DISCUSSION PAPER SERIES

IZA DP No. 2003

Income Mobility of Individuals in China and the United States

Niny Khor
John Pencavel

March 2006

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Income Mobility of Individuals in China and the United States

Niny Khor
Stanford University

John Pencavel
Stanford University
and IZA Bonn

Discussion Paper No. 2003
March 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 2003
March 2006

ABSTRACT

Income Mobility of Individuals in China and the United States

Though much has been written about annual income inequality in China, little research has been conducted on longer run measures of income inequality and on income mobility. This paper compares income mobility of urban individuals in China and the United States in the 1990s. The following questions are taken up. To what extent are measures of annual income inequality misleading indicators of long-run income inequality? How much income mobility was there in China in the first half of the 1990s and how did this compare with mobility in other countries? Have real income increases been greater for the poor or the rich? How important is the variation in permanent incomes in China and how has this changed?

JEL Classification: D31, D63, O15

Keywords: income inequality, income mobility, China, United States

Corresponding author:

John Pencavel
Department of Economics
Stanford University
Stanford, CA 94305-6072
USA
Email: pencavel@stanford.edu

INCOME MOBILITY OF INDIVIDUALS IN CHINA AND THE UNITED STATES

Niny Khor and John Pencavel*

I. Introduction

China's labor markets have moved from those identified with a highly centralized, state-directed, and planned economy towards a more decentralized market economy and, unlike the transformations in Eastern Europe and the former Soviet Union, this conversion has taken place without a drop in output. Employment by private enterprises has been growing and state enterprises have greater autonomy over whom to hire and what to pay. Nevertheless, in many respects, Chinese labor markets remain quite different from their Western counterparts. For instance, Western-style trade unions independent of the state do not exist.[1]

The movement away from a state-regulated economy in China appears to have been joined with an increase in annual income inequality. However, the meaning of this increase in inequality remains undetermined. If the rise in inequality in annual incomes is accompanied with more income mobility from year to year, income inequality measured over a longer interval of time may not have increased at all. The purpose of this paper is to encourage a focus on longer-term measures of income inequality, to initiate the analysis of income mobility in urban China, and to provide some comparisons with income mobility in other countries especially the United States.[2] Hence, our analysis is essentially comparative

---

* This paper is drawn from a longer manuscript on income inequality and income mobility in China presented on 1 October 2005 at the conference on China's Policy Reforms: Progress and Challenges at the Stanford University Center for International Development (SCID). We acknowledge the helpful comments on this previous paper of Gary Fields, Robert Flanagan, Denise Hare, David Metcalf, Yingyi Qian, and conference participants and we thank SCID for research support and encouragement. The comments of two anonymous referees further improved this draft. This research on this paper was supported by a grant from the Smith Richardson Foundation through the Stanford Institute for Economic Policy Research. [mobilityofincomesinChina.ppr]

[1] See Metcalf and Li (2005) on the state of labor unionism in China.

[2] There are many other investigations of the income structure in China, some using the very same surveys that we exploit here such as Gustafsson and Shi (2001), Khan and Riskin (1998, 2001), Knight

- relating China's income distribution and income mobility in the 1990s to that in the United States. We impose the same restrictions on Chinese and American income data to effect comparisons of income mobility between the two countries. It is difficult to draw conclusions about the meaning of any measured income inequality in an economy unless the material is placed in some comparative context such as that afforded by contrasting different economies at approximately the same time.

A fundamental (often latent) issue motivating much of this work is the enduring question of the extent to which alternative economic systems are directly associated with different degrees of income inequality. Some have suggested that market economies tend to generate more turbulence and year-to-year change in economic status so that measures of annual income can be a misleading indicator of inequality.[3] For this reason, our primary interest in this paper is in describing the amount of change in economic status in China and in contrasting this with the United States and some other countries. To what extent are measures of annual income inequality in China misleading indicators of long-run income

---

and Song (2003, 2005), and Meng (2004). However, we have not come across an analysis of income mobility in urban China of the sort undertaken here. There is Zhou's (2000) survey in 1993-94 of almost five thousand people asking for information on their incomes for eleven sundry years from 1955 to the 1990s, but he does not undertake the type of empirical work here and one wonders how accurate people's recollections can be of their incomes some forty years earlier.

[3] The classic statement is Milton Friedman's (1962): "Consider two societies that have the same distribution of annual income. In one there is great mobility and change so that the position of particular families in the income hierarchy varies widely from year to year. In the other, there is great rigidity so that each family stays in the same position year after year. Clearly, in any meaningful sense, the second would be the more unequal society. The one kind of inequality is a sign of dynamic change, social mobility, equality of opportunity; the other, of a status society.....[C]ompetitive free enterprise capitalism tends to substitute the one for the other. Non-capitalist societies tend to have wider inequality than capitalist, even as measured by annual income; in addition, inequality in them tends to be permanent, whereas capitalism undermines status and introduces social mobility", pp.171-2. The Marxist position is that capitalism generates growing income inequalities with a small and shrinking elite capturing a larger fraction of a society's resources. The facts on income distribution in the planned economies of the Soviet Union and East Europe and on the changes during the transitions are summarized in Flemming and Micklewright (2000).

inequality? How much income mobility was there in urban China in the first half of the 1990s and how did this compare with mobility in the U.S.? Have real income increases been greater for the poor or the rich? How important is the variation in permanent incomes in China compared with the U.S. and how has this changed? These questions form the focus of this paper.

II. Information on Incomes of the Same People in Successive Years

Data Sources

The source of data on China in this paper is the Chinese Household Income Project (Riskin, Renwei, and Shi (2000)), a survey conducted in 1996 of about 8,000 rural households (representing some 35,000 individuals) and almost 7,000 urban households (approximately 22,000 members).[4] This paper covers urban individuals only. The data are derived from larger samples designed by China's State Statistics Bureau (SSB), but the questions about income are different from the SSB's surveys. Nonresponse is rare although excluded from the urban sample are those without a formal certificate of residence (*hukou*), an increasingly serious omission over time as the size of this population grows. Individuals are asked to keep a record of their incomes and expenditures and they are asked to consult their records before providing information on incomes in previous years.

Details about the construction of our samples are provided in the Appendix. In all cases, to reduce the influence of errors of measurement that are most likely to affect the values of outlying observations, we trim the data by omitting the 0.5 percent of the lowest and the 0.5 percent of the largest values of income in any sample. This has the effect, of course, of reducing the values of indicators of income inequality (such as the Gini coefficient) that draw on information at all points in

[4] The Chinese Household Income Project is a joint research effort sponsored by the Institute of Economics, Chinese Academy of Social Sciences, the Asian Development Bank and the Ford Foundation. Additional support was provided by the East Asian Institute, Columbia University. An analysis of some of its findings is contained in Khan and Riskin (2001).

the income distribution. In some cases, we examined the impact of this trimming procedure and found it inconsequential for our inferences about inequality. We also "cleaned" the retrospective data when errors in the values of variables seemed manifest.

We are concerned not merely with income inequality in a single year, but how inferences about income inequality are affected by measuring incomes over a longer period of time. If a society resembles a caste system in which the income position that people occupy in one year is the same as that in another year, then a longer time perspective will have no affect on inferences about income inequality. On the other hand, if the society is characterized by a great deal of income mobility, measures of income inequality derived from comparisons of income in a single year may be very misleading as indicators of longer run income inequality.

Addressing these issues requires the use of data that follow the same people over time or that collect retrospective information.[5] Such data are available in the Chinese Household Income Project survey for 1996 which asks individuals in urban areas to report their "total income" not only for 1995, but for each of the previous five years so that the incomes of individuals are available each year from 1990 to 1995.[6] Our data set consists of individuals aged between 17 and 63 years in the initial year 1990. It turns out that, after trimming the data by omitting the top 0.5 percent and the bottom 0.5 percent in each year, all these individuals had a positive income.

The comparable information on individual incomes in the United States comes from the Panel Study of Income Dynamics (PSID), the primary longitudinal survey in the United States. The

[5] The significance of panel data to many pressing questions in poorer countries is stressed by Baulch and Hoddinott (2000) and Fields (2001, Ch. 6 & 7).

[6] Income consists of the sum of labor income, property income, transfer income and "income from household sideline production". This is before tax income. By far the largest fraction of total income is labor income.

procedures for the coding of wage income for the PSID were revised in 1993 and this impedes the tracking of the incomes of individuals from 1990 to 1995, the same years as those for individuals in China. Therefore, we selected an overlapping period where the definitions were unchanged, namely, the surveys from 1994 to 1999 that relate to the years 1993 to 1998. However, in 1997, the PSID moved from interviewing each year to biennial data collection and this implies we lack observations on income for the year 1997. Our sample of 1,847 individuals from the PSID excludes the over-sampled low income subset. We follow heads of households and the spouses of heads of households. We apply to the PSID the same definitions and criteria as those that were applied to the Chinese data: to be in the sample people must live in metropolitan areas of the country and be aged between 17 and 63 years in 1993. More details on the characteristics of the PSID sample are provided in the Appendix.

The Representativeness of the Chinese Sample

The retrospective information on incomes in China raises two obvious difficulties. The first is that respondents will recall their past incomes with error so that the correlation between incomes across years will embody errors in measurement.[7] The nature of these errors may take many different forms. Perhaps most likely is that, regardless of their true incomes, individuals will report the same incomes in different years or the same proportion of incomes in different years. If so, this will give the impression of more inertia in the income distribution than is really the case and our measures of income mobility will constitute the lower limit of true income mobility.

A second problem is that of non-response: not all individuals supply information on income in each and every year from 1990 to 1995. Among those aged 22 to 69 years in 1995, 88 percent provide usable income information for all five years. The survey does not provide explanations for the missing

[7] If there are imperfections in recalling incomes just five years ago, what must the errors be in a survey conducted in 1993-94 that asks individuals to recall their incomes in 1955, 1960, 1965, 1975, 1978, and more recent years (Zhou, 2000)?

data, but one may surmise that there are a number of possibilities: even though individuals are encouraged to check their records, some may not be able to recall their incomes in previous years and so, perhaps, choose not to respond; some (especially the young and old) may be out of the labor force for an entire year and may prefer not responding instead of reporting zero labor income; others may simply not make the effort to recall their previous incomes.

Non-responses create the same sort of problem as that of attrition in panel data: if non-response is not random, the sub-sample of people who provide information on incomes in all years is not representative of the entire urban adult population.[8] Because we have information on people in 1995, we may ask whether there is a systematic relationship between the probability of non-response and certain attributes of individuals. That is, define a variable, Z , that takes the value of unity for an individual who provides information on his income in every year from 1990 to 1995 and of zero otherwise. Relate this variable Z to a number of variables secured from the survey in 1996 including age, gender, marital status, Communist Party membership, ethnic minority, schooling, and a set of 1995 income dummy variables, one for each income decile. The estimated effects of changes in the right-hand side variables on the probability of full response are given by the coefficients in Table 1.

The probability of providing income information in all six years is a positive concave function of age and it is higher for married people and for Communist Party members. The probability of full response reaches a maximum at 46 years of age. Other things equal, a married individual is 14 percent more likely than a unmarried individual to provide information on incomes in all years. The probability of response also decreases with schooling: someone with a college education has a 10 percent lower

[8] A panel data study of rural China (Benjamin, Brandt, and Giles, 2005) lost almost one-third of households to attrition. Gibson *et al.* (2003) show that inequality estimates are substantially affected by the pattern of monthly non-responses to the Household Income and Expenditure Survey in urban China.

probability of providing income information in all years than someone with an elementary schooling. In part, this schooling relationship is the consequence of defining the set of people to be aged 22 to 69 years in 1995: some of these people will be in full-time schooling and may not supply information on their incomes if they earned nothing. Moreover, these highly schooled people may enter the labor force with relatively high incomes and their omission has the effect of understating income inequality.

Finally, the relationship between full income response between 1990 and 1995 and income in 1995 is non-linear.[9] The probability of providing full information rises mildly with income and reaches a maximum for those in the fourth income decile: those with income between the 30th and 40th income percentile have a four percent higher probability of reporting income than those with an income below the 10th percentile in 1995. Those with high incomes (those with incomes above the 60th percentile) have a response probability that is not significantly different from those with incomes below the 10th percentile. With people in the tails of the income distribution less likely to be providing income in all years, the measured income distribution will depart from the "true" income distribution: the measured income distribution describing those who report incomes in all years from 1990 to 1995 will reveal less inequality than the income distribution of the larger population of people reporting income in 1995.

This is confirmed by the information in Table 2 that compares the 1995 distribution of income among those who provided income in all six years with the 1995 distribution of income among the larger sample of those who provided income information in 1995. The 11,571 adults (aged between 22 and 69 years) who provide information on their income in 1995 reveal a slightly wider distribution of income in 1995 than the subset of 10,184 adults who provide full information on incomes in all six

---

[9] In the United States, attrition in the Panel Study of Income Dynamics has also been a nonlinear function of income with low income and high income people more likely to leave the panel than middle income people. See Becketti *et al.* (1988) and Fitzgerald *et al.* (1998). Attrition from the National Longitudinal Survey of Youth has been greater for those who have left school and are without jobs and, among those with jobs, greater for those with higher earnings. See MaCurdy *et al.* (1998).

years from 1990 to 1995.

The differences in the degree of inequality in the income distributions in Table 2 may not be viewed as large, but (together with the other relationships reported in Table 1) they do imply that our analysis below of income mobility among the people who provide income information in all six years will not be fully representative of the total set of adults. This is further supported by determining the degree of annual income inequality in each year of the 10,184 people who supply income information in all six years. These indicators of income inequality are given in Table 3 and they indicate that, for these 10,184 people, income inequality fell mildly between 1990 and 1995. This result is counter to the change in income inequality among the larger sample of urban individuals between 1988 and 1995 contained in the Chinese Household Income Project surveys for 1989 and 1996.[10]

The full responders are a more stable group of people (more likely to be married, to be middle-aged, and to be Communist Party members) who are drawn disproportionately from the middle part of the 1995 income distribution. Such differences between all individuals and those who provide data in all years are common to research on income mobility where attrition or non-response or administrative regulations result in analyses on samples not fully representative of the population. Nevertheless, it means that our statements below about income mobility need to be tempered by the recognition that our analytical sample is not entirely representative of the urban population in China.

Our analysis proceeds using the 10,184 urban Chinese individuals who provided information on their incomes in all years from 1990 to 1995. All values of income in different years are expressed in 1995 yuan by using the consumer price index as a deflator.[11] Similarly, the incomes of the 1,847

---

[10] See Khan and Riskin (1998) and Khor and Pencavel (2005).

[11] We do not believe there are no systematic regional price differences and we are aware that some authors have constructed regional price indices (e.g., Démurger, Fournier, and She, 2005). However, we do not know how much confidence to place in these regional price deflators and so we decline to

urban Americans we follow from the PSID from 1993 to 1998 are expressed in 1996 dollars by using the personal consumption expenditures price deflator.

III. Income Mobility

Income Quintiles and Income Clusters

Consider a familiar description of income mobility, one that uses information on the probability of an individual moving across income quintiles or income clusters. For this analysis, for urban China, we use the 10,184 individuals aged 22 to 69 years in 1995 reporting incomes in every year between 1990 and 1995. Cross-classify these people into income quintiles from I (the bottom quintile) to V (the top quintile) in 1990 and in 1995 with an equal number of people in each quintile.[12] Then five-by-five income transition tables are constructed where each element consists of p_{jk}, the fraction of individuals in quintile j in 1990 who occupy quintile k in 1995. The transition matrix for all 10,184 people is presented in Table 4 with those for men and women separately below it.

According to the top panel of Table 4, one-half (precisely, 0.496) of those who occupied the richest quintile of the population in 1990 remained in the same quintile six years later in 1995. Some 44 percent (0.439) of the poorest fifth in 1990 remained in the same quintile in 1995. Though this might suggest a good deal of immobility or persistence in relative incomes, the identical information may be presented with the opposite perspective: one-half of those in the highest income quintile in 1990 had fallen out of this quintile by 1995 and more than one-half (0.561) of those in the bottom quintile of incomes in 1990 had risen out of this quintile by 1995! Intriguingly, outside the top and bottom quintiles, the entries in the main diagonal tend not to be the largest in any row: for the second and third

---

apply them.

[12] In the event of any ties, people were allocated randomly to the adjacent quintiles so that the same number of observations were allocated to each quintile.

quintiles, a larger fraction of people fell one quintile than remained in the same quintile; in the fourth quintile, an equal fraction fell one quintile as stayed in that quintile. As a consequence - sadly, with an ordinal scale, all people cannot improve themselves - there is a tendency for a greater incidence of large increases in relative incomes than of large declines in relative incomes.[13] The transition matrices of men and women in the bottom two panels of Table 4 are remarkably similar.

A corresponding income transition matrix for 1,847 individuals in the United States is given in Table 5. Seventy percent of those who occupied the highest income quintile (quintile V) in 1993 were in the same quintile in 1998 while almost 60 percent (precisely, 59.1 percent) occupying the lowest quintile in 1993 remained in that same quintile in 1998. These entries are noticeably larger than the corresponding values for urban China. Though there is some indication of greater income mobility for women than men in the United States, the gender differences are second order.[14]

To help in comparing income mobility in China with that in other countries, consider three summary indicators of the income mobility embodied in the transition matrices: first, the average quintile move; second, the fraction who remain in the same quintile, also called the “immobility ratio”;

---

[13] Thus 9.9 percent of those with incomes in the second quintile in 1990 rose to the top quintile in 1995 whereas only 7.6 percent of those with incomes in the fourth quintile in 1990 fell to the bottom quintile in 1995. Or 4.9 percent of those in the bottom quintile in 1990 rose to the lowest quintile in 1995 while just 2.1 percent of those in the top quintile in 1990 fell to the bottom quintile in 1995. Nevertheless, we cannot reject the hypothesis that the transition matrices are symmetric. A maximum likelihood test of symmetry makes use of the statistic $\Omega = \Sigma_{i>j} (p_{ij} - p_{ji})^2 / (p_{ij} + p_{ji})$ which has a chi-square distribution with $q.(q - 1) / 2$ degrees of freedom (where q is the number of quantiles, here five). The hypothesis of symmetry for the transition matrices in Table 4 cannot be rejected with a high confidence. See Bishop, Fienberg, and Holland (1975, pp. 282-3).

[14] The immobility ratio (the fraction remaining in the same quintile) for U.S. men is 0.519 and that for women is 0.492. The average quintile move is 0.636 for men and 0.697 for women.

and, third, the fraction who remain in the same quintile plus the fraction who move one quintile.[15] By way of reference, if every entry to the transition matrix (that is, if every value for p_{jk}) were one-fifth (sometimes defined as a transition matrix corresponding to "perfect mobility"), the average quintile move would be 1.6, the immobility ratio would be 0.20, and the fraction who remain in the same quintile plus the fraction who move one quintile would be 0.52. Transition matrices and these indicators of income mobility are constructed for all urban people in China; for men and women separately; for those aged less than 30 years, those aged 30-50 years, and those aged more than 50 years; and for three different schooling groups. In addition, because the regional disparities within China are well known, we computed these indicators of income mobility for people in the "coastal" areas separately from people in the "interior".[16]

The columns under (1) in Table 6 list the three indicators of mobility for all the adults and for various sub-groups of people in China. For all (the first row of Table 6), the average quintile move is 1.06, the immobility ratio is one-third, and the fraction who remain in the same quartile plus the fraction

[15] The average quintile move is defined as

$$\frac{1}{5}\left\{\sum_{j=1}^{5}\sum_{k=1}^{5}(|j-k|)\,p_{jk}\right\}.$$

The fraction who remain in the same quintile is defined as $(5)^{-1}\Sigma_{j\,=\,1,..,5}\,(p_{j\,j})$.

[16] The "coastal" areas consist of the provinces of Liaoning, Jilin, Jiangsu, Zhejiang, Shandong, Guangdong, Fujian, and Beijing. Of course, Beijing does not have a sea coast which is why we place the word "coastal" in inverted commas. To avoid problems posed by geographic mobility, this coastal/interior analysis is undertaken for all those 8,973 individuals who reply to the survey in 1996 to have been employed in the same work unit for at least five years. This means that, from the full sample of 10,184 people, we drop 670 people who are living in the coastal provinces in 1995 and 541 people living in the interior. Insofar as changing jobs or locations is the mechanism by which income mobility is engendered, our focus on non-movers will overlook this mobility.

who move one quintile is 0.711.[17] There are relatively small differences in these indicators of mobility for the various sub-groups of the population: focusing on people aged 30-50 years (that is, deleting both young people at or close to their school years and older people near or at retirement) has little effect on mobility and different schooling groups have similar income mobility. Those living in the interior provinces reveal no less mobility (and perhaps a little more) than those in coastal areas though, for this particular analysis, we delete individuals who had not worked at the same unit for less than five years. The interaction among employment, geographic, and income mobility requires a more sophisticated analysis than we have effected here.

How do these numbers compare with those in other countries? Data definitions and procedures are not the same across countries and small differences in definitions may have large consequences for measures of income mobility. Nevertheless, to provide some sort of reference for these indicators of income mobility for China in Table 6, examine the entries to Table 7 which provide mobility indicators for earnings for certain other countries in addition to China. The notes beneath Table 7 identify the key data differences among the countries: for instance, the mobility data in the line "USA - 1" are constructed from the income transition matrix in Table 5 while the line "USA - 2" relates to weekly earnings between 1986 and 1991. Heavy reliance should not be placed on small differences in the values in Table 7 across countries.[18] However, the differences between China and the other countries

[17] Suppose we use not the 10,184 individuals reporting incomes in all six years but the 10,692 individuals providing income information in at least the two years 1990 and 1995. Then the average quintile move is 1.062, the immobility ratio is 0.332, and the "stayers + one movers" ratio is 0.709. In other words, the basic results of Table 6 are unaffected.

[18] A number of other studies provide comparisons across countries in income mobility. For instance, Burkhauser *et al.* (1998) compare income mobility in the United States with that in Germany in the 1980s. Generally, the degree of income mobility in the two countries is very similar: "The mobility patterns of U.S. and German men are virtually the same. Despite very different labor market institutions, and very different tax and transfer systems, the dynamic outcomes of working age men and women are remarkably similar" (p. 144). Comparing the United States, Germany, and Britain, Fabig

appear large, not small. For all four indicators of income mobility across quintiles in Table 7, China's mobility is substantially greater than that in other countries: the immobility ratio in urban China is about two-thirds of that in the other countries and the correlation coefficient is about three-quarters of the average of the other countries.

An alternative method to measure transitions across income quintiles is to measure incomes after eliminating systematic income differences associated with various exogenous factors that may be associated with mobility. To that effect, consider the degree of income mobility in China after controlling for differences among these individuals in their age, schooling, and gender. The indicators of mobility after holding constant these factors are shown in columns under (2) of Table 6.[19] These indicators suggest even greater mobility: the average quintile move rises from 1.06 to 1.14, the immobility ratio falls from 0.334 to 0.307, and the fraction who remain in the same quartile plus the fraction who move one quintile falls from 0.71 to 0.68. For the United States between 1993 and 1998, the corresponding indicators are 0.767 for the average quintile move, 0.464 for the immobility ratio, and 0.835 for the fraction who remain in the same quartile plus the fraction who move one quintile. The clear differences of income mobility between China and the United States remain.

The rate of transitions across income quintiles will not be independent of the degree of income inequality in a society. Thus an individual who experiences a given absolute increase in income will be more likely to move across quintiles in an economy with a narrow income distribution than an individual enjoying the same income increase in a society with a wide income distribution. The economies listed

---

(2000) draws somewhat different conclusions. Aaberge *et al.* (2002) compare income mobility in the United States with that in the Scandinavian countries in the 1980s and they claim that, while annual income inequality in the United States is greater, income mobility is similar.

[19] To be precise, in each calendar year, cross-section income equations were fitted that regressed the logarithm of income on age, age squared, years of schooling, and gender. The residuals from these equations form the basis of the indicators in the columns under (2) in Table 6.

in Table 7 exhibit notable differences in annual income inequality and this will affect the measured indicators of income mobility. To address this, suppose a transition matrix is defined not on the basis of income quintiles but on the basis of deviations from median income. Define five income clusters as follows: the lowest cluster consists of people with less than 0.65 of the median income; the second cluster consists of people with incomes between 0.65 and 0.95 of the median income; the third income cluster consists of people with incomes between 0.95 and 1.25 of the median income; the fourth cluster consists of people with incomes between 1.25 and 1.55 of the median income; and the fifth cluster consists of people with incomes above 1.55 of the median income. Obviously, if the median is the same in the two countries, the income cutoffs will be the same, but they will correspond to different fractions of individuals when income dispersion is different in the two countries. In an economy with a wide income distribution, more people will be in the income cluster of less than 0.65 of the median compared with an economy with a narrow distribution. However, now a given absolute income increase in two societies will be equally likely to cross the boundaries between income clusters.

The consequences for our indicators of income mobility in China of measuring transitions across income clusters rather than transitions across income quintiles are shown in the columns beneath (3) "income clusters" in Table 6. Measured income mobility is slightly less than that in previous columns - something to be expected given the relatively narrow income distribution in urban China in the first half of the 1990s - but the change is small. Table 8 compares these income-cluster indicators of income mobility for China with those for other countries. The degree of divergence between China and individual countries changes but the general pattern remains the same: the immobility ratio in China is now about 64 percent of the average of the other countries'. The general inference is that the degree of income mobility in urban China during the first half of the 1990s was unusually high, unusual, that is, by the standards of income mobility in rich economies.

The Effect of a Longer Time Perspective

An issue related to the degree of mobility across income quintiles or income clusters is the extent to which measures of income inequality are affected by computing incomes over a longer time horizon. The effect on indicators of income inequality of averaging incomes over a longer time horizon in China is shown in the upper panel of Table 9. The Gini coefficient for incomes in 1995 is 0.274 and it declines to 0.250 when incomes are averaged over the four years from 1992 to 1995. The ratio of income at the 90th to the 10th percentile falls from 3.55 in 1995 to 3.17 when incomes are averaged over the four years from 1992 to 1995. For the United States (the bottom panel of Table 9) income inequality also falls by comparable amounts when incomes are averaged over longer time horizons[20] and "permanent" inequality - the point at which Gini coefficients (or other indicators) plateau - is reached after about the same number of years in the two countries. The degree of income inequality when averaged over more than one year depends on the level of inequality in each year, on how this inequality changes over time, and on the correlation coefficients of income in different years.[21]

Correlation coefficients of income across the years in China are given in Table 10 which suggests not only that correlation coefficients fall as the distance between years increases but also that,

[20] Thus, the ratio of the Gini coefficient when incomes are averaged over four years to the Gini coefficient when income is computed for one year (1995 for China and 1998 for the United States) is 0.91 for urban China and 0.94 for the U.S. Again note that the incomes for the U.S. exclude the year 1997 when the PSID did not survey income.

[21] Consider the simple case of measuring income inequality over three years. Suppose income inequality is changing over time. In particular, if σ_j denotes the standard deviation of income in year j, suppose $\sigma_2 = \lambda_2.\sigma_1$ and $\sigma_3 = \lambda_3.\sigma_1$ where λ_3 and λ_2 are greater or less than unity. Then the coefficient of variation of "lifetime" (i.e., three year) income is equal to $\kappa.\sigma_1 / y'$ where y' denotes mean income over the three years and $\kappa = [1 + \lambda_2^2 + \lambda_3^2 + 2(\lambda_2.r_{12} + \lambda_3.r_{13} + \lambda_2\lambda_3 r_{23})]^{1/2}$ where r_{ij} is the correlation coefficient between years i and j. So the inequality of lifetime income depends positively on the year-to-year change in inequality (the λ's) and on the across-year coefficients of correlation of income. As noted in Table 7, the correlation coefficients of incomes across years tend to be lower in China than in other countries.

holding constant the gap between years, correlation coefficients are lower for the years in the mid-1990s than those in the early 1990s. In other words, correlation coefficients of incomes of the same individuals in different calendar years fall over time (being lower for 1994 and 1995 than for 1990 and 1991). Analogous correlation coefficients between incomes for the United States are reported in Table 11. Clearly the correlation coefficients for China tend to decline more rapidly than those for the United States as years intervening increase. This is illustrated in Figure 1 where the values of the correlation coefficients for all individuals in Tables 10 and 11 are graphed against the years intervening. The correlation coefficients for the United States (given by a circle in Figure 1) and those for China (given by a cross) both decline with years intervening, but the rate of decline is greater for China so that, after five years, the correlation coefficient for China is 0.541 while that for the United States is 0.749.

A striking feature of the correlation coefficients for China in Table 10 is that, holding constant the number of intervening years, correlation coefficients are lower for the years in the mid-1990s than those in the early 1990s. This suggests increasing income mobility in China in the 1990s. To determine whether this result survives a more thorough analysis, we use the data on the years from 1990 to 1995 to construct gender-age-calendar time cells and calculate the correlation coefficient between the average incomes of the people in these cells. That is, for men and women separately, we compute the correlation coefficient between the incomes of people at each age a and at age $a + j$ where $j = 1, 2, 3, 4, 5$.[22] Let the correlation coefficient between incomes for those aged a and $a + j$ observed in years

[22] Thus, when $j = 1$, there are five possible data points: those involving the years 1990 and 1991, 1991 and 1992, 1992 and 1993, 1993 and 1994, and 1994 and 1995. When $j = 2$, there are four data points: 1990 and 1992, 1991 and 1993, 1992 and 1994, and 1993 and 1995. When $j = 3$, there are three data points: 1990 and 1993, 1991 and 1994, and 1992 and 1995. When $j = 4$, there are two data points: one entails the pair of years 1990 and 1994 and the other involves the years 1991 and 1995. When $j = 5$, the only set of observations consist of those correlating incomes in 1990 and 1995. So there are fifteen correlation coefficients for each gender and each age. If age ranges from 17 years to 64 years (as measured in 1990), there are 48 ages and two genders so there are 1,440 (15 x 48 x 2) potential

t and $t + j$ for gender g be given by $\rho(g; a, a + j; t, t + j)$.[23] Then we specify a least-squares regression equation in which ρ is a function of age in 1990, gender, the number of years between a and $a + j$ (or, equivalently, between t and $t + j$), the average years of schooling of the people in each cell, and the calendar years when the correlation is computed:

$$(1) \quad \rho(g\ ; a, a + j\ ; t, t + j) = \gamma_0 + \gamma_1 (a) + \gamma_2 F + \gamma_3 (Schooling) + \gamma_4 (a - a') + \Sigma_t \delta_t . Y_t + u$$

where u is a stochastic error term, a measures years of age (as of 1990), F is a dummy variable taking the value of unity for a woman and of zero for a man, *Schooling* measures average years of schooling of individuals in the cell, and $a - a'$ represents the number of years between the two years (t and $t + j$) whose incomes are being correlated. Y_t is a dummy variable that takes the value of unity if the correlation coefficient involves year t and of zero otherwise. There are five calendar year dummy variables, one for each year from 1991 to 1995; the year 1990 is the reference year. This equation is estimated by weighted least-squares, using the number of individuals in each cell as weights. The resulting estimates of equation (1) are reported in Table 12.

These estimates suggest that the correlation coefficient between incomes is 0.01 higher for women than for men and it is slightly higher for those with more schooling: other things equal, someone with four years of schooling than another has a correlation coefficient that is almost 0.015 higher. Correlation coefficients fall gradually with age: other things equal, someone thirty years older than

---

correlation coefficients. In fact, some cells are empty so the effective data set of correlation coefficients is 1,410. Note that, in this exercise, we are not restricting the individual data to consist of those people who report incomes in each and every year though necessarily people must have incomes in the two years to which the correlation coefficients relate.

[23] So one observation for $\rho(g; a, a + j; t, t + j)$ is the correlation coefficient between the incomes of men in 1990 and 1992 aged 30 years in 1990. A second observation is the correlation coefficient between the incomes of men in 1993 and 1995 aged 30 years in 1990. These two observations of ρ both have a value of j of 2 but they are observed in two different pairs of years.

another will exhibit a 0.02 lower correlation coefficient.[28] Correlation coefficients decline sharply with the number of years intervening between the two years: other things equal, when the incomes are separated by four years the correlation coefficient is 0.25 lower than when the incomes are separated by merely one year. This is consistent both with income mobility being greater when measured over a longer interval of time and with errors in recalling incomes that increase with the time interval.[29]

The coefficient estimates attached to the calendar year dummy variables testify to a declining association over calendar time. To be specific, consider what the estimates imply about the correlation coefficients between incomes two years apart. Denote the predicted values of ρ for, say, men aged 40 years in 1990 with ten years of schooling in calendar years τ_1 and τ_2 as $\rho(\tau_1, \tau_2)$. Then ρ(1990, 1992) = 0.872, ρ(1991, 1993) = 0.841, ρ(1992, 1994) = 0.804, and ρ(1993, 1995) = 0.700. So, holding constant $a - a'$, correlation coefficients of incomes are falling over time suggesting an increasingly income mobile society. This is an important finding: in the first half of the 1990s, correlation coefficients between incomes across years in urban China fell over time.

IV. Changes in Real Income of People at Different Points of the Income Distribution

The analysis of income mobility across quintiles (or clusters) of the income distribution undertaken in the previous section is an analysis of movement across income thresholds. The degree to which these thresholds are exceeded or not exceeded is neglected. Moreover, the movement across

---

[28] The relationship with age is not non-linear: adding a quadratic term in age does not change our inferences about the weak negative association between age and the correlation coefficients. When a quadratic term in age is added, its estimates suggest a positive concave relationship, but the implied peak in age is younger than the youngest people in the sample.

[29] As noted earlier, an interpretation in terms of errors in reporting depends on the form that these errors take. If each individual reports not the true values of income but reports the very same income or the same proportion of income as years from 1996 increase, this would have the tendency of showing less income mobility than the true (unobserved) data. By contrast, we observe more income flux in China than in the United States for years far apart.

income quintiles (or clusters) describes relative movement and is consistent with the real income of all people falling or rising. Similarly, a given value of the coefficient of correlation of incomes across years is consistent with quite different values of the change in real income of the individuals.

We now turn to an analysis of the magnitude of real income increases at different points of the initial income distribution. China's urban economy grew rapidly in real terms in the first half of the 1990s and we ask whether those at the lower tail of the income distribution in 1990, the beginning of our data, experienced larger income gains by 1995 than those who started toward the upper tail of the income distribution. During this period, did the income process favor the relatively poor or the relatively rich in urban China?[30] The data we analyze are again the 10,184 people who provide information on their incomes in all six years from 1990 to 1995.

Order the 10,184 people by their incomes in 1990 and attach to each individual his or her percentile in the 1990 income distribution. Let individual *i*'s percentile in the 1990 income distribution be given by P^y_i. Each percentile of the 1990 income distribution is occupied by approximately 102 people (i.e, roughly 10,184 individuals divided by the one hundred percentile points). Let $ln[y_i(t)]$ be the logarithm of real income of individual i in calendar year t. Form for individual i the change in the logarithm of his real income between 1995 and 1990: $\Delta \ ln \ y_i = ln[y_i \ (1995) \ ] - ln[y_i \ (1990) \ ]$. The relation between the average value of $\Delta \ ln \ y_i$ at each percentile and P^y_i is given by the downward-sloping dotted line in Figure 2.[31]

The vertical axis of Figure 2 indicates that, for all percentiles in the 1990 income distribution,

---

[30] Fields *at al.* (2003, 2005) ask a similar question for other countries although their methods differ somewhat from ours.

[31] In Figure 2 the values of the change in income have been smoothed over five percentiles. There are 102 people in each percentile except for 86 people in the 100th percentile. In the case of ties, people are randomly assigned to adjacent percentiles which has the desirable effect of smoothing the data.

average real income increases are positive. The negative slope of the relationship means that those who experienced the larger increases in real income in the first half of the 1990s were those who started at the lower part of the income distribution in the year 1990. This is consistent with the hypothesis that China's economic growth disproportionately benefitted those with low incomes rather than those with high incomes though there are other interpretations of the data as we note shortly. In view of the fact that, over most of its range, the relationship graphed in Figure 2 does not depart egregiously from linearity, we may consider describing the relationship by means of a linear regression equation:

(2) $$\Delta \ln y_i = \gamma + \delta P^y_i + v_i$$

where γ and δ are parameters to be estimated and v_i is a stochastic term. The least-squares estimates of γ and δ are reported in column (1) of Table 13: the estimate of δ of -0.0086 implies that ten percentile points on the 1990 income distribution is associated with an 8.6 percent lower average income growth from 1990 to 1995.[32]

What is the appropriate inference from these relationships between initial income and the subsequent change in income? Is it that the income poor in urban China have benefitted more than the

---

[32] If γ is permitted to vary by demographic characteristics (age, schooling, and gender), the role of initial income remains relevant with the suggestion that those higher up the income order in 1990 experienced smaller income increases than those lower down. Of course, by introducing these demographic characteristics, we are moving away from the initial question: whether economic growth in the first half of the 1990s favored the relatively rich or the relatively poor. With demographic controls, we are not asking whether economic growth benefitted the poor but whether economic growth benefitted the poor after accounting for the links between income growth and these demographic characteristics. This is no longer the fundamental question to be addressed. Nevertheless, we mention these results to show that the magnitude of the link between $\Delta \ln y_i$ and P^y_i is robust with respect to the addition of these demographic characteristics. Other specifications lead to similar inferences. For instance, δ was also allowed to vary with these demographic characteristics. However, a joint test of the resulting interactions suggests that this amendment does not add significant explanatory power. Another specification relates the income change not to the 1990 income percentile but to the 1990 logarithm of income: $\Delta \ln y_i = \gamma' + \delta' [\ln y_i(1990)] + v_i'$. Again, the general inferences are similar to those reported in the text.

rich from economic growth? This depends on the meaning of our ordering of income in 1990. This rank of income in 1990 will be affected by transitory shocks that place some people above their permanent income and other people below their permanent income. Those in 1990 temporarily below their steady-state income will tend to experience a larger increase in their income from 1990 to 1995 while those in 1990 temporarily above their steady-state income will tend to experience a smaller increase in their income from 1990 to 1995. In this way, the estimates in column (1) of Table 13 do not provide a description of the relation between permanent (or long-run) income and subsequent income growth because they are subject to the familiar errors-in-variables problem.[33]

The relevance of these considerations is suggested by the relationship graphed by the dashed line in Figure 2 which relates income growth in China not to initial income but to final income. That is, plot the average change in the logarithm of income between 1990 and 1995 against the income percentile as measured in 1995. The dashed line in Figure 2 is decidedly positive: income changes in the early 1990s were least for those with low incomes by mid-decade. The corresponding least-squares estimates of equation (2) with P^y_i given by the income percentile in 1995 are reported in column (2) of Table 13. Now ten percentile points on the 1995 income distribution is associated with a 7.7 percent higher income growth from 1990 to 1995. The least-squares coefficient on the 1995 values of P^y_i is almost equal and opposite in sign to the coefficient on 1990 values of P^y_i. Again this is consistent with

[33] A way to depict this is to express the level of income in 1995 as a linear stochastic function of income in 1990: $ln[y_i(1995)] = a + b\, ln[y_i(1990)] + e_i$. Normally when autoregressive equations such as these are fitted, the least-squares estimates of b are positive but less than unity. If $ln[y_i(1990)]$ contains transitory components that affect income in 1990 but not income in later years, the ordinary least-squares estimates of b will be a downward biased measure of the association between the permanent components of income in different years. Upon subtracting $ln[y_i(1990)]$ from both sides of the previous equation, we have $ln[y_i(1995)] - ln[y_i(1990)] = a + b^*\, ln[y_i(1990)] + e_i$ where $b^* = b - 1$ and, if $0 < b < 1$, then b^* will be negative as we find. Though the estimates in Table 13 do not use $ln[y_i(1990)]$ on the right-hand side, the same logic applies to the income percentile, P^y_i.

a permanent-transitory income explanation: those in 1995 temporarily below their steady-state income will tend to have a smaller increase in their income from 1990 to 1995 while those in 1995 temporarily above their steady-state income will record a larger increase in their income from 1990 to 1995. The relationship does not trace out a link involving long-run incomes but a relationship strongly affected by transitory shocks to incomes.[34]

A better measure of permanent income is some central tendency of each individual's incomes over the entire six years.[35] Suppose then we form the average of each individual's annual income over the years from 1990 to 1995 and order this by percentile. With P^y_i now constructed using average income over the entire period, the relationship between average income growth and P^y_i is that graphed by the solid line in Figure 2. It does not change markedly from the fifth percentile to more than the fiftieth percentile after which point it declines.[36] Though the relationship graphed by the solid line in Figure 2 departs from linearity, upon estimating equation (2) with P^y_i computed using income averaged

---

[34] Expressed differently, suppose we write $ln[y_i(1990)] = c + d\ ln[y_i(1995)] + \zeta_i$ where c and d are parameters and ζ_i is a stochastic term. By the same logic as before, if $ln[y_i(1995)]$ contains transitory elements that raise or lower income in 1995 and in 1995 only, as an estimate of the association between permanent incomes, the least-squares estimate of d is likely to be biased downwards with a value of less than unity. Multiply the previous equation by minus unity and subtract $ln[y_i(1995)]$ from both sides: $ln[y_i(1995)] - ln[y_i(1990)] = c^* + d^*\ ln[y_i(1995)] + \zeta_i^*$ where $d^* = 1 - d$. If d is less than unity, d^* will be positive. In other words, when relating the growth of income to the income in the terminal year, an errors-in-variables interpretation suggests the relationship will be positive.

[35] An alternative procedure is to use a measure of initial year income that is predicted by a least-squares regression of initial year income on a number of covariates (such as schooling, age, and gender) as in Fields *et al.* (2003, 2005). The trouble with this procedure is that, at least in the context of China, these covariates remove a relatively small fraction of the variation in initial incomes. The consequence is that this procedure does not answer the basic issue of the extent to which economic growth benefits low income individuals. It replaces this issue with the question of the degree to which economic growth benefits those with a particular linear combination of variables that is loosely correlated with income.

[36] The value of $\Delta\ ln\ y_i$ at the fifth percentile is 0.91 and its value at the 57^{th} percentile is 0.92 after which it falls gradually to 0.73 at the 97^{th} percentile.

over the six years, the estimates are shown in column (3) of Table 13. Now the link between income percentile and income growth is much weaker though negative: the estimate of δ of -0.0016 implies that ten percentile points on the 1990-95 income distribution is associated with about a one-and-a-half percent lower income growth from 1990 to 1995. This small negative relationship between income growth and income level is consistent with the indicators of income inequality in Table 3: for these people, annual income inequality decreased slightly. So, for these people in China who reported their incomes in all six years, income inequality narrowed moderately and income growth in the early 90s mildly favored lower income individuals.

This interpretation of the relationships between income change and income level - that it is critical to discriminate between transitory and permanent income factors in computing this relationship - finds support when analyzing the 1,847 individuals from the PSID in the United States. For each individual, compute the logarithm of real income in 1988 minus the logarithm of real income in 1993: $\Delta \ln y_i$. First, rank these people by their incomes in 1993: P^y_i.[37] The resulting relationship between $\Delta \ln y_i$ and P^y_i in 1993 is graphed by the dotted line in Figure 3. This is negatively sloped (especially at lower percentiles) suggesting that those with lower incomes in 1993 tend to experience larger subsequent income increases. The least-squares estimates in column (4) of Table 13 suggest that, on average, ten lower percentiles on 1993 income is associated with a 2.4 percent greater income increase.

The dashed line in Figure 3 shows the relation between $\Delta \ln y_i$ and P^y_i when P^y_i is ranked by terminal year's incomes, 1998. This dashed line is positively sloped (again, principally at lower percentiles) and the fitted relationship reported in column (5) of Table 13 indicates that, on average, ten higher percentiles on 1998 income is associated with 4.3 percent greater income increase. The

[37] With 1,847 people, each percentile from the first to the ninety-ninth is occupied by 18 people with the remainder (65 people) occupying the 100th percentile.

negative relationship between $\Delta\ ln\ y_i$ and P^y_i when P^y_i is formed using initial incomes and the positive relationship between $\Delta\ ln\ y_i$ and P^y_i when P^y_i is constructed from final year's incomes demonstrate the principal role of transitory shocks in measuring these relations.

Finally, when P^y_i is based on the average of income over all years, the relation between $\Delta\ ln\ y_i$ and P^y_i is given in Figure 3 by the solid line which exhibits neither a dominant positive or dominant negative slope. According to the estimates in column (6) of Table 13, ten lower percentiles on average income is associated with less than a one percent greater income increase, a relationship that is barely significantly different from zero by conventional statistical criteria.[38]

These results on the association between the change in income and the level of income describe the central tendency. Behind each income percentile is a wide variety of different experiences. In other words, at each income percentile, for China, there are 102 individuals and Figure 2 and the regression equation (2) describes the average change in income among the 102 individuals at each income percentile. What about the variety of experiences at each percentile? One indicator of this is the difference between the maximum value of $\Delta\ ln\ y_i$ and the minimum value of $\Delta\ ln\ y_i$ at each income percentile. These values are graphed in Figure 4 against the income percentile where this income percentile is computed from the average of incomes over the six years. Evidently, there is a very wide range of experiences in China so that the typical smallest proportional change in income is a drop of almost ninety percent while the typical largest proportional increase in income is a gain of well over 600 percent. These gains and losses are similar across income percentiles so the extensive diversity of experiences is not at all strongly correlated with income. Expressed differently, although in these data

[38] Indeed, if γ is allowed to depend on demographic characteristics (age, schooling, and gender), the estimated coefficient on P^y_i (using the average of income over all years) is clearly not significantly different from zero on the usual tests. This weak relationship between P^y_i and $\Delta\ ln\ y_i$ is consistent with the very small changes in the Gini coefficient of individual incomes in the PSID over these years: the value of the Gini coefficient is 0.348 in 1993 and 0.340 in 1998.

for urban China there is a slight tendency for those with low permanent income to enjoy larger income gains in the first half of the 1990s than those with high permanent income, there was a substantial variation in experiences that is largely independent of income position.

V. The Relative Importance of the Variance in Permanent and Transitory Income

The meaning attached to any changes in annual income inequality depends on whether the permanent component or the transitory component of income inequality has changed. Given the pressure in earlier years in China to reduce income inequality, more recent reforms in China's labor markets aimed at linking earnings closer to skills are likely to cause an increase in the permanent variance in incomes and an increase in the permanent variance tends to result in a larger income gap between people when they are ordered by their incomes. In turn, this decreases income mobility across quintiles. By contrast, a move towards a market economy may well be associated with more turmoil and ferment in labor markets which may well result in more transitory income variance. Or freer labor markets may be characterized by less year-to-year arbitrariness and political favoritism in income determination with a corresponding decline in transitory income variance.

A larger transitory income variance generates more noise that will be manifested in greater year-to-year income mobility and a greater likelihood of change across income quintiles. If there are equal and opposite changes in the variance of permanent income and the variance of transitory income, there may be no change in the inequality of annual income and yet these developments may well portend important social developments. In other words, absence of change in annual income inequality provides no assurance that inequality in long-run permanent income is not changing. Conversely, changes in annual income inequality are compatible with no change in inequality in permanent income. The purpose of this section is to determine the relative magnitude of the transitory and the permanent component of income inequality and to ascertain whether any trends in these components can be identified for China in the first half of the 1990s and to compare them with those in the United States.[39]

The measurement of the permanent and transitory components of incomes rests crucially, of course, on the statistical process describing the evolution of incomes. The data at our disposal for China provide only six years of incomes and this severely restricts the generality of models that may be applied but we consider a simple process. Consider decomposing the logarithm of individual *i*'s real income in year *t* into permanent and transitory components

$$ln[y_i(t)] = ln[y^P_i(t)] + ln[y^T_i(t)] \;. \tag{3}$$

$ln[y^P_i(t)]$ denotes the logarithm of an individual's long-run or permanent income and $ln[y^T_i(t)]$ is the logarithm of an individual's transitory income.[40] Transitory income includes year-to-year unanticipated shocks to an individual's income and it incorporates errors in measuring income. These two elements - temporary components that are known to the individual and components that reflect the researcher's ignorance of true income - are conceptually quite distinct but here they are both embodied in $ln[y^T_i(t)]$. As usual, we assume that variations in permanent and transitory income are independent of one another. Transitory income is allowed to be serially correlated:

$$ln[y^T_i(t)] = \omega(t).ln[y^T_i(t\text{-}1)] + \varepsilon_i(t) \tag{4}$$

where ω is the first-order serial correlation parameter that is permitted to vary over time and $\varepsilon_i(t)$ is white noise. In this way, for example, an illness or a mishap that affects an individual's income in one year may well last into the following year. Or recall bias and measurement errors that affect incomes

---

[39] The points made in this paragraph are certainly not new. For a recent application to the distinction between chronic poverty and transient poverty in rural China, see Jalan and Ravallion (2000).

[40] This analysis was conducted also after removing the year-specific effects on the logarithm of incomes of systematic age, schooling, and gender effects. The qualitative results and the general quantitative implications were similar to those reported below because so little of the variation in incomes is removed by these variables. Thus the R^2 in these estimated equations ranged from 0.16 in 1995 to 0.09 in 1990.

in adjacent years are also incorporated in ω .

Permanent income is assumed to have the following structure:

(5) $$ln[y^P_i(t)] = \alpha + \mu(i, t).\overline{ln(y_i)}$$

where $\overline{ln(y_i)}$ is each individual's average value of the logarithm of income over the six years from 1990 to 1995. $\mu(i,t)$ is a parameter that varies across individuals and registers calendar year shocks to estimates of permanent income. α is a constant. The average value of an individual's income over a number of years is probably the best observable indicator of his permanent income.[41] The dependence of the parameter μ on time recognizes shocks that affect all individuals in a comparable manner. For example, new regulations governing labor market activities or movements in exchange rates may change the value of μ . The multiplicative parameter μ varies not only over time but also over individuals as they age which allows the meaning of any given value of $\overline{ln(y_i)}$ to change over the life cycle. Combining equations (3), (4), and (5) yields a first-order difference equation with time varying parameters

(6) $$ln[y_i(t)] = \alpha\,[1 - \omega(t)] + [\mu(i,t) - \omega(t).\mu(i,t\text{-}1)].\overline{ln(y_i)} + \omega(t).ln[y_i(t - 1)] + \varepsilon_i(t)$$

With just six years of data, a very general specification for the time-series variations in μ and ω is not possible. However, to address whether systematic movements are evident in the 1990s, assume μ and ω each follow a trend and μ varies nonlinearly with age, a , as expressed as a quadratic form:

$$\mu(i,t) = \mu_0 + \mu_1.T_t + \mu_2.(a_i) + \mu_3.(a_i)^2 \quad \text{and} \quad \omega(t) = \omega_0 + \omega_1.T_t$$

[41] Friedman's (1957) original work used a moving average of observed income as a measure of permanent income and, more recently, Gottschalk and Moffitt (1994) did something similar.

where T_t denotes a yearly time trend and a_i is individual i's years of age.[42] This specification of μ and ω yields

$$(7)\qquad \begin{aligned} ln[y_i(t)] = {} & \alpha(1-\omega_0) - \alpha\omega_1 T_t + \mu_0(1-\omega_0)\overline{ln(y_i)} + [\mu_1(1-\omega_0) - \mu_0\omega_1](T_t.\overline{ln(y_i)}] \\ & - \mu_1\omega_1[T_t^2.\overline{ln(y_i)}] + \mu_2(1-\omega_0)[a_i.\overline{ln(y_i)}] - \mu_2\omega_1[T_t.a_i.\overline{ln(y_i)}] \\ & + \mu_3(1-\omega_0)[a_i^2.\overline{ln(y_i)}] - \mu_3\omega_1[T_t.a_i^2.\overline{ln(y_i)}] + \omega_0\, ln[y_i(t-1)] \\ & + \omega_1\{T_t.ln[y_i(t-1)]\} + \varepsilon_i(t) \end{aligned}$$

This representation allows a decomposition of any changes in the inequality of incomes into permanent and transitory components. Thus the variance of permanent income is given by

$$(8)\qquad var\{ln[y^P{}_i(t)]\} = [\mu(i,t)]^2 .var[\overline{ln(y_i)}]$$

and, because movements in transitory income and permanent income are independent of one another, once $var\{ln[y^P{}_i(t)]\}$ is determined, so the variance in transitory income, $var\{ln[y^T{}_i(t)]\}$, may be computed. The variance of transitory income is proportional to the variance of $\varepsilon_i(t)$ as follows: $var\{ln[y^T{}_i(t)]\} = (1 - \omega^2)^{-1}.var[\varepsilon_i(t)]$, so that smaller values of ω are associated with a lower transitory income variance.

This model is not rich in parameters compared with some models of the income process that have been estimated recently.[43] Thus, the model specified above has merely seven structural parameters (including the intercept) but six years of data (with the first year effectively "lost" owing to the lagged income term) severely restricts what is practical. The number of cross-section observations is

---

[42] The model was estimated also with a dummy variable for the year 1995 (in addition to the linear time trend) entering both the expressions for μ and ω. An examination of the data suggested the year 1995 seemed a little different from the earlier years and perhaps there is less measurement error in 1995. The implications of this slightly more general model yielded inferences that were very close to those reported in Tables 14, 15, and 16.

[43] The state of the art is illustrated by Meghir and Pistaferri (2005).

considerable - over ten thousand in the case of China - but the identification of the stochastic income processes leans heavily on the time series movements in the variables and the years from 1990 to 1995 provide little opportunity to model these liberally. Though restrictive, this model allows us to determine the magnitudes of the permanent and transitory income variances and, more tentatively, to assess whether these have changed in China in the early 1990s.

The estimates of the parameters for urban China are given in the upper panel of Table 14 where the model is fitted to men and women separately in addition to all individuals together. The sign of μ_1 is positive suggesting a rising permanent variance over time and the sign of ω_1 is negative suggesting a smaller transitory income variance over time. (The magnitude of the transitory income variance depends not only on the value of ω, but also on the value of $var[\varepsilon_i(t)]$.) The coefficients attached to years of age imply that the factor, μ, relating average observed incomes to permanent income rises with age. In other words, average observed income provides a more accurate indicator of permanent income for older than for younger people, something eminently plausible.[44]

The implications of these parameters for the permanent income variance, the transitory income variance, and the ratio of the permanent income variance to the total income variance are reported in Table 15 for each year from 1990 to 1995. Because our specification allows for effects that vary with the ages of individuals, these estimates of the variance in permanent and transitory incomes are different in the same calendar year for people of different ages. Therefore, the values reported in Table 15 for

---

[44] More precisely, μ_2 is positive and μ_3 is negative suggesting μ follows a concave relationship with age, first rising and then falling. In fact, the estimates imply that μ reaches a maximum at an age older than seventy years, the oldest individuals in the sample. Hence, over all observed ages, μ rises with age. On individual t-tests, the coefficients attached to μ_2 and μ_3 would not be judged as significantly different from zero by conventional statistical criteria. However, a joint test on μ_2 and μ_3 is more meaningful and, on a likelihood ratio test, the hypothesis of no relation between μ and age is soundly rejected at less than the one percent level.

China vary not only by calendar year but also by age.[45] *var{ln[y* ^P_i *(t)]}*, the permanent income variance, rises over time while *var{ln[y* ^T_i *(t)]}* the transitory income variance tends to fall. Although the transitory variance is larger in 1995 than in previous years for younger people, this seems to be an aberration resulting from an unusually large variance in observed incomes in this single year, 1995, for this cohort. In general, the transitory variance tends to decline over time.

The share of the variance in the logarithm of income that is attributable to the permanent component is given in Table 15 by *S(t)* which grows over time: for individuals aged 33 years in 1990, it rises from 55 percent in 1990 to 81 percent for these individuals aged 38 years in 1995.[46] Note that there is both an age effect and a calendar year effect embodied in these estimates for China. Because the variance of permanent income is a positive function of μ, the positive estimate of μ_1 implies the rising importance of the variance in permanent income with respect to calendar time while the estimates of μ_2 and μ_3 imply the growing importance of permanent income variance with age.[47]

How do these estimates of the relative importance of permanent component of the variance of income compare with those for the United States? First, the model specified and estimated to the urban Chinese is fitted to the 1,847 urban Americans from the PSID. The estimates for the U.S. are contained in the bottom panel of Table 14. Because a likelihood ratio test cannot reject the hypothesis that μ is

---

[45] Those labelled "aged 24 years in 1995" are, more precisely, those aged from 22 to 26 years in 1995. Those labelled as "aged 38 years in 1995" and "aged 53 years in 1995" are, respectively, aged from 36 to 40 years and from 51 to 55 years in 1995.

[46] It tends to grow also for those aged 24 years in 1995 except for the outlier year 1995. As noted earlier, 1995 exhibits an singularly large income variance for this cohort in 1995.

[47] Differences in the variance of permanent income by age are induced not only by the estimates of μ_2 and μ_3, but also by the values of $var[\overline{ln(y_i)}]$ that differ by age. $var[\overline{ln(y_i)}]$ is 0.221 for those aged 24 years in 1995, 0.185 for those aged 38 years in 1995, and is 0.214 for those aged 53 years in 1995.

independent of age, the age variable is omitted in constructing estimates of the permanent income variance.[48] As is the case of urban China, these estimates for the United States imply a rising trend in μ (that is, μ_1 is positive) and a falling trend in ω (that is, ω_1 is negative). The implications of these estimates for the share of the variance in the logarithm of income attributable to the permanent component for the United States is given at the bottom of Table 15 by *S(t)*. The average of the values is 0.830 which is some ten percent larger than the average of the values for China. Remarkably, this 83 percent for urban United States is exactly the value that Friedman (1957) proposed for U.S. urban families.[49]

This permanent-transitory model has implications not only for variances but also for covariances which may be expressed in terms of correlation coefficients. Let $\rho_{ln}(t, t+\tau)$ denote the correlation coefficient between the logarithm of incomes in year t and year $t+\tau$ so the permanent-transitory model with first-order serial correlation in transitory income implies

(9) $$\rho_{ln}(t, t+\tau) = S + \omega^{\tau}(1 - S)$$

where, again, S is the share of the variance in the logarithm of income attributable to the permanent component. The performance of this estimated model may be assessed by comparing its implied values for the correlation coefficient of the logarithm of incomes with the observed values. Using the estimates of S and ω in Tables 14 and 15, the implied and observed values of $\rho_{ln}(t, t+\tau)$ are reported

---

[48] The point estimates of the permanent income variance are affected little by including the age variables. In the U.S. case, the estimates are formed by combining income observations on the 1,847 people for the years from 1993 to 1996. We lack income observations for the PSID in 1997.

[49] As an estimate of the relative importance of permanent income, Friedman (1957, p. 190) proposed 0.83 as a "reasonably typical value" for nonfarm consumer units. More recent estimates suggest a smaller value for the relative importance of permanent income. Using earnings data on men, Haider (2001) estimated S at about seventy percent in 1990 while Gottschalk and Moffitt (1994) offer a value of about two-thirds in the 1980s though these more recent estimates are not restricted to urban workers.

in Table 16 for China and in Table 17 for the United States.[50] Though some values of the implied correlation coefficients come close to the observed values, in general, the correspondence is unimpressive. For both countries, the observed correlation coefficients are a steeper negative function of years apart than are those using the estimates of equation (7): thus, for China, in most instances, the observed correlation coefficients are higher for adjacent years than those implied by equation (7)'s estimates; on the other hand, the observed correlation coefficients are lower for intervals of four or more years apart than those implied by the estimates of equation (7). In the case of the United States in Table 17, the estimated small and negative autocorrelation of incomes (that is, the negative estimates of ω) has the effect of checking the decline in the correlation between incomes in years farther apart.

To conclude, the share of the variance in the logarithm of income attributable to the permanent component in China is somewhat smaller than that implied for the United States. The importance of the permanent component appears to have been growing moderately in China which is consistent with the belief that wages have increasingly reflected the relative scarcity of human skills and, because these skills are durable, the variance of wages embodies a larger permanent component. In addition, in the early 1990s, the wage component of incomes rose while the contribution to income of housing and other subsidies fell[51] and this change in the structure of incomes may well have caused a change in the relative importance of the variance of permanent income. However, it must be emphasized that, as noted in Section II, our full respondents (that is, those who provided information on incomes in all six years) are a more stable group of individuals - they are more likely to be married, middle-aged, Communist Party members, and drawn from the middle of the income distribution. Therefore, at this

---

[50] Because the parameters have been estimated by permitting them to change over time, the values of ω and S entered into equation (9) for the correlation coefficient between log incomes in years t and $t + \tau$ are the simple average values of ω and S for year t and year $t + \tau$.

[51] See Khan and Riskin (2001), pp. 20-2.

stage of the inquiry, it would be courageous to extend these findings to the population as a whole.

VI. Conclusions

Measures of income inequality in a single year suggest that inequality increased between the late 1980s and the mid-1990s in China. Was this increase in annual income inequality accompanied by greater income mobility? Was it the variance of permanent or transitory incomes that increased? To address these and other issues, information on the incomes of the same people over time are needed.

The information in this paper on the incomes of the same people are restricted to urban residents. They suffer from the defect that the sample of respondents is not fully representative of the population. This defect is common to panel data. Married, middle-aged, people from the middle of the income distribution in 1995 in China are more likely to be represented in the analysis of income mobility than unmarried, younger and older, people with incomes towards the top and bottom of the income distribution. The income distribution of the people in the income mobility analysis exhibits less dispersion in income than the income distribution of all urban residents; furthermore, the income distribution of the individuals in the income mobility analysis tends to narrow rather than widen over the years before 1995. The income differences may not be viewed as large, but nevertheless they call for caution in drawing inferences from the sample of people in the income mobility study to the population of all urban residents.

Subject to these important qualifications, the analysis of income mobility in urban China in this paper reaches a number of conclusions. First, the research in Section III indicates that the level of income mobility across quintiles or clusters in China in the 1990s was greater than that in the United States or other high income countries. Second, the correlation coefficients between incomes in different years are lower - suggesting greater income mobility - than those in the United States when the correlation coefficients are computed between incomes three or more years apart. Third, the correlation coefficients between incomes in different years were falling in China: holding constant age, schooling, and gender, the correlation coefficient between incomes in 1993 and 1995 is 0.17 lower than the correlation coefficient between incomes in 1990 and 1992. Because falling correlation coefficients are associated with rising income mobility, this implies that income mobility increased in urban China from 1990 to 1995. Fourth, Section IV suggests there is a slight tendency for income increases to have been greater for those with low permanent incomes though this average relationship conceals a wide variety of very different experiences at each income level. In computing whether changes in income have tended to benefit those with lower or higher incomes, for both China and the United States, we demonstrate that it is very important to recognize how transitory shocks may affect our inferences. Section V estimates the relative importance of permanent income in China in the 1990s to be lower than that in the United States. In addition, the share of the variance in incomes attributable to the permanent component increased mildly in the 1990s in urban China. This is consistent with an economy moving closer to a system in which incomes are linked to durable skills.

It needs emphasizing that this paper has quantified the extent of income mobility in urban China in the first half of the 1990s, but it has not identified the mechanisms through which this mobility is effected. For instance, we have not traced the degree to which changes in income have taken place through the mobility of workers across firms compared with the changes in income of workers in the same firm. However, we may speculate how income mobility was realized in China in the early 1990s.

The extent of worker mobility across firms in urban China in the 1990s appears to have been low compared with that in other countries with the result that job tenures in China were long.[52] The

---

[52] Knight and Yueh (2004) calculate the median tenure of urban residents in China to be 19 years in 1999 (though that of rural migrants to urban areas is 3 years). The OECD (OECD, 1997) report median tenure in other countries in the mid-1990s to be substantially less than this in most cases: median tenure in the U.S.A. is 6.7 years, Britain 4.2 years, Japan 8.3 years, and Germany 10.7 years.

suggestion is that, even though labor mobility has been rising in China, little income mobility was effected in the early 1990s through job switches across firms. Therefore, year-to-year income mobility is to be sought principally in the changing incomes of people who remain with the same firm. The 1990s saw more authority granted to individual enterprises to set wages independently and there appears to be a positive relation between the profits of firms and the earnings of employees. Such rent-sharing is by no means singular to China but it manifests itself in an explicit form in China where bonus payments are linked to a firm's profits and where workers with long tenure seem to be rewarded disproportionately (see Knight and Song (2005), pp. 155-63).[53]

Though in the research to date we have not traced out the mechanism by which increases in income mobility were effected in urban China in the first half of the 1990s, there is information that (we conjecture) will allow us to be less agnostic. This is a research project for the future.

It must also be stressed that our results pertain to China in the early 1990s and the principal comparison here is with the United States at approximately the same time. The first half of the 1990s constituted a period of remarkable - almost singular - income growth in urban China so it would be rash to extend our general conclusions here to years less distinctive. Clearly more data for other years on individual incomes - preferably panel data - are needed to determine whether the first half of the 1990s was an unusual period for income mobility in China. In any event, the discussion on income inequality needs modification with less stress on the patterns of annual incomes and with greater emphasis on longer-run measures of income inequality.

---

[53] Knight and Song (2005, pp. 164-5) write, "The relationship between profits and wages also contributed to a widening of wage inequalities over the late 1990s. Not only were employees of loss-making firms paid less than those of profit-making firms in 1995 but also their wages grew less rapidly over the subsequent four years, and among them it was particularly the unskilled workers with low-paying characteristics who fell behind". Also see Meng (2002).

REFERENCES

Aaberge, Rolf, Anders Björklund, Markus Jäntti, Marten Palme, Peder J. Pedersen, Nina Smith, and Tom Wennemo. (2002). “Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States”, Review of Income and Wealth, 48 (4), December, pp. 443-69.

Baulch, Bob, and John Hoddinott. (2000). “Economic Mobility and Poverty Dynamics in Developing Countries”, The Journal of Development Studies, 36 (6), August, pp. 1-24.

Becketti, Sean, William Gould, Lee Lillard, and Finis Welch. (1988). “The Panel Study of Income Dynamics after Fourteen Years: An Evaluation”, Journal of Labor Economics, 6 (4), October, pp. 472-92.

Benjamin, Dwayne, Loren Brandt, and John Giles. (2005). “The Evolution of Income Inequality in Rural China”, Economic Development and Cultural Change, 53 (4), July, pp. 769-824.

Bishop, Yvonne, M.M., Stephen E. Feinberg, and Paul W. Holland. (1975). Discrete Multivariate Analysis: Theory and Practice. Cambridge, MA.: MIT Press.

Burkhauser, Richard, Douglas Holtz-Eakin, and Stephen Rhody. (1998). “Mobility and Inequality in the 1980s: A Cross National Comparison of the United States and Germany”, in Stephen P. Jenkins, Arie Kapteyn, and Bernard M. S. van Praag. (eds.), The Distribution of Welfare and Household Production: International Perspectives. Cambridge: Cambridge University Press, pp. 111-75.

Démurger, Sylvie, Martin Fournier, and Li She. (2005). “Urban Income Inequality in China Revisited, 1988-2002", unpublished paper, April.

Fabig, Holger. (2000) “Labor Income Mobility - Germany, the USA and Great Britain Compared”, in Richard Hauser, and Irene Becker. (eds.) The Personal Distribution of Income in an International Perspective, Springer-Verlag Berlin, pp. 31-55.

Fields, Gary S. (2001) Distribution and Development: A New Look at the Developing World, Cambridge, MA: MIT Press.

Fields, Gary S., Paul L. Cichello, Samuel Freije, Marta Menéndez, and David Newhouse. (2003). “Household Income Dynamics: A Four-Country Story”, The Journal of Development Studies, 40(2), December, pp. 30-54.

Fields, Gary S., María Laura Sánchez Puerta, Robert Duval Hernández, and Samuel Freije. (2005). “Earnings Mobility in Argentina, Mexico, and Venezuela: Testing the Divergence of Earnings and the Symmetry of Mobility Hypotheses”, unpublished paper, May.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. (1998). “An Analysis of Sample Attrition in

Panel Data", Journal of Human Resources, 33 (2), Spring, pp. 251-99.

Flemming, J.S., and John Micklewright. (2000). "Income Distribution, Economic Systems, and Transition", in Anthony B. Atkinson, and François Bourguignon. (eds.) Handbook of Income Distribution Volume 1, Amsterdam: Elsevier, pp. 843-918.

Friedman, Milton. (1957). A Theory of the Consumption Function, Princeton University Press, National Bureau of Economic Research.

Friedman, Milton. (1962). Capitalism and Freedom, Chicago: University of Chicago Press.

Gibson, John, Jikun Huang, and Scott Rozelle. (2003). "Improving Estimates of Inequality and Poverty from Urban China's Household Income and Expenditure Survey", Review of Income and Wealth, 49 (1), March, pp. 53-68.

Gottschalk, Peter, and Robert Moffitt. (1994). "The Growth of Earnings Instability in the U.S. Labor Market", Brookings Papers on Economic Activity, 2, pp. 217-72.

Gustafsson, Björn, and Li Shi. (2001). "The Anatomy of Rising Earnings Inequality in Urban China", Journal of Comparative Economics, 29, pp. 118-135.

Haider, Steven J. (2001). "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991", Journal of Labor Economics, 19 (4), October, pp. 799-836.

Jalan, Jyotsna, and Martin Ravallion. (2000). "Is Transient Poverty Different? Evidence for Rural China", The Journal of Development Studies, 36 (6), August, pp. 82-99.

Khan, Azizur Rahman, and Carl Riskin. (1998). "Income and Inequality in China: Composition, Distribution, and Growth of Household Income, 1988 to 1995", The China Quarterly, 154, June, pp. 221-53.

Khan, Azizur Rahman, and Carl Riskin. (2001). Inequality and Poverty in China in the Age of Globalization, New York: Oxford University Press.

Khor, Niny, and John Pencavel. (2005). "Income Disparities and Income Mobility in China", paper delivered at the Conference on China's Policy Reforms: Progress and Challenges, Stanford University Center for International Development, 1 October.

Knight, John, and Lina Song. (2003). "Increasing Urban Wage Inequality in China", Economics of Transition, 11 (4), pp. 597-619.

Knight, John, and Lina Song. (2005). Towards a Labour Market in China, Oxford: Oxford University Press.

Knight, John, and Linda Yueh. (2004). "Job Mobility of Residents and Migrants in Urban China", Journal of Comparative Economics, 32 (4), December, pp. 637-660.

MaCurdy, Thomas, Thomas Mroz, and R. Mark Gritz. (1998). "An Evaluation of the National Longitudinal Survey of Youth", Journal of Human Resources, 33 (2), Spring, pp. 345-436.

Meghir, Costas, and Luigi Pistaferri. (2004). "Income Variance Dynamics and Heterogeneity", Econometrica, 72 (1), January, pp. 1-32.

Meng, Xin. (2002). "Profit Sharing and the Earnings Gap between Urban and Rural-Migrant Workers in Chinese Enterprises", unpublished paper, April.

Meng, Xin. (2004). "Economic Restructuring and Income Inequality in Urban China", Review of Income and Wealth, 50 (3), September, pp. 357-79.

Metcalf, David, and Jianwei Li. (2005). "Chinese Unions: Nugatory or Transforming? An *Alice* Analysis", London School of Economics, unpublished paper, September.

O.E.C.D. (1996). "Earnings Inequality, Low-Paid Employment, and Earnings Mobility", Employment Outlook, July, pp. 59-108.

O.E.C.D. (1997). "Is Job Insecurity on the Increase in OECD Countries? " Employment Outlook, July, pp. 129-60.

Riskin, Carl, Zhao Renwei, and Li Shi. (2000). "Chinese Household Income Project, 1995" [computer file], ICPSR version, Amherst, MA, University of Massachusetts, Political Economy Research Institute [producer], 2000. Inter-university Consortium for Political and Social Research [distributor], Ann Arbor, Michigan, November.

Zhou, Xueguang. (2000). "Economic Transformation and Income Inequality in Urban China: Evidence from Panel Data", The American Journal of Sociology, 105 (4), January, pp. 1135-74.

APPENDIX

Chinese Data

The data for the Chinese households and individuals in this study come from the Chinese Household Income Project 1995, publicly available through the Inter-university Consortium for Political and Social Research and more fully described in the relevant codebooks. Total labor market income is the income variable used in the analysis of the incomes of individuals in urban areas. It excludes property income and transfer income and neglects any income taxes paid. Labor market income represents about 95 percent of gross annual income. To address outliers, the data were then trimmed by dropping the top 0.5 percent and the bottom 0.5 percent of individuals' incomes.

A considerable amount of time was spent in verifying the accuracy of the data and in "cleaning" the data to eliminate measurement error. Essentially this was effected by examining the sequence of income over the years for each individual and identifying unexpected or odd values. Sometimes these individuals were dropped from the sample. In other cases, it seemed reasonable to alter the recorded entry on income. This would occur when, for instance, zeros were missing in a single year. Our procedure was as follows, for each individual, the income series over the years from 1990 to 1995 was examined for unexpected or odd values. To do this, we compute the individual's real income in each year as a ratio of the average of the individual's real income over the six years. Next, we create flags for outliers: if real income in a year is greater than 200% of average real income between 1990 and 1995 or smaller than 33.3%, then we flag that observation as "unexpected". (We examined other thresholds too.) Using the observations that are not unexpected (in this case, n = 10,066), we fit a regression equation to predict the logarithm of income using average income from the other years (not including income from the year being predicted) and interactions with age groups (in 5 year intervals), educational levels, gender, and province dummy variables. We predict out of sample values for those observations flagged as unexpected. We 'guess' the right level of income by taking the original reported income and adjusting it by multiples of ten (the particular magnitude being determined by comparing the observation to the rest of the sequence). We compare the 'guess' with the predicted value from the regression. If the 'guess' estimate is closer to the predicted value than the original reported income, we use this value instead. This process results in 10,184 valid observations.

For all 1995 data, all respondents are asked two questions pertaining to schooling: the educational level attained and the number of years of schooling received (not including years spent on repeating a grade). There is a considerable variation in years of schooling reported for each level of educational attainment. See Khor and Pencavel (2005) for how this was addressed.

All monetary values are converted to 1995 values using the consumer price index for all respondents as given by the China Statistical Yearbook 1996 (Table 8-1, p. 255).

Data for the United States

The Panel Study of Income Dynamics data come from the online data center.[54] The characteristics of these U.S. data were limited to conform to those applied to the Chinese data. Owing to changes in reported income variables, our sample for the PSID includes the later survey years of

---

[54] http://psidonline.isr.umich.edu/

1994 to 1999.[55] Initially, this includes 38,141 individuals per year. The sample size was reduced because of a number of changes made to the PSID in 1997. Thus, the number of heads of households fell from 10,972 in 1994 to 7,176 in 1999.

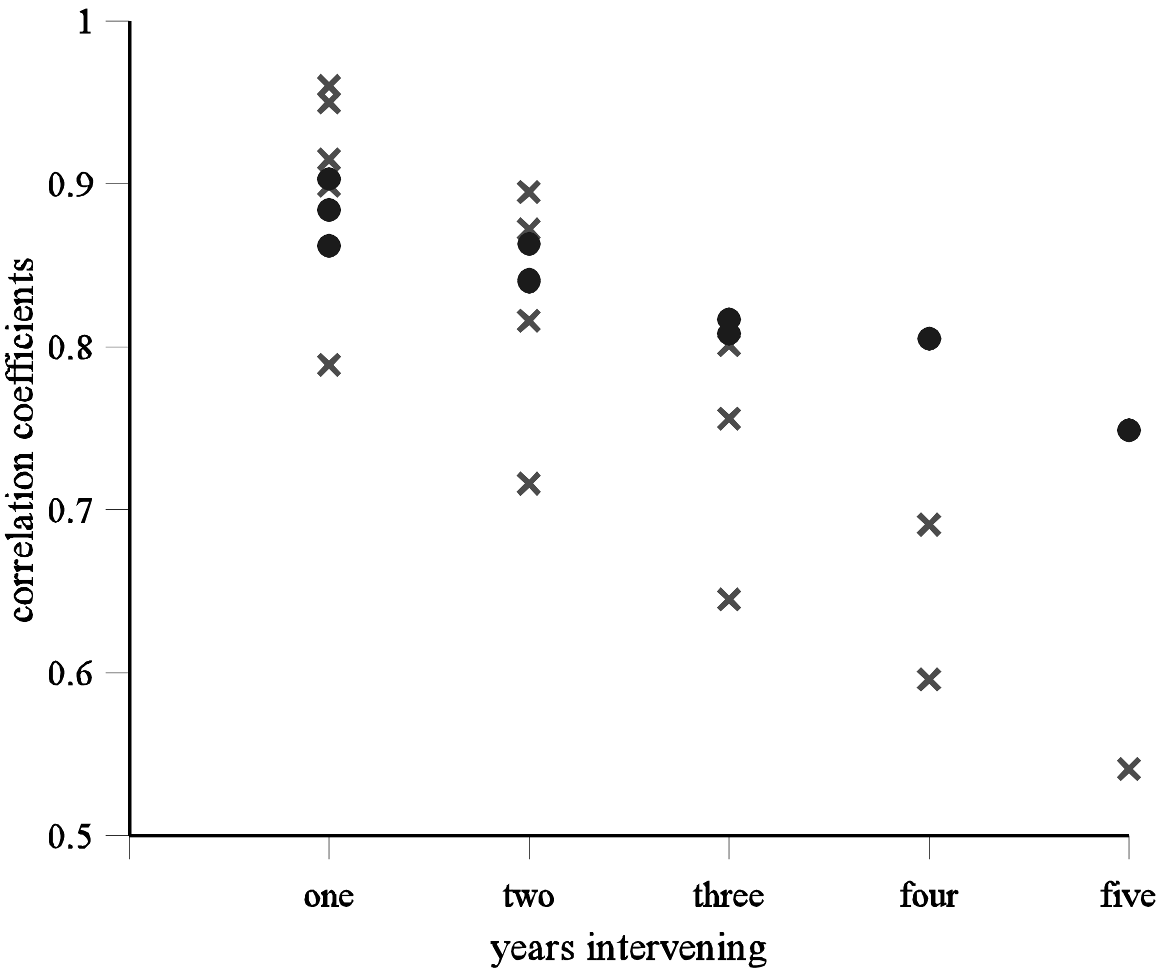
The income variable we use is from the Income Plus files which contain data on family income and its components, notably the labor earnings of the head and spouse. Because we seek to construct a balanced panel on income, we restrict the sample of the individuals to those who are always in the sample and who are either heads of households or their respective spouses. This results in the following composition of the sample:

relationship	men	women	total
household head	3,504	1,236	4,740
legal wife	0	2,556	2,556
cohabiting	0	51	51
total	3,504	3,843	7,347

Next we limit the sample by retaining only those who earn positive income for all the observed years. This leaves 2,374 men and 2,048 women. To maintain comparability with urban China, we drop 71 observations in rural areas and we lose 41 people by limiting people to those aged less than 70 years in 1998 (survey year 1999). We restricted the sample to those who report working continuously over all the sample years and this eliminates an additional 854 observations. Eliminating the over-sampled population further reduces the sample size by 1,482 observations. After trimming the data by dropping the top and bottom 0.5% of incomes, our final sample consists of 1,847 observations (1,038 men and 809 women).

[55] No survey of incomes was conducted in 1998.

Figure 1
Correlation Coefficients between Incomes by Years Intervening: China and the United States



The correlation coefficients for China are given by a cross and those for the United States by a circle.

Figure 2
Change in the Logarithm of Income between 1990 and 1995 Against Income Percentile: China

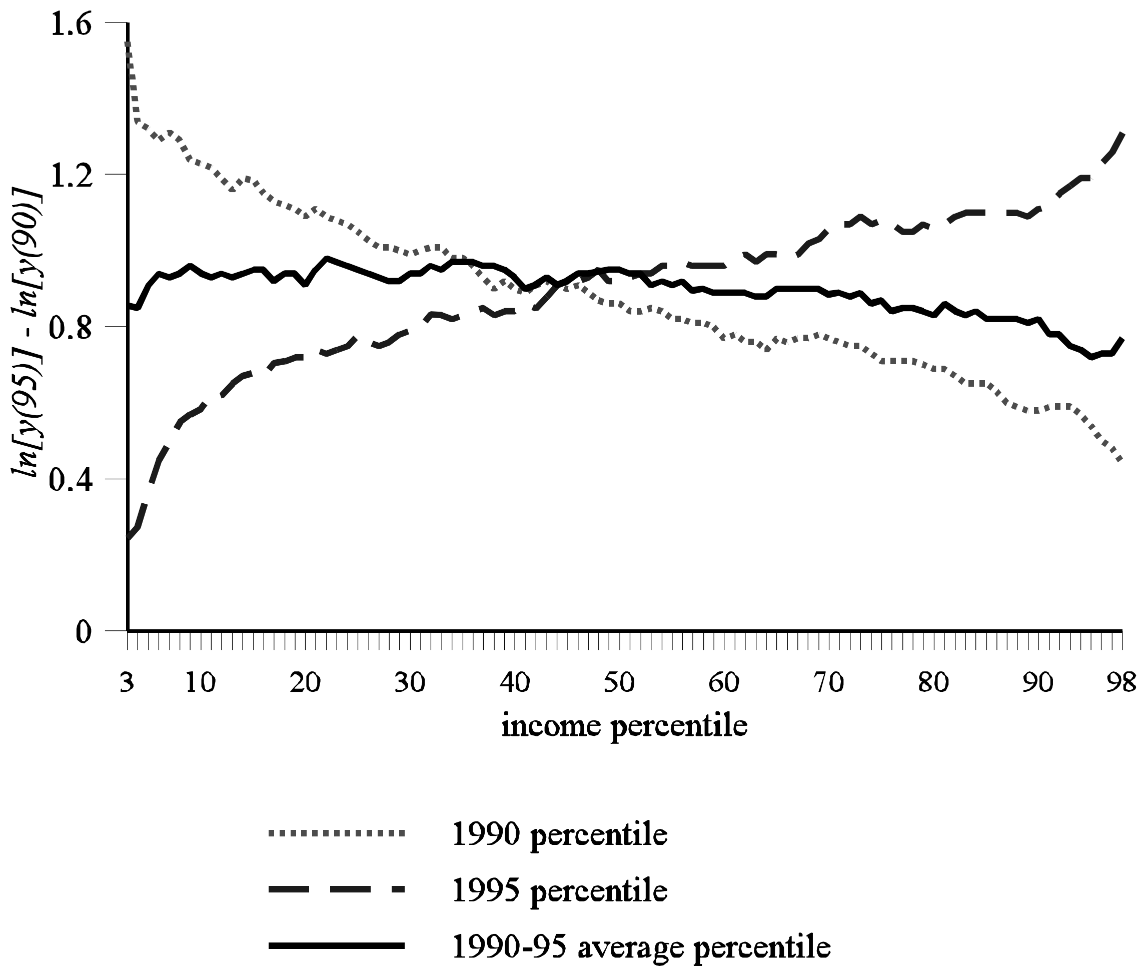


Figure 3
Change in the Logarithm of Income between 1993 and 1998 Against Income Percentile: U.S.A.

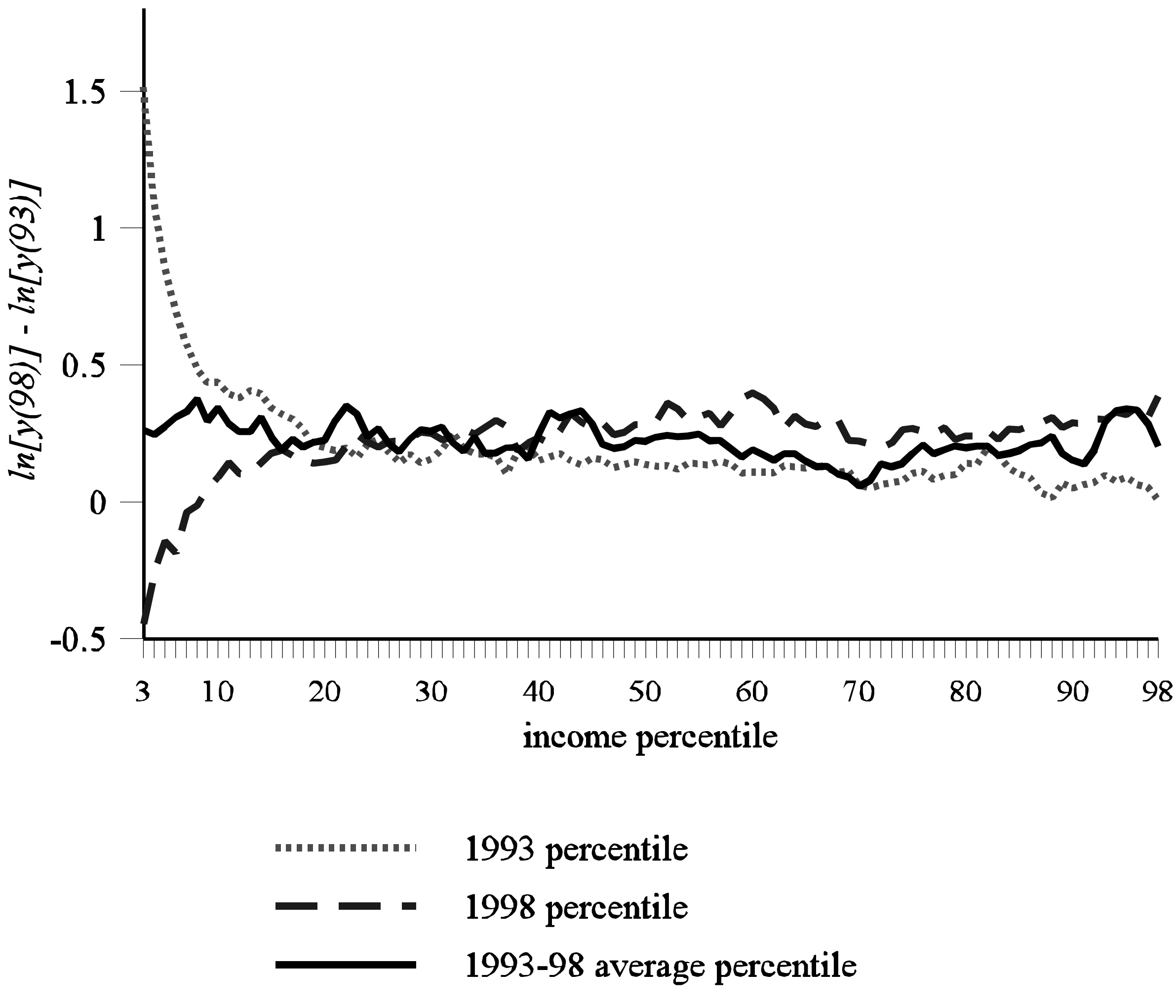


Figure 4
The Range of Values of the Change in the Logarithm of Income between 1995 and 1990: China
Against Average Income Percentile: China

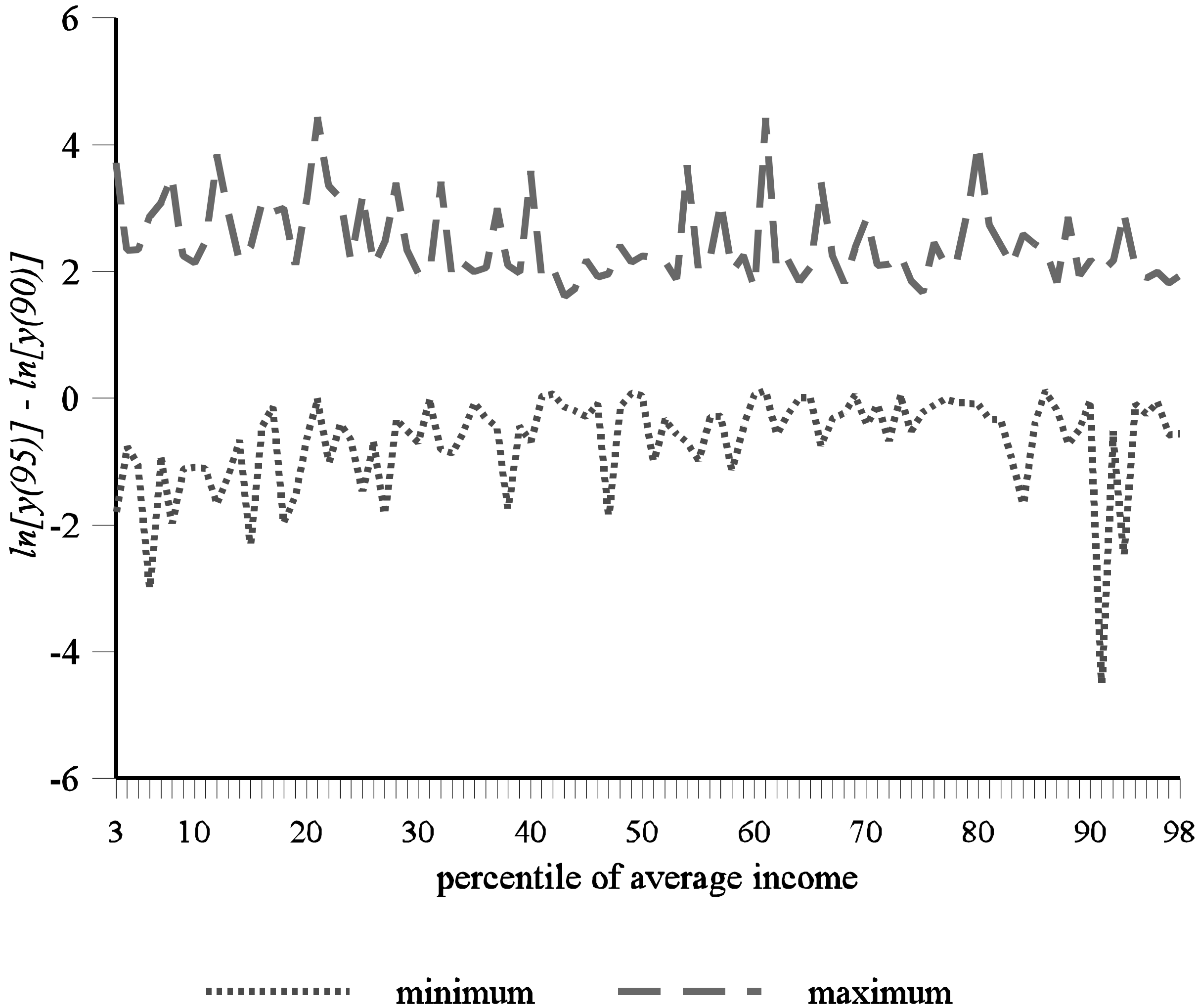


Table 1
Marginal Effects from Logit Estimates of the Probability of Providing Information on Income in All Six Years: China

age	0.026 (0.002)		< 10 percentile	reference
(age squared)/100	-0.028 (0.003)		10-20 percentile	0.032 (0.008)
woman	-0.009 (0.005)		20-30 percentile	0.042 (0.007)
married	0.144 (0.016)		30-40 percentile	0.043 (0.008)
Communist Party	0.019 (0.007)		40-50 percentile	0.038 (0.008)
ethnic minority	0.027 (0.011)		50-60 percentile	0.036 (0.008)
schooling 1	-0.104 (0.030)		60-70 percentile	0.013 (0.010)
schooling 2	-0.065 (0.023)		70-80 percentile	0.021 (0.010)
schooling 3	-0.049 (0.021)		80-90 percentile	-0.001 (0.011)
schooling 4	-0.033 (0.018)		> 90 percentile	0.008 (0.011)
schooling 5	-0.031 (0.017)			
schooling 6	reference			

There are 11,571 observations in this regression. Estimated standard errors are in parentheses. For continuous variables, marginal effects are partial derivatives while, for discrete variables, the effects are of a change in the value of the dummy variable from zero to unity. These effects are evaluated at the mean values of the right-hand side variables. “Age” measures years of age. All the other variables are dichotomous variables. “Woman” takes the value of unity for a woman, “married” takes the value of unity for a currently married individual, “Communist Party” takes the value of unity for someone who is a member of the Communist Party, and “ethnic minority” that takes the value of unity for someone who reports being an ethnic minority. “schooling1" takes the value of unity for someone with a college education, “schooling2" takes the value of unity for someone with a professional school education, “schooling3" takes the value of unity for someone with a middle level professional, technical or vocational school education, "schooling4" takes the value of unity for someone with an upper middle school education, and "schooling5" takes the value of unity for someone with a lower middle school education. "schooling6", the omitted dichotomous variable, refers to elementary or below elementary school. The variables taking the form “x-y percentile” are dichotomous variables that take the value of unity for someone with an income in 1995 in the percentile range between x and y. The lowest tenth percentile constitutes the reference category.

Table 2
Measures of Income Inequality in 1995 for Two Groups: China

	individuals providing information on income in all six years (nobs=10,184)	individuals providing information on income in year 1995 (nobs=11,571)
Gini coefficient	0.274	0.282
ratio of 90th to 10th percentile	3.551	3.742
coefficient of variation	0.520	0.534
standard deviation of logarithm of incomes	0.548	0.596

Table 3
Annual Income Inequality among Those who Provide Income Information in All Years from 1990 to 1995: China

	1990	1991	1992	1993	1994	1995
Gini coefficient	0.289	0.282	0.276	0.269	0.264	0.274
90th/10th percentile	4.000	3.711	3.636	3.500	3.415	3.551
coefficient of variation	0.553	0.540	0.530	0.515	0.506	0.520
standard deviation of log income	0.549	0.531	0.527	0.523	0.514	0.548

Table 4
Income Transition Matrix for China: All Individuals, Men, and Women, 1990-1995

All Individuals (10,184)						
		year 1995				
		I	II	III	IV	V
year 1990	I	0.439	0.219	0.177	0.115	0.049
	II	0.277	0.260	0.203	0.161	0.099
	III	0.187	0.242	0.227	0.208	0.136
	IV	0.076	0.206	0.249	0.249	0.220
	V	0.021	0.073	0.144	0.266	0.496

All Men (5,372)						
		year 1995				
		I	II	III	IV	V
year 1990	I	0.436	0.228	0.175	0.115	0.046
	II	0.274	0.268	0.187	0.174	0.097
	III	0.184	0.227	0.237	0.215	0.136
	IV	0.087	0.209	0.238	0.235	0.230
	V	0.020	0.067	0.162	0.262	0.491

All Women (4,812)						
		year 1995				
		I	II	III	IV	V
year 1990	I	0.436	0.200	0.179	0.141	0.045
	II	0.266	0.264	0.186	0.164	0.119
	III	0.186	0.251	0.244	0.184	0.135
	IV	0.089	0.219	0.232	0.255	0.204
	V	0.023	0.067	0.159	0.256	0.497

Table 5
Income Transition Matrix for the United States: All Individuals, Men, and Women, 1993-98

All Individuals (1,847)						
		year 1998				
		I	II	III	IV	V
year 1993	I	0.591	0.236	0.111	0.043	0.019
	II	0.290	0.417	0.192	0.076	0.024
	III	0.081	0.274	0.420	0.171	0.054
	IV	0.030	0.054	0.236	0.480	0.199
	V	0.008	0.019	0.041	0.230	0.704

All Men (1,038)						
		year 1998				
		I	II	III	IV	V
year 1993	I	0.615	0.216	0.111	0.034	0.024
	II	0.269	0.428	0.183	0.082	0.039
	III	0.087	0.260	0.409	0.192	0.053
	IV	0.019	0.087	0.245	0.457	0.194
	V	0.010	0.010	0.053	0.236	0.689

All Women (809)						
		year 1998				
		I	II	III	IV	V
year 1993	I	0.525	0.241	0.130	0.086	0.019
	II	0.315	0.395	0.185	0.068	0.037
	III	0.086	0.309	0.377	0.167	0.062
	IV	0.049	0.037	0.259	0.469	0.186
	V	0.025	0.019	0.049	0.210	0.696

Table 6
Indicators of Income Mobility for China

	(1)			(2)			(3)		
	income quintiles			income quintiles of residuals			income clusters		
	(A) ave move	(B) immob -ility ratio	(C) stayers + one movers	(A) ave move	(B) immob -ility ratio	(C) stayers + one movers	(A) ave move	(B) immob -ility ratio	(C) stayers + one movers
all	1.056	0.334	0.711	1.144	0.307	0.681	1.027	0.350	0.729
men	1.060	0.333	0.706	1.147	0.308	0.676	1.028	0.336	0.723
women	1.077	0.339	0.695	1.143	0.308	0.684	1.059	0.336	0.719
age < 30 years	1.096	0.325	0.700	1.133	0.307	0.690	1.130	0.318	0.686
age 30-50 years	1.099	0.321	0.699	1.147	0.307	0.681	1.048	0.339	0.727
age > 50 years	1.096	0.320	0.691	1.147	0.304	0.672	1.039	0.315	0.712
schooling: low	1.011	0.361	0.721	1.089	0.332	0.702	1.019	0.358	0.730
schooling: middle	1.069	0.331	0.705	1.158	0.298	0.678	1.040	0.330	0.725
schooling: high	1.137	0.309	0.673	1.220	0.285	0.647	1.085	0.317	0.708
coastal provinces	1.041	0.331	0.716	1.061	0.335	0.706	1.029	0.343	0.730
interior provinces	1.152	0.313	0.666	1.237	0.293	0.642	1.079	0.311	0.716

In this table, “ave move” means the average quintile or cluster move defined in footnote 15. The “immobility ratio” is the fraction who remain in the same quintile or in the same cluster. “stayers + one movers” is the fraction who remain in the same quintile or cluster plus the fraction who move one quintile or one cluster. “Low” schooling are those with lower secondary schooling or less. “Middle” schooling are those who left school after upper secondary or vocational school. “High” schooling are those with a college education.

Table 7
Indicators of Income Mobility across Income Quintiles
over Six Years in China, the United States, and Six Other Countries

	(A)	(B)	(C)	(D)
China	1.056	0.334	0.711	0.541
U.S.A. - 1	0.624	0.522	0.888	0.749
U.S.A. - 2	0.660	0.514	0.868	0.685
U.K.	0.660	0.514	0.868	0.726
Sweden	0.684	0.505	0.866	0.741
Italy	0.685	0.503	0.857	0.782
Germany	0.647	0.523	0.876	0.777
France	0.683	0.530	0.854	0.718
Denmark	0.812	0.462	0.810	0.615

(A) - average quartile move
(B) - immobility ratio
(C) - fraction of stayers plus those who moved one quartile
(D) - correlation coefficient between incomes six years apart

Notes: In China, mobility is measured from 1990 to 1995. For U.S.A. - 1, mobility is measured from 1993 to 1998. For all the other countries, mobility is measured from 1986 to 1991. In China and U.S.A.-1, income is annual; in U.S.A.-2 and Denmark, the income measure is gross weekly earnings; elsewhere it is gross monthly earnings. These are household surveys for China, Germany, Sweden, and the U.S.A.; administrative data for Denmark, France, and Italy; and a survey of establishments for Britain. French and Italian data omit government workers. Non-workers in either year excluded. The data for the countries other than China and U.S.A.-1 appear in OECD (1996).

Table 8
Indicators of Income Mobility across Five Income Clusters
over Six Years in China, the United States, and Six Other Countries

	(E)	(F)	(G)
China	1.027	0.350	0.729
U.S.A. - 1	0.657	0.544	0.872
U.S.A. - 2	0.784	0.478	0.828
U.K.	0.697	0.482	0.858
Sweden	0.468	0.616	0.937
Italy	0.524	0.556	0.931
Germany	0.541	0.553	0.929
France	0.506	0.605	0.917
Denmark	0.555	0.552	0.912

(E) - average cluster move (constructed the same way as average quintile move)
(F) - immobility ratio (i.e., fraction who remained in the same income cluster)
(G) - fraction of stayers plus those who moved one cluster

Notes: The notes beneath Table 6 apply also to this table with the added condition that, whereas the data for China and for U.S.A. - 1 describe all workers, those in the other rows describe full-time wage and salary workers only.

Table 9
Effects on Income Inequality of Measuring Income over Longer Time Horizons:
China and the United States

income averaged over years	average real income*	Gini coefficient	90^{th} to 10^{th} percentile	coeff. of variation	stan. dev. of log of incomes
CHINA, 1995-1990					
1995	Y 6,597.5	0.274	3.551	0.520	0.548
1995 & ‘94	Y 6,296.3	0.256	3.261	0.486	0.488
1995, ‘94, & ‘93	Y 6,263.3	0.251	3.204	0.478	0.473
1995, ‘94, ‘93, & ‘92	Y 6,181.0	0.250	3.170	0.476	0.466
1995, ‘94, ‘93, ‘92, & ‘91	Y 6,033.9	0.250	3.194	0.476	0.463
1995, ‘94, ‘93, ‘92, ‘91, & ‘90	Y 5,859.2	0.251	3.192	0.477	0.462
U.S.A., 1998-1993					
1998	$ 41,133	0.340	5.233	0.678	0.661
1998 & ‘96	$ 39,351	0.330	4.861	0.648	0.619
1998, ‘96, & ‘95	$ 38,290	0.324	4.661	0.632	0.602
1998, ‘96, ‘95, & ‘94	$ 37,531	0.321	4.563	0.625	0.597
1998, ‘96, ‘95, ‘94, & ‘93	$ 36,768	0.320	4.531	0.622	0.593

* Real income measured in 1995 prices for China (using the consumer price index) and in 1996 prices for the U.S.A. (using the personal consumption expenditures deflator).

Table 10
Correlation Coefficients between Incomes in Different Years: China 1990-95

All Workers					
	1991	1992	1993	1994	1995
1990	0.956	0.895	0.801	0.691	0.541
1991		0.950	0.864	0.756	0.596
1992			0.915	0.816	0.645
1993				0.902	0.716
1994					0.789

Men					
	1991	1992	1993	1994	1995
1990	0.956	0.890	0.793	0.674	0.527
1991		0.948	0.860	0.744	0.585
1992			0.911	0.807	0.634
1993				0.904	0.710
1994					0.782

Women					
	1991	1992	1993	1994	1995
1990	0.954	0.895	0.801	0.697	0.534
1991		0.950	0.862	0.760	0.587
1992			0.916	0.819	0.639
1993				0.894	0.707
1994					0.785

Table 11
Correlation Coefficients between Incomes in Different Years: United States, 1993-1998

All Workers					
	1994	1995	1996	1997	1998
1993	0.862	0.840	0.811	NA	0.749
1994		0.903	0.860	NA	0.805
1995			0.887	NA	0.815
1996				NA	0.841

Men					
	1994	1995	1996	1997	1998
1993	0.849	0.836	0.798	NA	0.721
1994		0.892	0.841	NA	0.775
1995			0.875	NA	0.785
1996				NA	0.806

Women					
	1994	1995	1996	1997	1998
1993	0.836	0.783	0.773	NA	0.713
1994		0.894	0.855	NA	0.799
1995			0.876	NA	0.820
1996				NA	0.872

NA means not available because no income survey was conducted for 1997.

Table 12
Weighted Least-Squares Estimates of the Variables Associated with the
Coefficients of Correlation of Income across Years in China

$$\rho(g\ ;\ a, a + j\ ;\ t, t + j) = \gamma_0 + \gamma_1.(a) + \gamma_2 F + \gamma_3 (Schooling) + \gamma_4 (a - a') + \Sigma_t \delta_t . Y_t + u$$

γ_0	1.067 (0.035)
$\gamma_1 (10)^{-3}$	-0.742 (0.211)
γ_2	0.011 (0.004)
γ_3	0.0036 (0.0028)
γ_4	-0.083 (0.002)
δ_{91}	-0.024 (0.005)
δ_{92}	-0.036 (0.005)
δ_{93}	-0.042 (0.005)
δ_{94}	-0.068 (0.005)
δ_{95}	-0.166 (0.005)
R^2	0.820

Table 13
Least-Squares Estimates of the Relation between Changes in Income and Percentiles of Income: China and the United States

	China			United States		
	(1)	(2)	(3)	(4)	(5)	(6)
constant	1.324 (0.010)	0.504 (0.010)	0.971 (0.011)	0.208 (0.009)	0.0035 (0.0259)	0.2727 (0.0265)
income percentile in initial year	-0.0086 (0.0002)			-0.0024 (0.0001)		
income percentile in final year		0.0077 (0.0002)			0.0043 (0.0004)	
percentile of income averaged over all years			-0.0016 (0.0002)			-0.0009 (0.0004)
R^2	0.207	0.166	0.007	0.124	0.050	0.003

The dependent variable is the logarithmic change in real income between 1995 and 1990 for China and between 1998 and 1993 for the United States. The “initial year” for China is 1990 and for the United States it is 1993. The “final year” for China is 1995 and for the United States it is 1998. “Averaged over all years” means for China the average over the six years from 1990 to 1995 and for the United States the average over the years 1993, 1994, 1995, 1995, and 1998 (as observations on income in 1997 are lacking).

Table 14
Estimates of the Parameters (with Estimated Standard Errors in Parentheses) of the Permanent-Transitory Income Model with First-Order Serial Correlation

	μ_0	μ_1	μ_2 (100)	μ_3 (1,000)	ω_0	ω_1	α
China							
all individuals	0.890 (0.004)	0.0225 (0.0002)	0.0163 (0.0131)	-0.0010 (0.0018)	0.476 (0.010)	-0.0461 (0.0031)	0.478 (0.028)
all men	0.865 (0.006)	0.0224 (0.0003)	0.0248 (0.0170)	-0.0023 (0.0023)	0.498 (0.013)	-0.0559 (0.0042)	0.678 (0.040)
all women	0.910 (0.006)	0.0223 (0.0003)	0.0053 (0.0216)	0.0011 (0.0032)	0.468 (0.015)	-0.0402 (0.0046)	0.337 (0.040)
United States							
all individuals	0.986 (0.005)	0.0029 (0.0003)	0*	0*	-0.030 (0.026)	-0.0268 (0.0140)	0.079 (0.048)

The asterisk denotes parameters constrained to be zero. For the United States, a likelihood ratio test cannot reject the null hypothesis that μ is independent of age.

Table 15
Estimates of the Variance of Permanent and Transitory Income: China and the United States

years	1990	1991	1992	1993	1994	1995
China : all aged 24 years in 1995						
$var\{ln[y^P_i(t)]\}$	0.177	0.186	0.195	0.205	0.214	0.224
$var\{ln[y^T_i(t)]\}$	0.105	0.100	0.077	0.065	0.035	0.086
$S(t)$	0.628	0.649	0.718	0.759	0.859	0.723
China : all aged 38 years in 1995						
$var\{ln[y^P_i(t)]\}$	0.148	0.156	0.164	0.172	0.180	0.189
$var\{ln[y^T_i(t)]\}$	0.124	0.085	0.072	0.055	0.040	0.044
$S(t)$	0.546	0.647	0.695	0.759	0.820	0.810
China: all aged 53 years in 1995						
$var\{ln[y^P_i(t)]\}$	0.172	0.181	0.190	0.199	0.209	0.218
$var\{ln[y^T_i(t)]\}$	0.123	0.098	0.066	0.054	0.041	0.035
$S(t)$	0.583	0.648	0.743	0.788	0.835	0.860
years	1993	1994	1995	1996	1998	
United States						
$var\{ln[y^P_i(t)]\}$	0.362	0.364	0.366	0.368	0.373	
$var\{ln[y^T_i(t)]\}$	0.178	0.059	0.035	0.064	0.064	
$S(t)$	0.670	0.861	0.913	0.852	0.853	

The variance in the logarithm of permanent income is $var\{ln[y^P_i(t)]\}$ and the variance in the logarithm of transitory income is $var\{ln[y^T_i(t)]\}$. The ratio of the variance in the permanent component of log income to the variance in total log income is given by $S(t) = var\{ln[y^P_i(t)]\} / var\{ln[y_i(t)]$. The estimated standard errors are approximately 0.0014 for $var\{ln[y^P_i(t)]\}$ and for $var\{ln[y^T_i(t)]\}$ and 0.005 for $S(t)$ for China. The estimated standard errors are about 0.003 for $var\{ln[y^P_i(t)]\}$ and for $var\{ln[y^T_i(t)]\}$ and about 0.008 for $S(t)$ for the United States.

Table 16
Correlation Coefficients between the Logarithm of Incomes across Years in China:
Comparison of Observed Correlations with Those Implied by an Estimated Permanent-Transitory Model with First-order Serial Correlation of Transitory Income

years	all aged 24 years in 1995		all aged 38 years in 1995		all aged 53 years in 1995	
	implied	observed	implied	observed	implied	observed
1990, 1991	0.802	0.908	0.779	0.909	0.790	0.944
1990, 1992	0.733	0.865	0.691	0.843	0.725	0.884
1990, 1993	0.714	0.760	0.676	0.743	0.707	0.809
1990, 1994	0.749	0.696	0.690	0.623	0.716	0.673
1990, 1995	0.677	0.519	0.680	0.445	0.724	0.483
1991, 1992	0.812	0.915	0.805	0.921	0.819	0.952
1991, 1993	0.748	0.809	0.747	0.826	0.760	0.887
1991, 1994	0.766	0.751	0.746	0.706	0.754	0.743
1991, 1995	0.690	0.565	0.732	0.515	0.757	0.553
1992, 1993	0.833	0.863	0.825	0.882	0.850	0.938
1992, 1994	0.813	0.816	0.785	0.778	0.813	0.806
1992, 1995	0.729	0.618	0.760	0.571	0.808	0.619
1993, 1994	0.869	0.860	0.856	0.881	0.871	0.885
1993, 1995	0.763	0.662	0.803	0.663	0.839	0.676
1994, 1995	0.847	0.764	0.865	0.729	0.889	0.743

Each entry under “observed” reports the correlation coefficient between the logarithm of incomes of individuals in the years listed in the first column. (These values are not the same as those given in Table 10 because (a) the entries in Table 10 represent the correlation of incomes, not the correlation of the logarithm of incomes and (b) the correlation coefficients in this table are specific to the ages listed.) Each entry under “implied” is the correlation coefficient implied by equation (9) which corresponds to a permanent-transitory model with first-order serial correlation of the transitory component.

Table 17
Correlation Coefficients between the Logarithm of Incomes across Years in the United States: Comparison of Observed Correlations with Those Implied by an Estimated Permanent-Transitory Model with First-order Serial Correlation of Transitory Income

	1993		1994		1995		1996	
	implied	observed	implied	observed	implied	observed	implied	observed
1994	0.755	0.810						
1995	0.792	0.773	0.879	0.873				
1996	0.761	0.739	0.857	0.829	0.871	0.856		
1998	0.762	0.677	0.857	0.791	0.883	0.804	0.855	0.824