. 4 (December): 783–804.

Vijverberg, Wim, and Donald Mead. 2000. "Household Enterprises." In *Designing Household Survey Questionnaires for Developing Countries: Lessons from Fifteen Years of the Living Standards Measurement Study*, ed. Margaret Grosh and Paul Glewwe, vol. 2, 105–37. Washington, DC: World Bank.

Wang, Sangui, Ying Xia, and Xiaozhan Liu. 1996. "Criteria to Measure Poverty in China" [*Zhongguo pingkun zhibiao yanjiu*]. In *Research on Economic and Technological Development in Agriculture* [*Nongye jingji yu keji fazhan yanyiu*]. Beijing: China Agriculture Technology Press.

World Bank. 2003. "China: Promoting Growth with Equity; Country Economic Memorandum." Report no. 24169-CHA, Poverty Reduction and Economic Management Unit, East Asia and Pacific Region, World Bank, Washington, DC.