

Sicular, Terry; Yue, Ximing; Gustafsson, Björn; Li, Shi

Working Paper

The urban-rural income gap and inequality in China

WIDER Research Paper, No. 2006/135

Provided in Cooperation with:

United Nations University (UNU), World Institute for Development Economics Research (WIDER)

Suggested Citation: Sicular, Terry; Yue, Ximing; Gustafsson, Björn; Li, Shi (2006) : The urban-rural income gap and inequality in China, WIDER Research Paper, No. 2006/135, ISBN 929190919X, The United Nations University World Institute for Development Economics Research (UNU-WIDER), Helsinki

This Version is available at:

<https://hdl.handle.net/10419/63438>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



UNITED NATIONS
UNIVERSITY

UNU-WIDER

World Institute for Development
Economics Research

Research Paper No. 2006/135

The Urban-Rural Income Gap and Inequality in China

Terry Sicular,¹ Ximing Yue,²
Björn Gustafsson,³ and Shi Li⁴

November 2006

Abstract

Using new household survey data for 1995 and 2002, we investigate the size of China's urban-rural income gap, the gap's contribution to overall inequality in China, and the factors underlying the gap. Our analysis improves on past estimates by using a fuller measure of income, adjusting for spatial price differences and including migrants. Our methods include inequality decomposition by population subgroup and the Oaxaca-Blinder decomposition. Several key findings emerge. First, the adjustments substantially reduce China's urban-rural income gap and its contribution to inequality. Nevertheless, the gap remains large and has increased somewhat over time. Second, after controlling for household characteristics, location of residence remains the most important factor underlying the urban-rural income gap. The only household characteristic that contributes substantially to the gap is education. Differences in the endowments of, and returns to, other household characteristics such as family size and composition, landholdings, and communist party membership are relatively unimportant.

Keywords: China, urban-rural inequality, income

JEL classification: D31, O18, P25

Copyright © UNU-WIDER 2006

¹University of Western Ontario, ²Chinese Academy of Social Sciences, ³Göteborg University, ⁴Beijing Normal University.

This study has been prepared within the UNU-WIDER project on Inequality and Poverty in China.

UNU-WIDER acknowledges the financial contributions to its research programme by the governments of Denmark (Royal Ministry of Foreign Affairs), Finland (Ministry for Foreign Affairs), Norway (Royal Ministry of Foreign Affairs), Sweden (Swedish International Development Cooperation Agency—Sida) and the United Kingdom (Department for International Development).

ISSN 1810-2611

ISBN 92-9190-919-X

ISBN-13 978-92-9190-919-3

Acknowledgements

We are grateful to Anthony Shorrocks and other participants of the UNU-WIDER project meeting on Inequality and Poverty in China, 26-27 August 2005, Helsinki, for their comments, and to two anonymous referees for their detailed suggestions. We thank the Swedish International Development Cooperation Agency (SIDA), the Ford Foundation, Centre for International Governance Innovation (CIGI), and the University of Western Ontario for funding that facilitated this work.

The World Institute for Development Economics Research (WIDER) was established by the United Nations University (UNU) as its first research and training centre and started work in Helsinki, Finland in 1985. The Institute undertakes applied research and policy analysis on structural changes affecting the developing and transitional economies, provides a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and promotes capacity strengthening and training in the field of economic and social policy making. Work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.

www.wider.unu.edu

publications@wider.unu.edu

UNU World Institute for Development Economics Research (UNU-WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Camera-ready typescript prepared by Lorraine Telfer-Taivainen at UNU-WIDER

The views expressed in this publication are those of the author(s). Publication does not imply endorsement by the Institute or the United Nations University, nor by the programme/project sponsors, of any of the views expressed.

1 Introduction

Studies of China's inequality almost universally report that the gap between urban and rural household incomes in China is large, has increased over time, and contributes substantially to overall inequality. According to most estimates mean per capita income in urban China is more than triple that in rural areas, giving China one of the highest urban-rural income ratios in the world. The size of this gap has been discussed in the Chinese official media, is noted in government and communist party reports, and is the motivation for recent major policy initiatives such as the 'Build a Socialist New Countryside' campaign, which aims to reduce the gap by boosting public spending in rural areas.

China's urban-rural income gap is often attributed to policies that have inhibited labour mobility, most importantly the household registration or *hukou* system. The household registration system was established during the Maoist period to control population movement. It has continued to the present and is reinforced by a range of complementary policies such as taxation of urban employers that hire migrants, prohibition of urban employment of migrants in some trades, and the denial of urban public services such as education to unregistered households.¹ In recent years the government has carried out reforms of the *hukou* system so as to allow greater mobility, but substantial barriers remain (Wang 2004). These barriers are thought to protect the welfare of registered urban residents, a politically sensitive group, but at the same time they create political concerns of a different sort.

Whether or not concerns about the urban-rural income gap are justified depends, among other things, on the true magnitude of the gap as well as on the factors that underlie the gap. To date a range of studies have examined China's urban-rural income gap (e.g., Knight and Song 1999; Yang and Zhou 1999; Zhao and Tong 2000; Shi 2004; Sicular et al. 2004; Benjamin et al. 2005). For several reasons, most linked to data constraints, past estimates of the income gap are likely biased.

First, most income data for China do not include certain components. One missing component is housing-related income, specifically, the imputed rental value of owner-occupied housing and imputed subsidies on publicly-owned rental housing. Housing-related income is likely to differ systematically between urban and rural areas, and it is of recent interest because in the late 1990s China privatized urban housing (Khan and Riskin 2007). Also missing from income is the value of household consumption of public services in areas such as education, health care, and local infrastructure. Consumption of public services is, again, likely to be higher in urban than rural areas, and so its exclusion would cause understatement of China's urban-rural gap. Second,

¹ See Chan and Zhong (1999) and Solinger (1999) for details on the *hukou* system and related policies.

most studies do not control for spatial differences in the cost of living. This is understandable, as systematic information on spatial price differences has been scarce. Still, if the cost of living in urban areas is substantially higher than that in rural areas, then the real gap in incomes may be smaller than that reported in the literature. Third, most estimates of China's urban-rural income gap are based on data that exclude unregistered migrants resident in urban areas. Since rural-to-urban migration is generally considered an important mechanism for narrowing the urban-rural income gap, excluding rural-to-urban migrants is problematic. Excluding this group, which in China has mean income below that of registered urban and above that of rural residents, causes overstatement of the urban-rural income gap. Including migrants is increasingly relevant in studies of China, because restrictions on migration have been loosened and migration has grown accordingly.

With these considerations in mind, here we recalculate the size of China's urban-rural income gap, estimate the contribution of the revised gap to overall inequality, and analyze the factors underlying the gap. For our analysis we use household and individual data from household income surveys for 1995 and 2002. These surveys were conducted under the auspices of the Chinese Academy of Social Sciences (CASS). They are large, nationally representative, and contain detailed information on household income and other relevant household and individual characteristics. The CASS data have certain advantages. They are relatively recent, and so provide more up-to-date information than is generally available. Other data with wide regional coverage, most notably from the National Bureau of Statistics's (NBS) household survey, are typically only available to researchers in tabulated or aggregated form. Alternative datasets that provide household-level survey data such as the China Health and Nutrition Survey (CHNS) have narrower regional coverage than the CASS survey. Finally, the CASS dataset is unusual in that it contains information on housing-related income components and on rural-to-urban migrants.

The first step of our analysis is to recalculate the size of the urban-rural gap and its contribution to national inequality. We do so for China as a whole and for its three major regions—the east, center and west. In this recalculation we make three modifications that bring our measurement of the gap closer to international best practice and allow more comparability with studies for other countries.

First, we use a fuller measure of income that includes housing-related components of income. Unfortunately, we cannot measure the implicit subsidies associated with household consumption of public services. From a theoretical standpoint this should be included to fully capture urban-rural differences, but the necessary information is unavailable. More generally, data on household consumption of public services is unavailable for most countries and rarely included in international calculations of

household income.² Second, we adjust for spatial differences in the cost of living. Costs of living can differ systematically among regions and between urban and rural areas, and so ideally studies of inequality should use incomes that have been adjusted using spatial price deflators. Due to lack of data on regional price levels, spatial price deflation is rare in studies on China. An exception is Ravallion and Chen (2004), which uses estimates of the urban and rural poverty lines to adjust for cost of living differences, and then recalculates national inequality. More recently, Benjamin et al. (2005) uses new spatial price indices from a study by Brandt and Holz (2004) to recalculate the level of national inequality. Neither of these studies, however, uses spatial price deflation for a detailed analysis of the urban-rural income gap. Here we use the Brandt and Holz (ibid.) spatial price indices to deflate incomes and then recalculate the urban-rural income gap and its contribution to inequality. Where relevant, we compare our findings to those of Ravallion and Chen (ibid.) and Benjamin et al. (ibid.).³ Third, we include rural-to-urban migrants. Past analyses exclude unregistered rural migrants living in cities. The 2002 CASS survey data contain information for a sample of rural migrants resident but not registered in urban areas. Using this sample, we are able to provide some indicative findings that include migrants. We discuss the coverage of our migrant sample in more detail below, as well as some broader methodological concerns regarding the measurement of China's urban and migrant populations.

The second step of our analysis is an investigation of the factors underlying the urban-rural gap. Here we use the Oaxaca-Blinder decomposition. The Oaxaca-Blinder method cannot identify how particular policies such as the household registration system contribute to the gap, but it gives information on the extent to which the gap reflects differences between urban and rural areas in household characteristics as opposed to simple location of residence. This method also gives a measure, albeit from a partial equilibrium perspective, of how large the gap would be if rural and urban groups had similar characteristics. Such information is useful from a policy perspective. For example, if differences in educational characteristics between rural and urban areas contribute substantially to the gap, as we find they do, then policy makers may wish to focus their attention on the determinants, and consequences, of education levels in the two sectors.

² In their review of household income survey data for twenty-five countries, Smeeding and Weinberg (2001) report that only one country collects information on consumption of public education services (Australia), and only three on government subsidized health care services (Australia, Germany and the USA). Note that some studies (e.g., Eastwood and Lipton 2000; Sahn and Stifel 2003) look directly at urban-rural gaps in levels of education, health, and other welfare-related variables. Some information on urban-rural gaps in such variables for China can be found in China Development Research Foundation and UNDP (2005) and Zhang and Kanbur (2005).

³ Note that these two studies use different data than those used here. Ravallion and Chen (2004) use tabulated data provided by the NBS, which has broader geographical coverage than the CASS dataset but is not at the household-level. Benjamin et al. (2005) use household-level data from the China Health and Nutrition Survey, which has narrower regional coverage than the CASS survey.

Several key findings emerge from our analysis. We find that after recalculation, the urban-rural income gap is substantially reduced. While including housing-related income components increases the income gap somewhat, adjusting for spatial price differences dramatically reduces it. Including migrants narrows the gap further. With these revisions, China still has a relatively large urban-rural income ratio, but that ratio is within the range of most other countries. It follows that these adjustments also reduce the contribution of China's urban-rural gap to overall inequality. After recalculating income and including migrants, we find that in 2002 the urban-rural gap contributes about one quarter of overall inequality, as compared to estimates of 50 per cent or more in most studies.

The Oaxaca-Blinder decomposition reveals that household and individual characteristics such as education, age, and household demographics, indeed contribute to the urban-rural income gap. Differences in the endowments of such characteristics, holding the returns to these characteristics constant, contributed about half of the income gap. Location of residence, including differences between urban and rural areas in the returns to household and individual characteristics, contributed the other half. Interestingly, the contribution of location declined between 1995 and 2002, although only modestly. This is consistent with the increase in spatial mobility during this time. The decomposition reveals further that education is the most important non-location characteristic underlying the urban-rural income gap. In 2002 differences in education levels between urban and rural areas contributed one quarter of the income gap. Differences in the endowments of and returns to other household characteristics such as family size and composition, landholdings, and Party membership are, on balance, less important. That said, in the long term education levels are endogenous, and current investments in human capital are likely affected by other household characteristics such as family size and composition.

2 Definitions and data

The data used for the analysis in this paper come from two rounds of the CASS Household Income Survey conducted in 1996 and 2003 for the reference periods of 1995 and 2002. These surveys were carried out under the direction of a team of researchers consisting of scholars at the Institute of Economics, Chinese Academy of Social Sciences, and researchers from other countries. The data were collected by the NBS using survey instruments designed by the project research team. A detailed description of the data can be found in Li et al. (2007). Here we point out some of the main features of the data set and discuss aspects most relevant to our analysis.

Regional coverage changed somewhat between the two years of the survey. To ensure comparability between the results for the two years, we use a subsample having the property that each location (province*rural, province*urban) was present in the survey for both years under investigation. The rural sample covers Anhui, Beijing, Gansu,

Guangdong, Guizhou, Hebei, Henan, Hubei, Hunan, Liaoning, Jiangsu, Jiangxi, Jilin, Shaanxi, Shandong, Shanxi, Sichuan, Yunnan, and Zhejiang. The urban sample covers Anhui, Beijing, Gansu, Guangdong, Henan, Hubei, Jiangsu, Liaoning, Shanxi, Sichuan, and Yunnan.⁴

Since urban residents were over-sampled in 1995 and under-sampled in 2002, we weight the urban and rural subsamples so that their population shares equal those in the total population according to official NBS census-based population data. With this adjustment, the sample distribution between rural and urban areas is consistent with the official population distribution between urban and rural areas for all of China. All analyses using the combined urban and rural samples use this population weight adjustment.

A limitation of most household survey data for China is that rural-to-urban migrants who do not have an urban residence permit are excluded. For 2002 the CASS survey includes a special sample of migrants, making it possible to produce more complete estimates for that year. In Section 5 below we describe the migrant sample and explore how including migrants influences the size of the rural-urban gap and its contribution to inequality. Section 5 also contains a more general discussion of China's urban population statistics.

The target variable for this study is household per capita disposable income.⁵ This includes cash income, retained in-kind income (important in rural China, particularly at the beginning of the period studied), and other income in kind (relevant in urban China in the past, although declining in importance in recent years). Net taxes and fees are subtracted.

⁴ The sample also includes Chongqing, which was part of Sichuan province in 1995 but became a separate province in 1997. For consistency, in the analysis Chongqing observations are treated as part of the Sichuan sample in both 1995 and 2002.

⁵ The advantages and disadvantages of using income as the target variable in studies of inequality have been discussed extensively in the literature (see, for example, Deaton 1997; Atkinson and Bourguignon 2000; Gradín et al. 2004; WIDER 2005). One disadvantage of using income is that income fluctuates over the lifecycle and can vary from year to year. If households can save and borrow, however, then in the face of such income fluctuations they can smooth consumption. Consumption expenditures, then, may better reflect expected permanent income. The use of consumption, however, also has its drawbacks. Consumption, like income, can fluctuate over time as needs can vary over the lifecycle. Consumption also depends heavily on the habits and preferences of individuals, so that some measured inequality will be spurious. From a practical standpoint, using consumption raises difficulties in the treatment of infrequently purchased consumer durables.

We would argue that certain features of China provide reasons to use income rather than consumption. In China financial markets are still undeveloped and households have limited opportunities to borrow and save. The theoretical advantages of using consumption, then, are not fully applicable. Availability of consumption data is also an issue, and where available Chinese consumption data count the entire cost of consumer durable purchases as current year expenses. Perhaps for these reasons, inequality measured over consumption per capita is often higher than that measured over income per capita, and few inequality studies for China use consumption data (we know of only two, Jalan and Ravallion 1998, and Wu and Perloff 2005). In view of these considerations, and so that our findings are comparable to most other studies of inequality in China, we use income as the target variable.

Most economists believe that income should include housing-related components. The NBS does not include these components in disposable income, nor do most other household income data for China. Our estimates of average household income in China use the NBS definition but add in housing subsidies and imputed rent. Income levels here are therefore higher than those obtained using the NBS definition. Depending on the distributional profile of housing subsidies and imputed rent, our definition of income could show larger or smaller inequality than the NBS definition. In fact, we find that including housing increases inequality, which is not surprising as higher-income and urban households tend to enjoy larger housing subsidies and imputed rents.

During the period under investigation, China carried out housing reform in urban China. In the past most urban households had lived in public housing and paid low rent, implying that they received rental housing subsidies. These subsidies largely benefited better-off households (Khan et al. 1993). In the late 1990s the government privatized urban public housing. By 2002 most urban residents owned their homes and no longer received rental housing subsidies. Rather, they now received the imputed rents from owned housing. For urban China and China as a whole, inclusion of housing components and changes in these components due to the housing reform could influence the measured urban-rural gap and inequality.⁶

Our analysis treats the household as the income-receiving unit. Disposable income of each household is then divided by the number of household members. Following what is now common practice in analysis of income distributions, we assign this household average to each member of the household. Individuals are thus the unit of analysis, and we abstract from intra-household allocation issues.⁷

Since price levels have changed over time, and differentially among provinces and between rural and urban areas, we use official provincial consumer price indices to express 2002 incomes in 1995 prices. Note that separate indices are available for rural versus urban areas in each province. We use these separate indices so that deflation factors can differ between urban and rural areas within provinces as well as among provinces. Prices differ not only across time, but also spatially at any point in time. This is especially true in a geographically large country like China. Analyses of income inequality for China typically do not adjust for spatial price differences because price data by region have been unavailable. A recent study for China by Brandt and Holz (2004) gives estimates of regional differences in the costs of living among provinces and between urban and rural areas. Their study uses raw regional price data for 1990 to

⁶ In addition, and probably at least as important as its impact on income distribution, housing reform has led to a redistribution of wealth. Changes in the distribution of wealth are not the topic of this paper, but interested readers will find discussion of this topic in Zhao and Ding (2007).

⁷ Some analyses of inequality use equivalence scales to adjust for differences in household composition and size. Unfortunately, no recent estimates of equivalence scales for urban and rural China are available, and we do not have the information needed to estimate them.

calculate baseline spatial cost of living indices for that year. The 1990 spatial price indices are then extrapolated to later years using provincial urban and rural consumer price indices.

The Brandt-Holz spatial price deflators have some limitations. One is that their estimates of housing costs are based on the costs of housing construction materials, and the difference in the costs of construction materials between urban and rural areas is typically smaller than the difference in costs of housing services. For this reason, the Brandt-Holz estimates may understate the price differential between urban and rural areas. Also, they only have raw price data for 1990, and they use a basket of consumption quantities for 1990. The accuracy of extrapolations from 1990 will obviously decline the longer the intervening time period because the structure of consumption and also the quality of goods and services consumed changes over time. Here we are extrapolating a fairly long way, to 2002. Despite these limitations, the Brandt-Holz estimates provide an opportunity to correct, albeit imperfectly, for spatial price differences, and to see how such corrections affect the level and composition of inequality. Below we present findings calculated both with and without spatial price adjustments. In most cases the differences are substantial.

3 The urban-rural income gap: magnitude and trends

Table 1 gives average household per capita income for all of China and separately for urban and rural households. The statistics in this table exclude the migrant subsample, which is incorporated starting in section 5. Table 1 provides two measures of the urban-rural income gap, the ratio of urban to rural mean incomes (relative gap) and the difference between urban and rural mean incomes (absolute gap). For both 1995 and 2002 the first columns give income calculated according to the NBS definition, which excludes housing components of income. The second columns give NBS income plus housing components of income. These numbers are in current prices with no spatial price adjustments.⁸ Not surprisingly, adding in housing-related income increases mean incomes for both the rural and urban samples. Urban incomes increase more than rural incomes because the imputed value of urban owner-occupied housing and housing subsidies exceed those in rural areas. Consequently, including this component enlarges the urban-rural income ratio, by 10 per cent in 1995 and by 6 per cent in 2002. In ensuing sections, we only present findings calculated using the more complete measure of income that includes housing components of income.

⁸ The numbers in Table 1, including those that follow the NBS definition of income, are calculated using the CASS survey data. Due to differences in sample size and geographical distribution, our numbers using the NBS income definition differ somewhat from the numbers published by the NBS, which are calculated using data from the NBS household surveys. The NBS reports an urban-rural income ratio of 2.7 for 1995 and 3.3 for 2002 (NBS 2003).

Table 1: Mean household per capita incomes: national, urban, rural, and urban-rural gap (unit: yuan, except for the ratios)

	1995		2002		
	Unadjusted	PPP	Unadjusted	PPP	PPP, 1995 prices
National	2,969	2,596	5,930	5,121	4,686
Urban	5,878	4,379	10,396	7,913	7,240
Rural	1,779	1,866	3,063	3,329	3,046
Ratio of urban to rural	3.31	2.35	3.39	2.38	2.38
Urban minus rural	4,099	2,514	7,333	4,584	4,194

Notes: PPP numbers are adjusted for spatial price differences using the Brandt-Holz spatial costs of living estimates. The numeraire is the nationwide average cost of living for a joint basket of consumer goods, which we calculate as the weighted average of the Brandt-Holz mean urban and average rural costs of living. Weights are the urban and rural current population shares. Choice of population weights affects income levels somewhat, but not the ratios or inequality levels.

In the last column, 2002 incomes are deflated to 1995 prices using NBS consumer price indices for each provincial urban and rural location, and then adjusted for spatial price differences using the 1995 spatial cost of living estimates from Brandt-Holz. This is equivalent to first converting 2002 incomes into nationwide average PPP terms using the spatial cost of living estimates from Brandt-Holz, and then deflating using the CPI for the nationwide average cost of living between 1995 and 2002.

The urban-to-rural ratios for the last two columns are the same, because income values in these two columns differ by a constant factor, which is the deflation factor for the nationwide average cost of living between 1995 and 2002.

Whether including housing components or not, at current, unadjusted prices the urban-rural income ratio is substantial, close to or exceeding 3 in both years. This is high by international standards. Eastwood and Lipton (2004) give ratios for other Asian countries in the 1990s that fall between 1.3 and 1.8, with the Philippines a high outlier at 2.17. Similarly, Knight and Song (1999: 338) give urban-rural ratios for income and consumption in twelve countries, mostly in Asia but also in the Middle East and Africa. China's ratio exceeds those in all the other countries listed except Zimbabwe and South Africa. Note that most of the ratios for other countries reported in these sources include housing components of income but are not adjusted for spatial price differences.

The next columns give income adjusted to control for spatial differences in the cost of living. Yuan units in these columns reflect purchasing price parity with national average consumer prices over both urban and rural areas. We refer to incomes after adjustment for spatial price differences as purchasing price parity or PPP incomes. To allow comparison with 1995, for 2002 Table 1 also gives PPP incomes in constant 1995 prices. Adjustments for spatial price differences reduce the relative gap substantially because costs of living are higher in urban areas. According to Brandt and Holz's cost of living estimates, prices in urban areas were on average 36 per cent higher than in rural areas in 1995 and 39 per cent higher in 2002. With spatial price deflation the relative gap declines markedly from 3.1 to 2.2 in 1995 and from 3.2 to 2.3 in 2002. Even so, China's ratios remain relatively high by international standards.

Comparison of the PPP figures in constant prices (shown in the last columns for both 1995 and 2002) reveals that China's urban-rural income gap has increased little over

time. Between 1995 and 2002 the adjusted relative gap rose by a mere 1 per cent. The absolute gap, however, increased by 64 per cent from 2360 to 3867 yuan in 1995 constant prices.

China's urban-rural gap is not uniform regionally. As shown in Table 2, the relative gap is highest in the West, where in both 1995 and 2002 the unadjusted ratios exceeded 4, as compared to 3 or less for the center and east. As above, adjusting for spatial price differences greatly reduces the relative gaps. Urban/rural differentials in the cost of living are highest in the west, so that PPP adjustments narrow the gap more in the west than elsewhere. Nevertheless, even in PPP terms the west's urban-rural income ratios remain well above 3, as compared to around 2 in the center and east.

Table 2: Regional differences in income per capita and the urban-rural gap (unit: yuan, except for the ratio)

	1995		2002		
	Unadjusted	PPP	Unadjusted	PPP	PPP, 1995 prices
Western provinces	2,140	2,016	4,137	3,816	3,491
urban	5,036	4,260	8,582	7,344	6,719
rural	1,168	1,262	2,006	2,124	1,944
ratio of urban to rural	4.31	3.38	4.28	3.46	3.46
urban minus rural	3,868	2,998	6,576	5,220	4,776
Central provinces	2,240	2,177	4,555	4,384	4,011
urban	4,172	3,399	7,941	6,741	6,167
rural	1,559	1,747	2,652	3,059	2,799
ratio of urban to rural	2.68	1.95	2.99	2.20	2.20
urban minus rural	2,614	1,652	5,289	3,682	3,369
Eastern provinces	4,259	3,415	8,509	6,767	6,191
urban	7,498	5,109	13,013	9,006	8,240
rural	2,537	2,514	4,526	4,786	4,379
Ratio of urban to rural	2.96	2.03	2.88	1.88	1.88
Urban minus rural	4,961	2,595	8,487	4,220	3,861

Notes: See Table 1. Western provinces include Sichuan, Guizhou, Yunnan, Shaanxi and Gansu. Central provinces include Jiangsu, Zhejiang, Shandong and Guangdong. Eastern provinces include Beijing, Hebei, Liaoning, Jiangsu, Shandong and Guangdong.

Between 1995-2002 the relative gap rose in the west and center, but declined in the east. These trends in the west and center indicate that those parts of China where poverty is most concentrated are falling farther behind, at least in relative terms. Yet trends in the East, China's most developed region, hint that perhaps in the long term as China becomes more developed, the urban-rural gap could stabilize or even narrow.

4 The contribution of the urban-rural gap to inequality

The standard method of measuring the contribution of spatial income differences to inequality is decomposition of inequality by subgroup. Discussion of this approach and its application to the analysis of spatial inequality are available elsewhere (see Shorrocks 1984; Shorrocks and Wan 2005), so here we summarize only the main elements. Subgroup inequality decomposition is typically carried out using inequality indices from the entropy family. We employ two commonly used entropy measures, the Theil L (Mean Logarithmic Deviation) and the Theil T. The Theil L is defined as

$$I_{TL} = \frac{1}{n} \sum_{i=1}^n \ln \left(\frac{\mu}{y_i} \right) \quad (1)$$

and the Theil T as

$$I_{TT} = \frac{1}{n\mu} \sum_{i=1}^n \left[\ln \left(\frac{y_i}{\mu} \right) \right] y_i \quad (2)$$

where μ is mean income, y_i income of the i^{th} individual, and n the total number of individuals. These inequality indices can be decomposed among subgroups using the general formula

$$I = \sum_{g=1}^k w_g I_g + I(\mu_1, \mu_2, \dots, \mu_k) \quad (3)$$

where w_g is a weight attached to the g^{th} group, I_g inequality within the g^{th} group, and μ_g mean income of the g^{th} group. Equation (3) states that overall inequality is equal to the weighted sum of inequality within each subgroup plus inequality measured across mean incomes of the subgroups. The weighted sum of inequality within each subgroup is referred to as ‘within-group’ inequality. Inequality measured across mean incomes of the subgroups is referred to as ‘between-group’ inequality. Since we are interested in the contribution to inequality of the urban-rural income gap, we divide the sample into urban and rural subgroups. The contribution of the urban-rural income gap to inequality is the between-group component of the decomposition and equals inequality measured across mean incomes of the urban and rural groups.

Table 3 gives values of the two Theil indices and the results of inequality decompositions for 1995 and 2002. These are calculated using both unadjusted and PPP incomes. The overall level of inequality shows no clear trend between 1995 and 2002. The Theil L increases slightly, while the Theil T decreases. This is true regardless of whether incomes are adjusted for spatial price differences. The contrasting trends in the Theil L and Theil T indices arise because the underlying Lorenz curves for these two years cross.

Table 3: Equality decomposition by urban and rural subgroups

	1995				2002		
	Theil L		Theil T		Theil L		Theil T
	Unadjusted	PPP	Unadjusted	PPP	Unadjusted	PPP	Unadjusted
Total	0.381	0.274	0.416	0.298	0.389	0.288	0.373
Between	0.165	0.083	0.175	0.088	0.183	0.092	0.177
Within	0.216	0.192	0.241	0.210	0.206	0.196	0.196
Contribution of between and within effects (%)							
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Between	43.4	30.1	42.1	29.5	47.0	32.0	47.4
Within	56.6	69.9	57.9	70.5	53.0	68.0	52.6

Note: The PPP figures are comparable across years, because deflation involves multiplication by a constant, and the inequality indices and decomposition are scale invariant.

Adjustments for spatial price differences substantially reduce the level of overall inequality. The extent of the reduction is similar for the two indices. In 1995 the price adjustment reduces inequality by about 27 per cent and in 2002 by about 25 per cent. Roughly one quarter of inequality in unadjusted incomes, then, is attributable to spatial price differences. This finding is consistent with that of Ravallion and Chen (2004), who also find that correcting for spatial price differences reduces overall inequality.⁹ The fact that correcting for spatial price differences reduces inequality reflects that spatial price differences are positively correlated with levels of income.¹⁰

The lower half of Table 3 shows the percentages of inequality contributed by between-versus within-group inequality. The results for the Theil L and Theil T are very similar. For unadjusted incomes, between-group inequality contributes about 40 per cent of total inequality in 1995, increasing to 45 per cent in 2002. These numbers would suggest that the urban-rural gap is an increasingly important source of inequality, approaching half of the total.¹¹ Adjusting for spatial price differences, however, reduces the contribution of between-group inequality noticeably, to less than 30 per cent in 1995 and about 30-

⁹ Ravallion and Chen (2004) find that using the urban and rural poverty lines to correct for urban-rural differences in the cost of living reduces the Gini coefficient by about 12 per cent in both 1995 and 2001. This reduction is less than ours, but we use spatial price deflators that are more finely disaggregated, capturing and price differences between urban and rural areas in each province and also among provinces. Their spatial price deflation only differentiates between urban and rural areas for the nation as a whole. Also, they use a different inequality index. Note that Benjamin et al. (2005) provide estimates of the Gini coefficient with spatial deflation, but they do not provide the undeflated numbers for comparison.

¹⁰ The correlation between incomes and costs of living for 1995 is 0.92, and for 2002 0.85.

¹¹ Kanbur and Zhang (1999) also calculate the contribution of the urban-rural income gap to overall inequality for 1995 and report a between contribution of 71 per cent, much higher than our estimate. They calculate inequality using provincial-level data. Such an approach understates the importance of intra-urban and intra-rural inequality, as inequality among provincial means will always be lower than inequality among households or individuals.

32 per cent in 2002.¹² In real terms, then, perhaps one-third of all inequality is due to the urban-rural gap; furthermore, the contribution of the real income gap has increased somewhat over time.

Disaggregating by region provides further information on the structure underlying the urban-rural income gap's contribution to inequality. Table 4 gives inequality decompositions for each of the three regions. For simplicity, the table contains only results calculated using PPP incomes. The regional differences are marked. In the west, between-group inequality contributes roughly half of total inequality, as compared to less than a quarter in the east (bottom of Table 4). The center lies in between.

Table 4: Inequality decomposition of PPP incomes by urban and rural subgroups, eastern, central and western provinces

	1995		2002	
	Theil L	Theil T	Theil L	Theil T
Western provinces				
total	0.317	0.433	0.346	0.335
between	0.162	0.178	0.184	0.188
within	0.154	0.255	0.162	0.147
Central provinces				
total	0.164	0.163	0.214	0.213
between	0.047	0.051	0.076	0.077
within	0.117	0.112	0.138	0.136
East provinces				
total	0.264	0.255	0.227	0.218
between	0.060	0.062	0.050	0.049
within	0.204	0.193	0.178	0.169
Contribution of between and within effects (%)				
West provinces				
total	100	100	100	100
between	51	41	53	56
within	49	59	47	44
Central provinces				
total	100	100	100	100
between	29	31	35	36
within	71	69	65	64
Eastern provinces				
total	100	100	100	100
between	23	24	22	22
within	77	76	78	78

Note: See Tables 2 and 3.

¹² Adjusting for spatial price differences increases the contribution of between-group inequality, because less of between-group than within-group inequality is due to spatial price differences (as shown by the indices for between- and within-group inequality in the top half of Table 3).

Indeed, the absolute levels of between-group inequality in the east and center are relatively low (top of Table 4). In the east and center, then, if policy makers wish to reduce inequality, they should focus their efforts on income differentials within urban or within rural areas. The situation is different in western China. Overall inequality is markedly higher in the west, and the numbers in Table 4 suggest that the reason for this, and the distinguishing feature of inequality in the west, is the high level of between-group inequality. The level of between-group inequality in western China is two to three times that in the other regions, while within-group inequality is roughly similar to that in the other regions. Concerns about the urban-rural gap, then, should focus on western China.

5 Urbanization, migrants, and the rural-urban gap

During the reform period the level of urbanization in China has increased substantially. As shown in Table 5, according to official statistics the urban population share rose from about 26 per cent in 1990, to 29 per cent in 1995, and further to 39 per cent in 2002. This increase holds implications for estimates of the urban-rural gap and inequality. Mean incomes for urban and rural areas depend on who is classified as urban and rural. Also, in calculation of inequality using sample survey data, urban and rural samples are typically assigned weights based on their shares in the national population.

Table 5: Urbanization in China

	Urban population as % of total	Urban natural rate of increase
1990	26.41	1.043
1995	29.04	0.923
2000	36.22	0.510
2001	37.66	na
2002	39.09	na

Source: NBS (1996 and 2003); Chan and Hu (2003).

Growth in China's urban population is the result of two trends, natural increase in the urban population and reclassification of the rural population. Reclassification occurs when rural residents migrate to urban places and when rural places (and their resident populations) are reclassified as urban places.¹³ All of these mechanisms have

¹³ Reclassification can also occur if the definition of urban places changes, which in fact it has. The NBS adopted a new definition of urban places for the 2000 census that replaces the definition adopted for the 1990 census and used during the 1990s. This change in definition is fairly complex, and we refer interested readers to the literature for details (see, for example, Zhou and Ma 2003). Starting with the 2002 statistical yearbook, the NBS has been publishing data for the 1990s that is adjusted to conform to the new definition of urban places. Some recent studies, however, criticize the NBS adjustments and provide alternative population estimates (Chan and Hu 2003; Zhou and Ma 2003). In their thorough analysis, Chan and Hu (2003) conclude that the NBS number for the urban population in 1995 (29.04 per cent) is too low. They propose an alternative estimate of 31.72 per cent, almost three percentage points higher than the NBS number. Using Chan and Hu's alternative estimate for 1995, we have recalculated

contributed to China's urban population growth, but migration appears to be the most important. Chan and Hu (2003) note that the urban natural rate of increase has been low (Table 5). They estimate that in the 1990s the natural rate of increase of the urban population contributed only about one-third of total growth in the urban population. They estimate further that of total growth in the urban population in the 1990s, 22 per cent was due to reclassification of rural places. The remaining 55 per cent was due to migration. In other words, during the 1990s migration likely contributed more than half of China's urban growth.

How should the reclassification of rural into urban places be treated in analyses of inequality? Most would argue that place classifications should remain unchanged, that is, that an area classified as rural in one or more years should be counted as rural for the duration of the period under study, or that an area classified as urban in one or more years should be counted as urban for the duration. So, for example, residents of a rural area where farmland is converted to industrial and other nonagricultural uses would be counted as either rural or urban for the entire period. Chan and Hu (2003), however, point out that one reason for allowing classifications to change is that in some cases the reclassification is driven by migration from villages to towns prior to their redesignation as urban.

The NBS population statistics incorporate changes in the classification of rural places, and almost all studies of China follow suit primarily because the data required to keep classifications constant are unavailable. An exception is Benjamin et al. (2005), which provides an alternative estimate of the urban-rural income gap that keeps place classifications constant. That study concludes that reclassification tends to slow the convergence of mean incomes between urban and rural places, because reclassified rural areas tend to be those that have experienced the fastest income growth. Residents of the now richer, once-rural places are counted as urban, and residents of those places that grow more slowly and remain relatively poor continue to be counted as rural.

While in principle it would be desirable to redo our analysis holding place classifications constant, our data do not permit it. Benjamin et al. (2005) conduct the analysis using panel data, which makes it possible to keep the urban and rural classifications unchanged over time. The CASS survey data are not panel. In our analysis, then, we are constrained to rely on the standard NBS approach, under which rural places are reclassified as urban when they evolve to meet the criteria used to delineate urban places. As mentioned above, however, less than a quarter of the growth in China's urban population is thought to be caused by place reclassifications.

inequality levels and the contributions of between- and within-group inequality. Using these alternative estimates has little impact on the results, so in this paper we use the NBS population statistics for our calculations.

More important is the treatment of migrants. Rural-to-urban migration is the major factor causing China's rising levels of urbanization. Researchers universally agree that rural migrants who have moved to urban places should be counted in the urban population. Unfortunately, migrant populations are difficult to count, and estimates of inequality for China do not adequately incorporate this group. Most household surveys for China, including the NBS household survey, do not count the great majority of migrants because the sample frames are based on place of registration, not place of residence, and most migrants are not registered in the cities in which they live. Consequently, unregistered migrants living in cities are absent in urban survey samples.

Unregistered rural-to-urban migrants are also largely absent from the rural samples. When an entire rural household migrates, it is not included in the rural survey because no family member is present to be surveyed. Individuals who migrate without their families are often also not included in the rural sample. Rural surveys only count such individuals if they reside at home for a substantial portion of the year (more than six months in the NBS survey), or are the primary source of income for their rural households. These criteria apply to the CASS urban and rural subsamples, which follow the NBS sampling frame and use the same criteria for household membership.

In order to improve coverage of migrants, the CASS project team designed the 2002 survey so that it included a subsample of rural migrants living in urban areas. This migrant subsample contained 2,005 households and 5,327 individuals.¹⁴ Due to sampling frame limitations, the migrant sample was drawn from urban resident committees, that is, from standard urban residential neighborhoods.¹⁵ Migrants not living in such residential neighbourhoods—e.g., those living on construction sites, factories, and on the street—are not included in the sample. Included are migrants who live in urban residential neighborhoods, that is, migrant individuals and families that live in apartments or other urban housing or who rent rooms in such buildings, as well as those who live with urban families (e.g., nannies). This group includes both short- and long-term migrants, but likely contains a disproportionately high share of long-term migrants and also of migrant families. With these limitations in mind, the migrant subsample provides information that can be used to explore the effects of including migrants in analysis of the urban-rural gap.

In order to analyze the urban-rural gap, we must assign a weight to the migrant sample that reflects its share in the larger population. Information on the number of migrants in

¹⁴ These households were selected from all the provinces, but not from all the cities, in the urban survey. As rural-to-urban migrants are concentrated in large cities, all the provincial capital cities, plus one or two middle-sized cities in each of the provinces, were selected for the migrant survey. For more details about the migrant subsample, see Li et al. (2007).

¹⁵ The CASS migrant subsample is drawn from a sample frame based on urban addresses and regardless of place of origin and length of residence in the city, but with the restriction that the individual be registered as rural, not urban.

China's cities is weak, but a few studies provide estimates. Using data from China's 2000 census, in which efforts were made to count migrants in their place of residence, Liang and Ma (2004) estimate that in 2000 the migrant population resident in cities was equal to 13 per cent of the urban population.¹⁶ This estimate includes urban-to-urban migrant households and so may be high; however, the number of migrants likely increased between 2000 and 2002, so that Liang and Ma's number would understate the migrant population in the year of study here.

Mo (2004) gives estimates for 2002 based on a special, nationwide survey of rural households that included detailed questions about employment and labour movement. This study reports a number for rural-to-urban migrant workers equal to 16 per cent of the urban population.¹⁷ This number does not include dependents. The CASS migrant survey data gives 24 per cent of the members in migrant households as dependents, which we use to adjust the Mo (2004) figure. On this basis we obtain an estimate of the rural-to-urban migrant population equivalent to 21 per cent of the urban population. This estimate is probably high, as the dependency ratio in the CASS migrant subsample likely includes a disproportionate number of dependents. Based on this information, we conclude that in 2002 the migrant share in China's urban population likely fell between 13 and 21 per cent. In most of our calculations we use the mid-point in this range, 17 per cent, as the migrant share of the urban population. We also conduct sensitivity analyses to explore how the choice of the migrant population share affects the findings.

Table 6 gives mean household per capita incomes including migrant households, with migrants assumed to constitute 17 per cent of the urban population. After adjusting for spatial price differences, mean income per capita of migrant households is about 60 per cent below that of the urban registered and about 40 per cent above that of the rural sample.¹⁸ Not surprisingly then, including migrants reduces the size of the urban-rural gap. With migrants the relative gap is 2.12 (PPP incomes), as opposed to 2.27 without migrants (see Table 1). For both unadjusted and PPP incomes, including migrants reduces the relative gap by about 10 per cent.

Table 7 gives inequality levels and decompositions with and without the migrants, and also with and without PPP adjustments. The impact of including migrants is fairly modest. Including migrants does not substantially change the level of overall inequality for China. The Theil L registers a decrease and the Theil T a small increase, indicating

16 This estimate includes all inter-county migrants in cities and towns who have resided in their destination for 6 or more months and who do not have local household registration status.

17 In Mo (2004) migrants are defined as workers employed outside their township of residence for more than six months; movements for marriage, study, and to join the army are excluded. The 16 per cent figure includes only rural migrants employed in urban areas.

18 The gap in income between registered urban residents and rural-to-urban migrants does not fully capture the gap in their economic welfare, as migrants have little access to urban social services and on average work longer hours than do registered urban residents.

that the Lorenz curves cross. For both indices the level of between-group inequality declines and that of within-group inequality increases. Including migrants causes the between-group contribution to total inequality in PPP incomes to decline by about four percentage points. This change is smaller than that resulting from spatial price deflation.

Table 6: Mean household per capita incomes, including migrants, 2002 (unit: yuan, except for the ratios)

	Unadjusted	PPP
National	5,559	4,927
Urban	9,323	7,267
urban registered	9,996	7,773
urban migrant	6,083	4,831
Rural	3,143	3,425
Ratio of urban migrant to registered	0.61	0.62
Ratio of migrant to rural	1.64	1.61
Ratio of urban to rural	2.97	2.12

Notes: Population weights are rural 60.91 per cent, urban non-migrant 32.37 per cent, and urban migrant 6.72 per cent. These shares maintain the official urban/rural population shares for 2002, but now migrants constitute 17.2 per cent of the urban population (see Khan and Riskin 2005; Liang and Ma 2003). Price adjustments are explained in the notes to Table 1. Mean incomes for rural and urban subsamples differ slightly here from those in Table 1 because the original survey sample is randomly re-sampled to achieve the desired urban, rural and migrant proportions or weights, and also because including the migrant sample changes the distribution of observations among regions.

Table 7: Inequality decomposition with and without migrants, 2002

	Theil L		Theil T	
	Excl. migrants	Incl. migrants	Excl. migrants	Incl. migrants
	Unadjusted incomes			
Total	0.389	0.359	0.372	0.353
Between	0.183	0.147	0.177	0.144
Within	0.206	0.212	0.196	0.209
Contribution of between and within effects (%)				
Total	100.0	100.0	100.0	100.0
Between	47.0	41.0	47.4	40.9
Within	53.0	59.0	52.6	59.1
	PPP incomes			
Total	0.288	0.272	0.273	0.265
Between	0.092	0.071	0.092	0.071
Within	0.196	0.201	0.181	0.194
Contribution of between and within effects (%)				
Total	100.0	100.0	100.0	100.0
Between	32.0	26.0	33.8	26.9
Within	68.0	74.0	66.2	73.1

Notes: Migrants are included in the urban subsample and the decomposition is carried out between two groups, urban (including migrants) and rural.

The data in Table 7 show the impact of including both migrants and spatial price deflation on the inequality decomposition. With both these adjustments, overall inequality declines by about 25 per cent. This decline in overall inequality is due primarily to spatial price deflation. With both adjustments the level of between-group inequality falls by more than 50 per cent. Its contribution to total inequality declines from 45 per cent to only 26 per cent. Most of this decline is again due to spatial price deflation.

Table 8: Household characteristics of individuals in the regression samples, 1995 and 2002

Variable	Urban		Rural		Ratio of urban to rural
	Mean	Std dev.	Mean	Std dev.	
					1995
Income per capita	5880	6532	1779	1449	3.31
Income per capita (PPP)	4381	4988	1866	1361	2.35
Average education of working age adults	10.29	2.53	6.17	2.09	1.67
Average age of working age adults	39.63	7.89	35.49	5.74	1.12
Household size	3.37	0.88	4.74	1.38	0.71
% of household members of working age (16-65 yrs)	77.59	18.57	70.10	21.24	1.11
% of working age members in the party	22.43	28.94	5.26	13.29	4.26
% of family members of ethnic minority	3.33	14.86	5.52	18.82	0.60
Contracted farm land per capita (mu)	0	–	1.16	1.14	–
No. of observations	16,279				39,785
					2002
Income per capita	10396	6813	3064	2537	3.39
Income per capita (PPP)	7913	4635	3330	2671	2.38
Average education of working age adults	10.87	2.53	7.06	2.00	1.54
Average age of working age adults	40.56	7.29	37.08	6.55	1.09
Household size	3.25	0.85	4.42	1.24	0.74
% of household members of working age (16-65 yrs)	80.72	19.22	75.29	19.95	1.07
% of working age members in the party	23.90	29.57	7.16	15.63	3.34
% of working-age members in poor health	0.23	3.34	0.66	4.92	0.35
% of family members of ethnic minority	3.42	14.76	6.72	21.12	0.51
Contracted farm land per capita (mu)	0	–	1.36	1.70	–
No. of observations	21,103				32,874

Notes: The statistics in this Table are calculated over individuals rather than households. One can interpret them as weighted household averages, with the weights being the number of household members. The number of observations is the number of individuals in households surveyed, adjusted to correct for over-sampling of urban households in 1995 and of rural households in 2002. As urban households were over-sampled in the 1995 survey and rural households over-sampled in 2002 survey, the samples have been adjusted so that the proportion of urban to rural individuals is equal to the national averages as given by the NBS in each of the two years. This was done by increasing the number of observations in the under-represented sector through random sampling of the original sample for the under-represented sector. Income values shown here are in current yuan. Information on health was not collected in 1995. Health status is self-reported.

Table 8 gives sensitivity analyses using different migrant shares in the urban population. All numbers here are based on the PPP incomes. For purposes of illustration, we use the following shares: 0 per cent, the share implicitly assumed by analyses that do not include migrants; a low estimate of 13 per cent; our mid-range choice of 17 per cent; and a high estimate of 21 per cent. We also show results for a yet higher share of 25 per cent so as to show the possible impact of an even larger migrant population. As expected, increasing the migrant share reduces the urban-rural income ratio. With zero migrants the ratio is 2.27. With a migrant share of 25 per cent, the ratio falls to 2.06. This ratio remains fairly high by international standards.

Overall inequality shows no clear upward or downward trend as the migrant population share increases. The Theil L and decreases and the Theil T increases. Between-group inequality declines steadily as the migrant share rises. Yet the impact of changes in the migrant share is not overly large. A near doubling in the migrant share from 13 per cent to 25 per cent causes the contribution of between inequality to fall by fewer than 4 percentage points, from about 27 per cent to 24 per cent. These results demonstrate that including migrants can have an impact on measured patterns of inequality. The impact, however, is fairly modest, even when using relatively high estimates of the size of the migrant population. This might reflect that migrants tend to have characteristics more similar to urban residents (younger, better educated, smaller households), so that movement of this subset of the rural population does not reduce the urban-rural gap as much as would movement of ‘average’ rural residents. Note, moreover, that these calculations hold constant the incomes of other groups. That is, the calculations do not take into account the fact that migration can affect the incomes of those remaining behind in rural areas as well as of the registered urban population.

6 Factors underlying the urban-rural income gap

Even after the adjustments outlined above, the urban-rural income gap in China remains relatively large and contributes substantially to overall inequality. The gap reflects a variety of factors, including differences in household characteristics and also in economic environments and policies. The Oaxaca-Blinder decomposition provides an empirical methodology for investigating some of the factors that underlie the gap. This method allows us to calculate the extent to which income differences between the urban and rural groups reflect differences in individual characteristics as opposed to other factors. We carry out the decomposition first without migrants and then with migrants included in the urban sample.

The Oaxaca-Blinder method is often used to analyze differences in earnings or the returns to labour. Here we use it to analyze differences in per capita income, including both labour earnings and other income. We analyze income rather than labour earnings because a large portion of income in China, especially in urban areas, is non-labour income. In addition, non-labour income accounts for well over 40 per cent of the

absolute gap between urban and rural incomes (Table 9). Analysis of labour earnings alone would therefore miss much of the story. Table 9 shows the composition of urban and rural incomes. Labour earnings include wages and net income from self employment, and non-labour income includes asset income, pensions, net government transfers, housing components of income, and private transfers and remittances. Of non-labour income, pensions, net government transfers, and housing income have been most important. Net government transfers declined substantially between 1995 and 2002.

Table 9: Per capita income OLS regression estimates (dependent variable: Ln household per capita income, unadjusted and PPP)

Variable	1995		2002	
	Urban	Rural	Urban	Rural
Average education of working-age adults	0.0437***	0.0510***	0.0313***	0.0458***
Education squared	0.0001	-0.0019***	0.0023***	0.0009*
Average age of working-age adults	0.0419***	0.0427***	0.0207***	0.0306***
Age squared	-0.0004***	-0.0005***	-0.0001**	-0.0004***
Household size	-0.2838***	-0.2076***	-0.3400***	-0.2318***
Household size squared	0.0201***	0.0102***	0.0288***	0.0129***
% of household members of working age (16-65 yrs)	0.0045***	0.0037***	0.0030***	0.0031***
% of working-age members in the party	0.0013***	0.0024***	0.0024***	0.0033***
% of working-age members in poor health			-0.0036***	-0.0028***
% of family members of ethnic minority	-0.0019***	-0.0012***	0.0012***	0.0011***
Contracted farm land per capita (mu)		0.0346***		0.0349***
Land squared		-0.0045***		-0.0003*
No. of observations	16,279	39,785	21,103	32,874
F-statistic (unadjusted)	713.37	918.10	821.42	747.11
Adjusted R ² (unadjusted)	0.45	0.40	0.44	0.41
F-statistic (PPP)	457.82	787.42	575.51	803.39
Adjusted R ² (PPP)	0.35	0.36	0.35	0.42

Notes: ***indicates significance at the 1% confidence level, **at 5% and *at 10%. Spatial price adjustments do not affect the estimated coefficients of the variables shown here, only those of the provincial dummy variables and constant term. Spatial price adjustments also affect the regressions' explanatory power somewhat, so we provide F-statistics and adjusted R²s for both cases. The constant term and estimated coefficients for provincial dummies are not shown due to space limitations. These coefficients were, for the most part, significant. Observations represent individuals rather than households. The number of observations is the number of individuals in households surveyed, adjusted to correct for over-sampling of urban households in 1995 and over-sampling of rural households in 2002 (see text). The percentage of working-age adults that is male was included as an explanatory variable in an initial regression, but was not significant so was dropped. Information on health was not collected in 1995. Health status is self-reported.

The Oaxaca-Blinder decomposition requires two steps. The first step is to estimate income equations separately for the two groups. These equations typically take the form $\ln(y^g) = \alpha_g + \beta_g X^g + \varepsilon^g$ for $g = u, r$ (4)

where g indicates the group (urban or rural here), y is a vector of per capita incomes of individuals, and X a matrix of individual characteristics.

The second step is to use the regression results to decompose the difference in mean incomes between the groups. The difference in mean log incomes between the higher income urban and lower income rural group can be written as

$$\overline{\ln y^u} - \overline{\ln y^r} = (\hat{\alpha}_u - \hat{\alpha}_r) + (\hat{\beta}_u \bar{X}^u - \hat{\beta}_r \bar{X}^r) = (\hat{\alpha}_u - \hat{\alpha}_r) + \hat{\beta}_u (\bar{X}^u - \bar{X}^r) + (\hat{\beta}_u - \hat{\beta}_r) \bar{X}^r \quad (5)$$

The first term in the right-hand side of equation (5) gives the portion of the urban-rural income gap due to differences in the constants. The second term gives the portion due to differences between the two groups in their endowments of characteristics. The third term is the portion due to differences in the estimated regression coefficients or ‘returns’ to characteristics. The first and third terms are typically considered the ‘unexplained’ portion of the gap, and the second term the ‘explained’ portion of the gap.

Equation (5) uses the coefficients of the richer (urban) group as weights for the differences in characteristics and uses the mean poorer (rural) characteristics as the weights for the differences in coefficients. This is the standard approach. The reverse decomposition would be

$$\overline{\ln y^u} - \overline{\ln y^r} = (\hat{\alpha}_u - \hat{\alpha}_r) + \hat{\beta}_r (\bar{X}^u - \bar{X}^r) + (\hat{\beta}_u - \hat{\beta}_r) \bar{X}^u \quad (6)$$

This reverse decomposition uses the rural coefficients to weight the differences in characteristics and uses mean urban characteristics to weight the differences in coefficients. We present results for both the standard and reverse decompositions.

6.1 Estimation of the income equations for the urban and rural subgroups

A variety of characteristics can influence per capita household incomes (Miles 1997; Gustafsson and Li 1998, 2001: 44-83; Knight and Song 1999; Morduch and Sicular 2000). These include demographic characteristics such as household size, the proportion of dependents versus working-age household members, the ethnic composition of household members, and the age of household members. The education of household members may also be important, as it influences the returns to labour and also to some assets. Household assets generate income. Holdings of many assets, however, are dependent on the level of household income and so endogenous. In China an important asset that is not dependent on the level of household income is farm land allocated to households by villages under the household responsibility or contracting system. Such land is allocated administratively by the village or township on the basis of household size, and reallocations are infrequent.

Another set of factors considered potentially important in explaining household incomes in China is political status and connections (Bian and Logan 1996; Morduch and Sicular 2000; Lam 2003). Political status and connections are difficult to measure directly, but may be associated with the presence of a communist party member or cadre within the household. Here we focus on Party membership, as cadre status is often attached to employment, and so disentangling the extent to which political connections as opposed to the wages from cadre employment explain income is difficult. Note that Party membership's relationship with income could reflect not only political connections, but also unobserved ability or ambition that may be associated with Party membership (Gerber 2000; Lam 2003). Finally, location of residence is commonly thought to affect income levels, especially in China where mobility is limited. Here we include provincial dummy variables to capture the effects of location.

Table 10a: Household characteristics of individuals in the regression samples, 1995

Variable	Urban		Rural		Ratio of urban to rural
	mean	standard deviation	mean	standard deviation	
income per capita	5633	6444	1810	1462	3.11
income per capita (PPP)	4256	4938	1898	1373	2.24
average education of working-age adults	10.27	2.54	6.20	2.07	1.66
average age of working-age adults	39.48	7.88	35.50	5.74	1.11
household size	3.37	0.88	4.73	1.37	0.71
% of household members of working age (16-65)	77.48	18.54	70.15	21.24	1.10
% of working-age members in the Party	22.32	28.79	5.28	13.31	4.23
% of family members that are ethnic minority	3.35	14.94	5.12	18.20	0.66
contracted farm land per capita (mu)	0.0	0.0	1.17	1.13	-
no. of observations	21378		34682		

Tables 10a and 10b present descriptive statistics on per capita income and household characteristics for the urban and rural subsamples. For 2002 we give two columns of urban statistics, one excluding and one including the migrant subsample (assumed to constitute 17 per cent of the urban population). Household characteristics differ noticeably between the urban and rural groups, suggesting that these variables explain at least part of the urban-rural income gap. Mean education for the urban sample is about 50 per cent higher than that for the rural sample. Urban households tend to be older and smaller, and they contain proportionately more working-age members. They also have a higher incidence of Party membership and a lower proportion of members with poor health or minority ethnicity. Only rural households have farm land.

Table 10b: Household characteristics of individuals in the regression samples, 2002

Variable	Urban without migrants		Urban with migrants		Rural		Ratio of urban (no migrants) to rural	Ratio of urban (with migrants) to rural
	mean	standard deviation	mean	standard deviation	mean	standard deviation		
Income per capita	9937	6473	9346	6533	3145	2597	3.16	2.97
Income per capita (PPP)	7712	4503	7268	4650	3426	2725	2.25	2.12
average education of working-age adults	10.85	2.56	10.35	2.77	7.11	1.97	1.53	1.46
average age of working-age adults	40.49	7.37	39.44	7.64	37.07	6.56	1.09	1.06
household size	3.25	0.84	3.21	0.86	4.41	1.24	0.74	0.73
% of household members of working age (16-65)	80.35	19.17	79.75	19.33	75.47	19.92	1.06	1.06
% of working-age members in the Party	23.92	29.66	20.43	28.64	7.20	15.67	3.32	2.84
% of working-age members in poor health	0.26	3.71	0.61	5.05	0.65	4.90	0.40	0.93
% of family members that are ethnic minority	3.48	14.83	4.06	16.24	6.08	20.23	0.57	0.67
contracted farm land per capita (mu)	0.0	0.0	0.0	0.0	1.36	1.73	-	-
no. of observations	20,173		25,272		33,803			

Notes: The statistics in these tables are calculated over individuals rather than households. One can interpret them as weighted household averages, with the weights being the number of household members. The number of observations is the number of individuals in households surveyed. Urban means including migrants are calculated using a migrant share in the urban population of 17%. Income values shown here are in current yuan. Information on health was not collected in 1995. Health status is self-reported.

The regression results for urban and rural income equations appear in Table 11. Estimation is carried out using OLS. We include squared terms for education, age, household size, and land to allow for potential nonlinearities. Spatial price adjustments do not affect the estimated coefficients for the variables shown in this table, but they alter the estimated constant term and coefficients for the provincial dummy variables (not shown). They also influence the overall explanatory power of the equations. We report F-statistics and adjusted R^2 statistics for both cases.

Table 11: Contributions of individual explanatory variables to the PPP urban-rural gap, standard and reverse decomposition (%)

	Standard decomposition (%)					
	1995			2002		
	Total	Endowment	Coefficient	Total	Endowment	Coefficient
Average education of working age adults	25.9	21.1	4.8	26.8	29.6	-2.9
Average age of working age adults	20.4	5.5	15.0	3.9	4.7	-0.7
Household size	2.5	16.1	-13.6	-2.9	12.2	-15.0
% of household members of working age	10.7	3.8	6.9	0.8	1.7	-0.8
% of working-age members in the party	2.0	2.6	-0.6	3.5	4.2	-0.7
% of working-age members in poor health				0.1	0.2	-0.1
% of family members of ethnic minority	0.0	0.4	-0.4	-0.3	-0.4	0.1
Contracted farm land per capita (mu)	-3.1	0.0	-3.1	-4.9	0.0	-4.9
	Reverse decomposition (%)					
Average education of working age adults	25.9	8.4	17.5	26.8	24.9	1.9
Average age of working age adults	20.4	1.1	19.4	3.9	0.6	3.3
Household size	2.5	17.8	-15.4	-2.9	15.4	-18.2
% of household members of working age	10.7	3.0	7.7	0.8	1.8	-1.0
% of working-age members in the party	2.0	4.6	-2.6	3.5	5.8	-2.3
% of working-age members in poor health				0.1	0.1	0.0
% of family members of ethnic minority	0.0	0.3	-0.2	-0.3	-0.4	0.0
Contracted farm land per capita (mu)	-3.1	-3.1	0.0	-4.9	-4.9	0.0

Notes: For education, age, household size, and land, the contributions shown are the sum contributions of the linear and squared terms. Due to rounding, numbers do not always add up exactly.

The estimated coefficients are almost all highly significant and for the most part of the expected signs. The coefficients clearly differ between the urban and rural samples. Including migrants moves most of the urban coefficients values closer to those of the rural sample, although differences remain. For example, in 2002 the relationship between education and income in rural areas is close to linear, while for the urban sample with and without migrants the marginal returns to education increase with the level of education. Also, the returns to Party membership are higher in rural areas than in urban areas. Differences in returns to characteristics, then, may also contribute to the urban-rural income gap.

6.2 Decomposition of the urban-rural wage gap

Tables 12a-c contain summary results from Oaxaca-Blinder decompositions of the urban-rural income gap for 1995 and 2002. Both the standard and reverse decompositions are shown, and the decomposition is carried out both for unadjusted and PPP incomes. For 2002 we present decomposition results with migrants included in the urban sample. As explained in the table notes, the Oaxaca-Blinder method cannot identify the separate contributions of the constant term and indicator variables. Therefore the Tables give only the sum contribution of the constant and provincial dummy variables.

Table 12a: Decomposition of the difference between mean urban (excluding migrants) and rural incomes, 1995

	Standard Decomposition		Reverse Decomposition	
	unadjusted	PPP	unadjusted	PPP
Difference in ln incomes	1.169	0.848	1.169	0.48
Contributions to difference (values):				
Constant term & provincial dummies	0.708	0.387	0.708	0.387
Other explanatory variables, of which:	0.461	0.461	0.461	0.461
coefficients	0.020	0.020	0.174	0.174
endowments	0.441	0.441	0.286	0.286
Contributions to difference (%):				
Constant term & provincial dummies	60.6	45.6	60.6	45.6
Other explanatory variables, of which:	39.4	54.4	39.4	54.4
coefficients	1.7	2.4	14.9	20.5
endowments	37.7	52.0	24.5	33.7

Table 12b: Decomposition of the difference between mean urban (excluding migrants) and rural incomes, 2002

	Standard Decomposition		Reverse Decomposition	
	unadjusted	PPP	unadjusted	PPP
Difference in ln incomes	1.205	0.887	1.205	0.887
Contributions to difference (values):				
Constant term & provincial dummies	1.039	0.722	1.039	0.722
Other explanatory variables, of which:	0.165	0.165	0.165	0.165
coefficients	-0.313	-0.313	-0.238	-0.238
endowments	0.479	0.479	0.405	0.405
Contributions to difference (%):				
Constant term & provincial dummies	86.2	81.4	86.2	81.4
Other explanatory variables, of which:	13.7	18.6	13.7	18.6
coefficients	-26.0	-35.3	-19.8	-26.8
endowments	39.8	54.0	33.6	45.7

Table 12c: Decomposition of the difference between mean urban (including migrants) and rural incomes, 2002

	Standard Decomposition		Reverse Decomposition	
	unadjusted	PPP	unadjusted	PPP
Difference in ln incomes	1.111	0.793	1.111	0.793
Contributions to difference (values):				
Constant term & provincial dummies	0.484	0.168	0.484	0.168
Other explanatory variables, of which:	0.625	0.625	0.625	0.625
coefficients	0.192	0.192	0.272	0.272
endowments	0.435	0.435	0.354	0.354
Contributions to difference (%):				
Constant term & provincial dummies	43.6	21.2	43.6	21.2
Other explanatory variables, of which:	56.3	78.8	56.3	78.8
coefficients	17.3	24.2	24.5	34.3
endowments	39.2	54.9	31.9	44.6

Notes: Numbers in these tables may not add up exactly due to rounding. The standard decomposition weights endowment differences between the two groups by the urban group's estimated coefficients and weights differences in coefficients by rural mean endowments. The reverse decomposition weights endowment differences by the rural group's coefficients and weights differences in coefficients by urban mean endowments. Some explanatory variables are uniformly equal to zero for the urban subgroup. These variables include a few provincial dummy variables and, importantly, farm land. Urban households have no contracted land. In principle, the contributions of these variables should be attributed entirely to differences in endowments, as is done by the reverse decomposition. The standard decomposition attributes the contributions of these variables entirely to differences in coefficients, which does not make much sense. We therefore prefer the reverse decomposition results. As discussed by Jones (1983) and Oaxaca and Ransom (1999), when dummy variables are included in the regression equations, the constant terms and the coefficients of the dummy variables will depend on the choice of reference group or groups for the dummy variables. For this reason, identification of the separate contributions of the constant terms and dummy variables is impossible in the decomposition, and we do not present them separately.

We begin with discussion of the results without migrants. Here we are mainly interested in the results for PPP incomes. As most studies do not adjust for spatial price differences, however, some comments about how spatial price adjustments affect the results may be of interest. As noted above, spatial price deflation reduces the urban-rural gap. In 2002, for example, the gap in unadjusted ln incomes is 1.205 and in adjusted ln incomes 0.887 (Table 12b). In the decompositions, this reduction in the gap is fully matched by the reduction in the sum contribution of the constant term and provincial dummy variables. That is, correcting for spatial price differences only affects the contributions of the constant term and provincial dummy variables and does not affect the contributions of other explanatory variables. This reflects the fact that adjusting for spatial prices only alters mean differences among locations, which differences are captured by the constant term and provincial dummy variables.

While correcting for spatial price differences does not change the absolute size of the contributions of non-geographic explanatory variables, it increases their proportional contributions to the gap. In 2002, for example, spatial deflation increases the percentage contribution of non-geographic explanatory variables from 14 per cent to 19 per cent in

2002. More generally, if incomes are not adjusted for spatial price variation, the proportional contribution of location and the constant term will be overstated, and of other explanatory variables such as age, education, and so on, understated.

For PPP incomes, in 1995 about 46 per cent of the urban-rural gap in ln incomes was due to location dummy variables and the constant term, and 54 per cent due to differences in the returns to and endowments of non-geographic explanatory variables. This contribution of non-geographic explanatory variables dropped markedly between 1995 and 2002, from 54 per cent to only 19 per cent. This decline is due to the fact that in 2002 the differences in coefficients between the two groups had become substantially inequality decreasing. In 1995, differences in the coefficients widened the gap, while in 2002 they reduced the gap by more than 25 per cent. As shown in Table 13b, this negative contribution of the 2002 coefficients is largely attributable to the returns on household size, which are more negative in urban than rural areas. The returns to most other variables are higher in urban than in rural areas.

Endowments of non-geographic household characteristics contribute between one third and one half of the urban-rural income gap. Tables 13a-c give the separate contribution of each such characteristic to the income gap. Education endowments have a large contribution, especially in 2002 when they account for more than one quarter of the income gap. All else held constant, if average education levels in rural areas were increased to be on a par with those in urban areas, then in 2002 the urban-rural income gap would decline by 26 to 30 per cent. The only other characteristic for which endowments have a large contribution to the income gap is household size. In 1995 differences in household size account for 16-19 per cent of the gap, and in 2002 for 13-16 per cent of the gap. This endowment effect, however, is offset by the negative contribution of differences in the coefficient on household size.

The contributions of the endowments of most other variables are small. Farm land endowments of rural households reduce the income gap by 5 per cent or less in both years. The higher incidence of Party membership in urban China increases the gap by 6 per cent or less in both years.

Including migrants in the urban sample has a noticeable impact on the decomposition results (Tables 12c and 13c). Including migrants reduces the per cent contribution of the constant term and provincial dummy variables markedly. The contribution of the coefficients, however, increases and becomes positive. The contribution of endowments remains relatively unchanged.

Table 13a: Contributions of individual explanatory variables to the PPP urban-rural gap, 1995 (%)

	Total	Standard Decomposition		Reverse Decomposition	
		endowment	coefficient	endowment	coefficient
Average education of working-age adults	30.8	22.9	7.1	9.4	20.6
average age of working-age adults	15.1	5.7	9.4	1.1	14.0
household size	-2.8	16.0	-18.9	18.5	-21.5
% of household members of working age	12.9	4.1	8.7	3.2	9.7
% of working-age members in the party	2.4	2.9	-0.6	4.7	-2.4
% of family members that are ethnic minority	0.3	0.3	-0.0	0.2	0.0
contracted farm land per capita (mu)	-3.3	0.0	-3.4	-3.4	0.0

Table 13b Contributions of individual explanatory variables to the PPP urban-rural gap, 2002 (%)

	Total	Standard Decomposition		Reverse Decomposition	
		endowment	coefficient	endowment	coefficient
average education of working-age adults	25.4	30.3	-5.0	26.2	-0.8
average age of working-age adults	-3.3	4.8	-8.0	0.7	-3.9
household size	-2.7	13.1	-15.8	16.3	-18.9
% of household members of working age	0.8	1.7	-0.9	1.8	-1.0
% of working-age members in the party	3.7	4.4	-0.7	6.1	-2.4
% of working-age members in poor health	0.1	0.1	0.0	0.1	0.0
% of family members that are ethnic minority	-0.1	-0.4	0.3	-0.2	0.2
contracted farm land per capita (mu)	-5.3	0.0	-5.3	-5.3	0.0

Table 13c Contributions of individual explanatory variables to the PPP urban-rural gap including migrants, 2002 (%)

	Total	Standard Decomposition		Reverse Decomposition	
		endowment	coefficient	endowment	coefficient
average education of working-age adults	39.2	31.8	7.4	24.5	14.7
average age of working-age adults	18.0	3.8	14.2	0.2	14.1
household size	17.3	13.4	3.9	15.4	-2.0
% of household members of working age	7.2	2.0	5.3	1.6	5.5
% of working-age members in the Party	3.4	4.0	-0.6	5.0	-1.6
% of working-age members in poor health	0.0	0.0	0.0	0.0	0.0
% of family members that are ethnic minority	-0.3	-0.1	-0.1	-0.3	-0.1
contracted farm land per capita (mu)	-5.9	0.0	-5.9	-5.9	0.0

Notes: The notes to Table 12 apply. For education, age, household size, and land, the contributions shown are the sum contributions of the linear and squared terms. Due to rounding, numbers do not always add up exactly.

7 Conclusion

In this paper we have explored China's urban-rural income gap. Several key findings emerge. First, China's urban-rural income gap is large by international standards, even after various adjustments such as fuller measurement of income, spatial price deflation and including migrants in the urban sample. Still, these adjustments, especially spatial price deflation, reduce the size of the gap substantially. With respect to trends over time, we find the adjusted relative gap widened little between 1995 with 2002. This conclusion differs from that reported elsewhere.^{19, 20}

Second, the contribution of the urban-rural income gap to overall inequality has been relatively large and has increased somewhat, although again its level is reduced by the adjustments. If we use 2002 unadjusted incomes and include migrants, between-group

¹⁹ Note that the Brandt and Holz spatial price indices likely understate the difference in housing prices between urban and rural areas, and perhaps increasingly so over time if urban areas have experienced speculative housing bubbles.

²⁰ As mentioned earlier, our income variable does not include the value of household consumption of subsidized public services because no data are available on such consumption. Inclusion of this component would likely increase the size of the urban-rural income gap. It would likely also affect measured trends in the gap, although the direction of the effect is not clear. Statistics for welfare indicators such as infant mortality rates, life expectancies, and education levels give a mixed picture. Overall, however, they do not indicate deterioration in the relative status of the rural population between 1995 and 2002. See China Development Research Foundation and UNDP (2005), National Bureau of Statistics (2003), World Bank (2003), and Zhang and Kanbur (2005).

inequality contributes more than 40 per cent of overall inequality. If we further correct for spatial price differences, the contribution declines to 26 per cent. With or without adjustments, this contribution is large relative to that in other countries. Shorrocks and Wan (2005) review international evidence on this question. Citing available studies based on household-level data and using similar methodology to that used here, they report that the contribution of the urban-rural gap ranges from less than 20 per cent in Greece to 26-30 per cent in the Philippines. Eastwood and Lipton (2004) give estimates for earlier years for developing countries. Excepting China, in all cases the contribution of the urban-rural gap is less than 25 per cent of total inequality.²¹ All of these estimates are calculated using nominal prices, unadjusted for spatial price differences. Our unadjusted estimates for China exceed the highest numbers for other countries by ten percentage points.

Third, regional differences in China's urban-rural gap are large. The urban-rural income gap is much larger in western China than in the eastern or central regions, as is its contribution to inequality. Indeed, the urban-rural gap's contribution to overall inequality in the east and center is fairly small. These regional differences suggest that efforts to bridge the urban-rural divide should target the west. Further research is required to identify what sorts of targeted interventions would be most effective, but some recent studies provides complementary evidence. Fan et al. (2004) and Zhang and Fan (2004) examine the impact of public investments on regional poverty and inequality in GDP per capita. Their findings suggest that public investment targeted to western China would have the most impact, especially investments in rural education and agricultural research and development. Investments in irrigation and poverty loans, however, would be less effective.

Our analysis highlights several measurement issues. One is spatial differences in prices and the cost of living, which has a substantial impact on the measured size of the gap and its contribution to inequality. Our findings here parallel those in Brandt and Holz (2004). As the study of income inequality is ultimately interested in real differences in incomes, spatially adjusted estimates of the urban-rural gap and its contribution are most meaningful.

A second measurement issue is the delineation of urban versus rural populations. Here various problems arise, but probably most important for China is the treatment of migrants. Including migrants in the urban sample reduces the size of the urban-rural income gap and that gap's contribution to inequality, but only modestly. Including migrants has little impact on the overall level of inequality, because lower between-group inequality is offset by higher within-group inequality. Migration increases inequality within urban areas, which brings with it a new set of challenges.

²¹ For China they refer to a study by Zhang (1997), which gives a contribution of 38 per cent in 1988. This contribution is comparable to the unadjusted contribution in this study. Zhang's estimate for 1988 is calculated using household data from an earlier round of the CASS survey.

Further research and better data are needed to fully explore the impact on inequality measurement of including migrants, but these results provide some indication of the magnitude and direction. The impact is, however, noticeably smaller than that of correcting for spatial price differences. Efforts to improve information on geographic price differences, then, are equally important.

What explains the urban-rural gap? Differences in endowments of household characteristics contribute roughly half the gap in PPP ln income. Most important here is education. Differences in education levels contribute 25-30 per cent of the gap. These estimates imply that, all else equal, if rural education levels were increased to be on a par with urban levels, the urban-rural income gap would decline by 25-30 per cent.

Location of residence contributes the other half of the PPP income gap. Here location's contribution is defined as the sum contribution to the gap of differences between urban and rural areas in the constant terms, coefficients on provincial dummy variables, and coefficients on household and individual characteristics. Spatial price deflation makes a difference here, reducing location's contribution by more than 10 percentage points. Over time the contribution of location declines somewhat, which is consistent with increased mobility and market integration.

Our analysis points to the need for further research in several areas. One is education. Studies on education in China generally report large differences in the levels of education not only between urban and rural areas but also among provinces (Hannum et al. 2005) Such spatial differences in education likely reflect multiple factors, including differences in incomes, in public expenditures on education, and in patterns of migration. Evidence provided here and elsewhere suggests that the private returns to education are also lower in rural areas (Cai et al. 2005; Yue et al. 2007). Further information is needed to understand why private returns to education differ geographically. To what extent, for example, do such geographic differences reflect differences in the industrial structure of employment and specificity of human capital? To what extent might they reflect correlation with unobserved community or individual characteristics?

A second topic for further research is spatial location. Why, after controlling for observed characteristics, does location of residence remain so important in explaining income differences? The *hukou* or household registration system and related policies that continue to hinder rural-to-urban movement are obvious culprits. Yet the persistence of urban-rural gaps in other countries suggests that even without such artificial restrictions, migration is unlikely to eliminate the urban-rural income gap or to equalize the returns to education and other individual characteristics.

China's urban-rural income gap has shown little sign of declining despite substantial easing of the restrictions on migration and the growing number of migrants. A variety of

factors could contribute to the persistence of spatial differences. One factor is non-labour income, which accounts for nearly half the income gap. Migration is not likely to reduce gaps in some forms of non-labour income such as housing-related income and pensions. Also, migration may not be able to eliminate the gap because variables other than income may affect decisions to move. Other relevant factors include access to community networks and support systems, farm labour requirements, job discrimination, incomplete information about living conditions and employment opportunities, higher costs of living (especially housing) in cities, and access to schooling and other public services.

References

- Atkinson, A.B. and F. Bourguignon (2000). 'Introduction: Income Distribution and Economics', in A.B. Atkinson and F. Bourguignon (eds) *Handbook of Income Distribution* Vol. 1, Elsevier North Holland: Amsterdam.
- Benjamin, D., L. Brandt, J. Giles, and S. Wang (2005). 'Income Inequality in the Transition Period', in L. Brandt and T.G. Rawski (eds) *China's Great Transition: Origins, Mechanisms and Consequences of the Post-Reform Economic Boom*, Department of Economics, University of Pittsburgh: Pittsburgh.
- Bian, Y., and J.R. Logan (1996). 'Market Transition and the Persistence of Power: The Changing Stratification System in Urban China', *American Sociological Review* 61(5): 739-58.
- Brandt, L., and C.A. Holz (2004). 'Spatial Price Differences in China: Estimates and Implications', mimeo, University of Toronto: Toronto.
- Cai, F., A. Park, and Y. Zhao (2005). 'The Chinese Labour Market', in L. Brandt and T.G. Rawski (eds) *China's Great Transition: Origins, Mechanisms and Consequences of the Post-Reform Economic Boom*, Department of Economics, University of Pittsburgh: Pittsburgh.
- Chan, K.W. and Y. Hu (2003). 'Urbanization in China in the 1990s: New Definition, Different Series, and Revised Trends', *China Review* 3(2): 49-71.
- Chan, K.W. and L. Zhong (1999). 'The Hukou System and Rural-Urban Migration in China', *China Quarterly* 160: 818-65.
- China Development Research Foundation and UNDP (2005), China Human Development Report 2005: Towards Human Development with Equity, www.hdr.undp.org/docs/reports/national/CPR_China/China_2005_en.pdf
- Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Johns Hopkins University Press: Baltimore.
- Eastwood, R., and M. Lipton (2000). 'Rural-Urban Dimensions of Inequality Change', *WIDER Working Paper* 200, UNU-WIDER: Helsinki.

- Eastwood, R., and M. Lipton (2004). 'Rural and Urban Income Inequality and Poverty: Does Convergence between Sectors Offset Divergence within Them?', in G.A. Cornia, *Inequality, Growth and Poverty in an Era of Liberalization and Globalization*, Oxford University Press for UNU-WIDER: Oxford.
- Fan, S., L. Zhang, and X. Zhang (2004). 'Reforms, Investment and Poverty in Rural China', *Economic Development and Cultural Change* 52(2): 395–421.
- Gerber, T.P. (2000). 'Membership Benefits or Selection Effects? Why Former communist party Members Do Better in Post-Soviet Russia', *Social Science Research* 29(1): 25-50.
- Gradín, C., O. Cantó, and Coral del Rio (2004). 'Inequality, Poverty and Mobility: Choosing Income or Consumption as Welfare Indicators', *Instituto de Estudios Fiscales Working Paper* 18/04, Universidade de Vigo: Vigo.
- Gustafsson, B., and S. Li (1998). 'Inequality in China at the End of the '80s: Locational Aspects and Household Characteristics', *Asian Economic Journal* 12(1): 35-63.
- Gustafsson, B., and S. Li (2001). 'A More Unequal China? Aspects of Inequality in the Distribution of Equivalent Income', in C. Riskin, R. Zhao, and S. Li, *China's Retreat from Equality: Income Distribution and Economic Transition*, M.E. Sharpe: Armonk NY.
- Hannum, E., J. Behrman, and M. Wang (2005). 'Human Capital in China', in L. Brandt and T.G. Rawski (eds) *China's Great Transition: Origins, Mechanisms and Consequences of the Post-Reform Economic Boom*, Department of Economics, University of Pittsburgh: Pittsburgh.
- Jalan, J., and M. Ravallion (1998). 'Transient Poverty in Postreform Rural China', *Journal of Comparative Economics* 26(2): 338-57.
- Jones, F.L. (1983). 'On Decomposing the Wage Gap: A Critical Comment on Blinder's Method', *Journal of Human Resources* 18(1): 126-30.
- Kanbur, R., and X. Zhang (1999). 'Which Regional Inequality? The Evolution of Rural-Urban and Inland-Coastal Inequality from 1983 to 1995', *Journal of Comparative Economics* 27(4): 686-701.
- Khan, A.R., K. Griffin, C. Riskin, and R. Zhao (1993). 'Household Income and Its Distribution in China', in K. Griffin and R. Zhao (eds) *The Distribution of Income in China*, Macmillan: Basingstoke.
- Khan, A.R., and C. Riskin (2007), 'Growth and Distribution of Household Income in China Between 1995 and 2002', in B. Gustafsson, S. Li and T. Sicular (eds) *Inequality and Public Policy in China*, Cambridge University Press: New York.
- Knight, J., and L. Song (1999). *The Urban-Rural Divide: Economic Disparities and Interactions in China*, Oxford University Press: New York.

- Lam, K.C. (2003). 'Earnings Advantage of Party Members in Urban China', mimeo, Business Research Centre, Hong Kong Baptist University, Hong Kong.
- Li, S., C. Luo, Z. Wei, and X. Yue (2007). 'Appendix: The 1995 and 2002 Household Surveys: Sampling Methods and Data Description', in B. Gustafsson, S. Li and T. Sicular (eds) *Inequality and Public Policy in China*, Cambridge University Press: New York.
- Liang, Z., and Z. Ma (2004). 'China's Floating Population: New Evidence from the 2000 Census', *Population and Development Review* 20(3): 467-88.
- Miles, D. (1997). 'A Household Level Study of the Determinants of Incomes and Consumption', *The Economic Journal* 107(440): 1-25.
- Mo, R. (2004). *2003-2004 Nian: Zhongguo Jiuye Baogao*, Ministry of Labour and Public Security Press: Beijing.
- Morduch, J., and T. Sicular (2000). 'Politics, Growth and Inequality in Rural China: Does it Pay to Join the Party?', *Journal of Public Economics* 77(3): 331-56.
- NBS (National Bureau of Statistics, China) (1996). *China Statistical Yearbook 1996*, China Statistical Press: Beijing.
- NBS (National Bureau of Statistics, China) (2003). *China Statistical Yearbook 2003*, China Statistical Press: Beijing.
- Oaxaca, R.L., and M.R. Ransom (1999). 'Identification in Detailed Wage Decompositions', *Review of Economics and Statistics* 8(1): 154-7.
- Ravallion, M. and S. Chen (2004). 'China's (Uneven) Progress Against Poverty', *World Bank Policy Research Department Working Paper* 3408, World Bank: Washington DC.
- Sahn, D., and D. Stifel (2003). 'Urban-Rural Inequality in Living Standards in Africa', *Journal of African Economies* 12(4): 564-97.
- Shi, X. (2004). 'Urban-Rural Income Differentials Decomposition in China 1990s', master's thesis, China Center for Economic Research, Peking University, Beijing.
- Shorrocks, A., (1984). 'Inequality Decomposition by Population Subgroups', *Econometrica* 52(6): 1369-85.
- Shorrocks, A., and G. Wan (2005). 'Spatial Decomposition of Inequality', *Journal of Economic Geography* 5(1): 59-81.
- Sicular, T., Y. Zhao, and X. Shi (2004). 'Urban-Rural Income Inequality in China in the 1990s', mimeo, China Center for Economic Research, Peking University, Beijing.
- Smeeding, T.M., and D.H. Weinberg (2001). 'Toward a Uniform Definition of Household Income', *Review of Income and Wealth* 47(1): 1-24.

- Solinger, D. (1999). *Contesting Citizenship in Urban China: Peasant Migrants, the State, and the Logic of the Market*, University of California Press: Berkeley.
- Wang, F. (2004). 'Reformed Migration Control and New Targeted People: China's Hukou System in the 2000s', *China Quarterly* 177: 115-32.
- World Bank (2003). 'China: Promoting Growth with Equity', *Country Economic Memorandum Report* 24169-CHA, World Bank: Washington DC.
- WIDER (World Institute for Development Economics Research) (2005). World Income Inequality Database (WIID), v 2.0a: User Guide and Data Sources, June, www.wider.unu.edu/wiid/wiid.htm
- Wu, X., and J. Perloff (2005). 'China's Income Distribution, 1985-2001', *Review of Economics and Statistics* 87(4): 763-75.
- Yang, D.T., and H. Zhou (1999). 'Rural-Urban Disparity and Sectoral Labour Allocation in China', *Journal of Development Studies* 35(3): 105-33.
- Yue, X., T. Sicular, S. Li, and B. Gustafsson (2007). 'Explaining Incomes and Inequality in China', in B. Gustafsson, S. Li and T. Sicular (eds) *Inequality and Public Policy in China*, Cambridge University Press: New York.
- Zhang, P. (1997). 'Income Distribution during the Transition in China', *WIDER Working Paper* 138, UNU-WIDER: Helsinki.
- Zhang, X., and R. Kanbur (2005). 'Spatial Inequality in Education and Health Care in China', *China Economic Review* 16(2): 189-204.
- Zhang, X., and S. Fan (2004). 'Public Investment and Regional Inequality in Rural China', *Agricultural Economics* 30(2): 89-100.
- Zhao, R., and S. Ding (2007). 'The Distribution of Wealth in China', in B. Gustafsson, S. Li and T. Sicular (eds) *Inequality and Public Policy in China*, Cambridge University Press: New York.
- Zhao, X.B., and S.P. Tong (2000). 'Unequal Economic Development in China: Spatial Disparities and Regional Policy Reconsideration, 1985-1995', *Regional Studies* 34(6): 549-61.
- Zhou, Y., and L.J.C. Ma (2003). 'China's Urbanization Levels: Reconstructing a Baseline from the Fifth Population Census', *China Quarterly* 173: 176-96.