Income inequality in today's China

Yu Xie[a,b,1] and Xiang Zhou[a]

[a]Department of Sociology, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104; and [b]Department of Sociology, Institute for Social Science Survey, Peking University, Beijing 100871, China

Contributed by Yu Xie, February 20, 2014 (sent for review December 31, 2013; reviewed by David Grusky and Andrew Walder)

Using multiple data sources, we establish that China's income inequality since 2005 has reached very high levels, with the Gini coefficient in the range of 0.53–0.55. Analyzing comparable survey data collected in 2010 in China and the United States, we examine social determinants that help explain China's high income inequality. Our results indicate that a substantial part of China's high income inequality is due to regional disparities and the rural-urban gap. The contributions of these two structural forces are particularly strong in China, but they play a negligible role in generating the overall income inequality in the United States, where individual-level and family-level income determinants, such as family structure and race/ethnicity, play a much larger role.

Since its beginning in 1978, China's economic reform has led not only to rapid economic growth but also to a large increase in economic inequality. Although scholars continue to debate about precise estimates (1), the consensus is that income inequality in China has now reached a level much higher than that in the United States (2). As we will discuss below, the Gini coefficient for family income in China has now reached a level above 0.5, compared with 0.45 in the United States in 2010. This finding is significant because China had a very low level of income inequality as recently as in the late 1980s (3). Ordinary persons in China know about this increase, as they have personally experienced it in their own lives (4). Although ordinary Chinese people seem to tolerate the high inequality (4–6), they also recognize it as a social problem needing to be addressed. In fact, out of a number of social issues given, respondents in a 2012 national survey rated economic inequality (more precisely, the "rich-poor gap") the most severe, above corruption and unemployment (7).

In this paper, we wish to address two research questions. (*i*) How high is income inequality in today's China? (*ii*) Why is it so high? The first question appears to be a simple fact that could be answered by government statistics. Unfortunately, this is not true for China. For a variety of complicated reasons, ranging from politics to practical difficulties, government statistics on Chinese well-being have been questioned for their accuracy (8). This concern is exacerbated by the long-standing concealment practices of China's National Bureau of Statistics (NBS), responsible for constructing and releasing government data on China, such that no original microlevel data are accessible to any independent researcher that could be used to corroborate the macrolevel statistics it releases. In the case of income inequality, the NBS stopped releasing the Gini coefficient after it reached 0.41 in 2000 (8). It was not until an economist claimed that the Gini coefficient had reached the shockingly high level of 0.61 that the NBS, in early 2013, released the Gini coefficients for recent years, which were slightly under 0.5 (2, 9). What is the true level of income inequality in today's China? Scholars continually and heatedly debate this question (1). In this paper, we contribute to this debate by computing estimates of the Gini coefficient using seven newly available nationally representative surveys, including microlevel data from a very large survey conducted by the NBS in 2005. The other six surveys were independently conducted by four university-affiliated survey organizations with a transparent sampling framework. We use data from these seven sources to estimate Gini coefficients for recent years and compare them both with those in China's past and those in other countries, particularly the United States.

The second objective of this paper is to explain why income inequality by international standards is so high in today's China. Past research has exclusively focused on the contributions of income components and determinants of temporal changes in China's overall income inequality (3, 10, 11). If we wish to understand why China's inequality is so high, however, we need to know what special features of the Chinese context might account for the inequality. With this goal in mind, we adopt a different research strategy and compare today's China not with its own past but with another large, important, and well-studied country with a high level of economic inequality—the United States. In the last 20 y, the United States has also experienced increasing economic inequality (12), but the increase in the United States has been much smaller than that in China. In this paper, we focus on two nationally representative surveys in 2010 and compare five determinants of income inequality between China and the United States. We test a hypothesis proposed by both Xie (13) and Wang (14) that China's current high income inequality is heavily driven by structural forces attributable to the Chinese political system, most notably the urban-rural gap and the regional variation in economic wellbeing.

Data

For our main analyses throughout this paper, we focus on the 2010 baseline survey of the China Family Panel Studies (hereafter CFPS 2010). The 2010 survey of the CFPS is the baseline wave of a large-scale, almost nationally representative panel survey project conducted by the Institute of Social Science Survey at Peking University (extensive information about the survey can be found at www.isss.edu.cn/cfps/). The 25 provinces of China (excluding Inner Mongolia, Xinjiang, Tibet, Hainan, Ningxia, and Qinghai) covered by the CFPS represent about 95% of the Chinese population in mainland China (15). Through a multistage probability sampling procedure, this survey completed interviews with

Significance

We document a rapid increase in income inequality in China's recent past, capitalizing on newly available survey data collected by several Chinese university survey organizations. By now, China's income inequality not only surpasses that of the United States by a large margin but also ranks among the highest in the world, especially in comparison with countries with comparable or higher standards of living. We argue that China's current high income inequality is significantly driven by structural factors attributable to the Chinese political system, the main structural determinants being the rural-urban divide and the regional variation in economic well-being.

Author contributions: Y.X. designed research; X.Z. performed research; X.Z. analyzed data; and Y.X. and X.Z. wrote the paper.

Reviewers: D.G., Stanford University; and A.W., Stanford University.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: yuxie@umich.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1403158111/-/DCSupplemental.

14,798 sampled households and all individuals living in these households. For each family, a list of questions were asked to solicit information on labor income, business income, investment income, transfer income, and income from other sources, from which we constructed total family income as our variable of interest. To cross-validate the level of income inequality in today's China, we also calculated Gini coefficients of family income from six additional data sources: the China 2005 1% Population Intercensus Survey (also called the 2005 mini-census, hereafter Mini-Census 2005), the 2010 and 2012 Chinese General Social Surveys (CGSS 2010 and CGSS 2012), the 2011 Chinese Household Finance Survey (CHFS 2011), the 2012 baseline wave of the China Labor Force Dynamic Survey (CLDS 2012), and the 2012 wave of the China Family Panel Studies (CFPS 2012). Detailed information about these surveys is provided in *SI Text*. In calculating the Gini coefficients, we excluded small proportions of families who did not report a positive income. Table 1 summarizes these data sets, with organization name, sample coverage, sample size, measure of family income, and estimated Gini coefficients. In particular, we also report Gini coefficients with purchasing power parity (PPP) adjustment that control for differences in cost of living between rural and urban areas and across provinces (see Table S1 for more details). For comparability with data from other countries, we use unadjusted Gini coefficients for the remainder of the paper.

Table 1. Seven recently collected, nationally representative survey data sets in China and the corresponding estimates of the Gini coefficient

Data source	Name of organization	Coverage of provinces	Sample size (no. families)	No. families with positive income	Measure of income	Gini coefficient (without PPP adjustment)	Gini coefficient (with PPP adjustment)
Mini-Census 2005	National Bureau of Statistics of China	All 31 provinces of Mainland China	973,159	779,849	Sum of each family member's self-reported monthly income multiplied by 12	0.483 (0.496)	
CGSS 2010	Renmin University of China and the Hong Kong University of Science and Technology	All 31 provinces of Mainland China	11,785	10,260	Self-reported total family income in 2009	0.545	0.518
CGSS 2012	Same as above	29 provinces of Mainland China (excluding Tibet and Hainan)	11,765	10,326	Self-reported total family income in 2011	0.539 (0.563)	0.515 (0.537)
CFPS 2010	Peking University	25 provinces of Mainland China (excluding Inner Mongolia, Xinjiang, Tibet, Hainan, Ningxia, Qinghai)	14,798	13,837	Total family income in 2009 from all sources	0.530 (0.541)	0.517 (0.525)
CFPS 2012	Same as above	Same as above	13,316	11,785	Total family income in 2011 from all sources	0.532 (0.526)	0.526 (0.517)
CHFS 2011	Southwestern University of Finance and Economics	25 provinces of Mainland China (excluding Inner Mongolia, Xinjiang, Tibet, Hainan, Ningxia, Fujian)	8,438	8,092	Total family income in 2010 from all sources	0.611 (0.633)	0.588 (0.609)
CLDS 2012	Sun Yat-sen University	28 provinces of Mainland China (excluding Tibet, Chongqing and Hainan)	10,612	9,735	Self-reported total family income in 2011	0.536	0.531

For CFPS 2010, CFPS 2012, and CHFS 2011, total family income is aggregated from a list of items for labor income, net business income, investment income, transfer income, and income from other sources, whereas for CGSS 2010, CGSS 2012, and CLDS 2012, family income is measured by one single question in which the respondent reported total family income in the previous year, which may come from all possible sources. Wherever information on family size is available, Gini coefficients for family income per capita were also calculated and are shown in parentheses. For PPP-adjusted Gini coefficients, incomes are adjusted for differences in cost of living between rural and urban areas and among provinces, which are from the geographic price indices in 2000 reported by Brandt and Holz (33) updated to 2010 using provincial rural and urban inflation rates published by the NBS (34).

How High Is Income Inequality in China Today?

The Gini coefficient is a well-understood measure of inequality (16). Based on the distribution of an outcome variable among ranked-ordered units in a population, the Gini coefficient measures the degree of inequality in the distribution of total resources. A Gini coefficient of 0 expresses perfect equality, with all units receiving an equal share. A Gini coefficient of 1 means maximal inequality where one unit has all of the resources. In Fig. 1, we show our estimates of the Gini coefficient of family income for recent years in China from the seven new sources discussed above, represented by red squares. We also display the Gini coefficient published by Li et al. (17) based on the 2007 survey of the Chinese Household Income Project (denoted by CHIP 2007), which was conducted by the NBS. For official statistics, we use purple triangles to show the Gini coefficients from 2003 to 2012 reported by an NBS official in a press conference in January 2013 (2). It is evident that the Gini coefficients based on independent, university-based surveys uniformly surpass the official figures, especially for the most recent years, i.e., 2010, 2011, and 2012, ranging from 0.530 based on CFPS 2010 to 0.611 based on CHFS 2011.

To provide background, we also present the Gini coefficients for earlier years in black circles, taken from the database of the World Institute for Development Economics Research of the United Nations University (hereafter UNU-WIDER) (18). To ease interpretation, we draw a nonparametric local polynomial regression (LOESS) curve to depict the trends in income inequality in China during the last four decades. For comparison, we also present a LOESS regression curve for the United States in the same period, based on data released by the US Census Bureau (19). Fig. 1 clearly shows that, since the 1980s, the rise of income inequality has been far more dramatic in China than in the United States. According to the smoothed trends, the Gini coefficient in China was around 0.30 in 1980, but by 2012 it had nearly doubled to 0.55, far surpassing the level of 0.45 in the United States.

The rise of income inequality in China has coincided with China's rapid economic growth since the beginning of the economic reform in 1978. These parallel increasing trends in China's recent past have caused a large portion of ordinary Chinese to think that an increase in income inequality automatically accompanies economic development and thus is a necessary price for economic growth (4, 13). In both economics and sociology literature, it has also been proposed that the relationship between economic development and income inequality takes the form of an inverted U: inequality first increases in the early stages of development and then declines in the later stages (20–23).

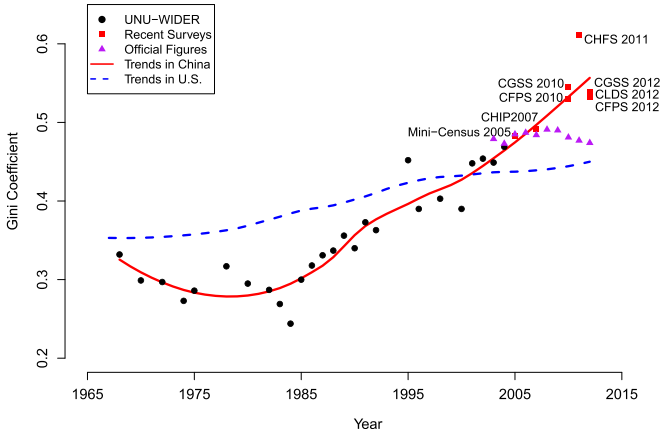

Fig. 1. Trends in Gini coefficient of family income in China and the United States. Gini coefficients from UNU-WIDER and CHIP 2007 are based on household disposable income per capita. Gini coefficients of family income in the United States from 1967 to 2012 are provided by the US Census Bureau (19). Trends in both countries were fitted using LOESS. The official figures in China were not included in fitting the LOESS curve.

When we examine the empirical pattern across countries between economic development and income inequality using the latest available data, we indeed find the expected inverted U relationship. In Fig. 2, we plot the Gini coefficients for family income in 136 countries between 1995 and 2012 (y axis) against the logarithm of gross domestic product (GDP) per capita (x axis) in these countries in 2012. The data came from the World Factbook released by the Central Intelligence Agency (24). Each green circle in the figure represents a country. The inverted U relationship is shown by a quadratic regression line (dashed line) that was fitted to the cross-national data. We observe that income inequality in the United States is significantly above the level predicted by the fitted curve.

Has China's recent past followed the upward segment of the general inverted U relationship between economic development and income inequality? To answer this question, we superimpose in Fig. 2 the trend data for China from 1980 to 2012, with a LOESS regression line depicting the relationship between log GDP per capita and Gini coefficient for China over this period. At the same level of economic development (as measured by logged GDP per capita), Fig. 2 shows that China's increase in Gini has been much faster than what is expected from the regression line estimated from the cross-national data. In fact, China's Gini continued to grow after it crossed the regression line in 2002, when China's GDP per capita reached US$2,866, and its Gini reached 0.45. As in Fig. 1, the Gini coefficients for the years before 2005, appearing in black circles, were taken from the UNU-WIDER database. As in Fig. 1, we use red squares to present the Gini coefficients estimated from the eight new sources, including Mini Census 2005 and CHIP 2007. Fig. 2 indicates that the temporal relationship between income inequality and economic development in China in the last three decades has clearly followed a different pattern than what is estimated from the cross-national data, with the former having experienced much faster increases than international comparison would suggest. In the most recent years, although there are still other countries with higher Gini coefficients than China at the same level of economic development, China's income inequality has far exceeded the average for other countries at similar levels.

Using multiple data sources, we established that China's income inequality since 2005 has reached very high levels, seen from the perspectives of both China's past (Fig. 1) and international comparisons (Fig. 2). The next question is what may account for China's very high levels of income inequality in the post-2005 period. We now turn to regression analyses of microlevel data to understand why income inequality in China has been so high in recent years.

Why Is Income Inequality So High in China?

To aid interpretation, we compare the microlevel structure of income inequality in China with that in the United States, which serves as a reference in this study because its inequality structure has been well studied (12). We pair CFPS 2010 with the March supplement of the Current Population Survey in 2010 (March CPS 2010). These two data sets are fairly comparable because both surveys contain comprehensive measurements of income for each interviewed family, including labor income, net business income, investment income, transfer income, and income from other sources. Total family income is defined as the sum of these items. For each country, we examine the extent to which income inequality is mediated by five explanatory variables: (*i*) region, (*ii*) area type, (*iii*) education, (*iv*) race/ethnicity, and (*v*) family structure. We made the measurements of the variables as comparable across the countries as possible. Region is defined as

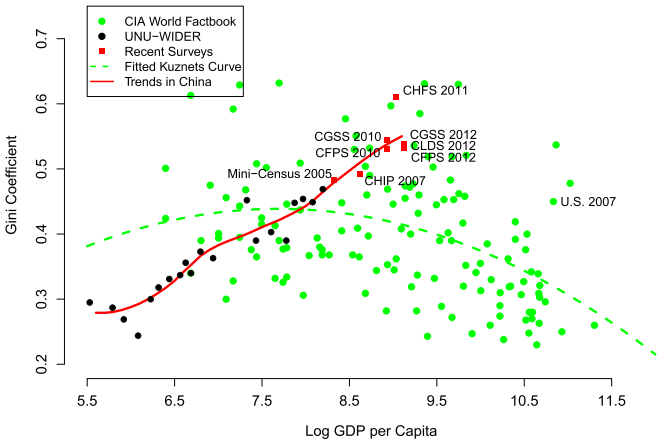

Fig. 2. GDP per capita and Gini coefficients of family income in 136 countries and their trends in China. GDP per capita are measured by purchasing power parity (current international dollar). Historical data on GDP per capita in China are provided by the China Statistical Yearbook 2012 (34). GDP per capita and Gini coefficients for different countries are provided by the CIA World Factbook (24). The relationship between log GDP per capita and Gini coefficient in China was fitted using LOESS. The Kuznets curve was fitted as a quadratic function:

$$\text{Gini} = -0.271 + 0.185 \times \log(\text{GDP per capita}) - 0.012 \times [\log(\text{GDP per capita})]^2.$$

province in China and state in the United States. Area type corresponds to a dichotomous variable for rural and urban in China and a trichotomous variable for metropolitan, non-metropolitan, and not identified in the United States. Both education (with a nearly identical 6-category classification) and race/ethnicity (23 categories for China and 38 categories for the United States) are measured for the family head. For both countries, we adopt a five-category classification to characterize the family structure: (*i*) primary-individual family, (*ii*) single-parent family, (*iii*) married couple with no children, (*iv*) married couple with child(ren), and (*v*) extended family. Detailed descriptions of these explanatory variables are provided in Table S2. To remove the confounding effect of family size, we use family income per capita—total family income divided by family size—as our dependent variable and weight all of the following analyses by family size. Altogether, 14,798 families were interviewed in CFPS 2010, and 88,957 were interviewed in March CPS 2010. After excluding those families with nonpositive income or missing data, we have 12,523 observations from CFPS 2010 and 85,564 observations from March CPS 2010.

To assess the relative importance of different factors in explaining income inequality, we first construct a simple linear regression model, with logged income as the dependent variable and each of the five explanatory variables discussed above as the independent variable

$$\log(Y_i) = X_i\beta + \varepsilon_i, \qquad [1]$$

where Y_i denotes family income per capita for the ith family, X_i denotes the row vector for an explanatory variable with β as the coefficient vector, and ε_i is the unexplained residual. With this model, we essentially measure income inequality with the variance of logged income. When income follows a log-normal distribution, this measure has a strictly monotonic relationship with the Gini index (16, 25). The variance measure is particularly useful for our purpose because it can be easily decomposed into between-group and within-group components. In fact, because all of the explanatory variables considered here are categorical, the above regression is equivalent to one-way ANOVA, and R^2 gauges the degree to which the overall inequality can be explained by a given factor. We estimate parallel regressions for China and the United States and compare the corresponding R^2s. Because we include only one factor at a time as the independent variable in predicting income, we denote these R^2s as bivariate R^2s.

We present the bivariate R^2 results in Fig. 3, using solid squares for China and hollow squares for the United States. Relative to the United States, income inequality in China is far more explainable by regional disparities and the rural-urban divide (i.e., area type). Specifically, about 12% of the overall income inequality in China can be attributed to differences across provinces, whereas variation across states in the United States accounts for no more than 2% of the overall inequality. Similarly, the rural-urban divide accounts for more than 10% of the total inequality in China, whereas it contributes virtually none in the United States. These results confirm a conclusion drawn in past studies that a large portion of China's income inequality is attributable to structural forces such as the urban-rural divide and regional variation, both of which have been maintained by political structures in large part to promote economic growth (13, 14, 25, 26). Fig. 3 also shows a notable difference between China and the United States in how income inequality is affected by family structure and race/ethnicity of the family head. Family structure is far more important in the United States, accounting for more than 12% of the total inequality, compared with only 2% in China. This contrast is partly due to a much larger proportion of single-parent families in the United States (7.9% of the US population vs. 1.9% of the Chinese population), which are subject to a significant disadvantage in average income in the United States (see Tables S3 and S4 for detailed statistics). In addition, the role of education is similarly important in China and the United States: in both countries, around 15% of income inequality can be explained by the level of education of the family head.

Robustness Analysis

One drawback of the preceding one-way ANOVA is that it only considered one factor at a time. We know that different determinants of income are correlated and thus share explanatory power in common. For example, in China, education tends to be higher in urban areas than in rural areas and higher in coastal, more developed regions than in inland, less developed regions. Hence, the bivariate R^2s from different explanatory variables may be confounded. To corroborate our finding, we resort to multivariate linear regressions. First, we regress logged income on the five explanatory variables together (see Tables S3 and S4

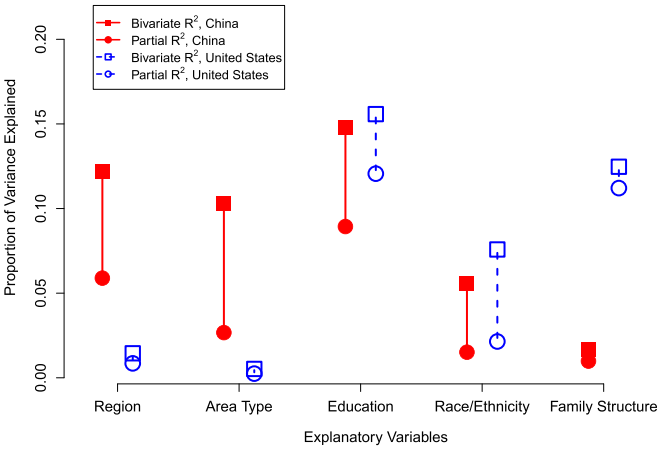

Fig. 3. Bivariate and partial R^2 for different predictors of income in China and the United States. The dependent variable is log family income per capita. All regressions were weighted by family size.

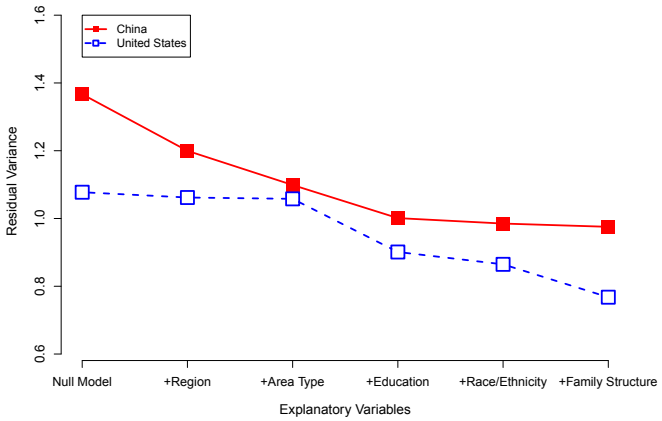

Fig. 4. Residual variances for different models of logged income in China and the United States. The dependent variable is log family income per capita. All regressions were weighted by family size.

for detailed results). Then, we alternately exclude one of the five independent variables from the full model and reestimate the restricted regression, yielding five auxiliary models. To assess the unique contribution of variable K, we calculate the partial R^2 (27) as

$$\text{partial } R^2 = \frac{R^2 - R^2_{-K}}{1 - R^2_{-K}}, \qquad [2]$$

where R^2 denotes the R^2 for the full model, and R^2_{-K} denotes the R^2 for the model that includes all variables except K. Hence, partial R^2 gauges the proportion of the remaining variation of logged income that can be explained by variable K when all of the other factors are taken into account. Fig. 3 displays these partial R^2s with solid and hollow circles, respectively, for China and the United States. We observe that most partial R^2s are markedly smaller than the corresponding bivariate R^2s, indicating correlated explanatory power among different determinants of income. However, the main contrasts between China and the United States remain true: (*i*) compared with the United States, inequality in China is heavily structured by regional variation and the rural-urban divide; (*ii*) compared with China, inequality in the United States is more influenced by family structure and race/ethnicity of the family head; and (*iii*) education of the family head is very important in predicting family income in both China and the United States.

To further understand how the five proposed social determinants contribute to income inequality differently between China and the United States, we now examine the levels of residual variance for a set of nested regression models as we progressively add region, area type, education, race/ethnicity, and family structure to the regression model of logged income. The results are shown in Fig. 4, again with solid squares representing China and hollow squares representing the United States. Consistent with results using the Gini measure, the overall variance of log family income per capita (i.e., null model) is much higher in China than in the United States. Second, confirming an earlier finding from Fig. 3, region and area type account for a substantial part of income variation in China but not in the United States. As a result, when only these two factors are taken into account, residual inequality is very similar between China and the United States. However, when education is included in the regression model, the size of residual inequality again becomes lower in the United States. Finally, when race/ethnicity and family structure are also factored in, residual inequality is markedly higher in China than in the United States.

Conclusion

The preceding analyses of recently available survey data from China have led us to draw two conclusions. First, China's income inequality has reached very high levels in recent years, with the Gini coefficient well above 0.50 around 2010: high both from the perspective of China's past and in comparison with other countries at similar stages of economic development. They are also substantially higher than what has been acknowledged in the government statistics. Second, a substantial part of China's high income inequality is due to two structural forces at work: a large regional variation and a large gap between rural and urban residents. The contributions of these two structural forces are particularly strong in China, in light of the fact that they play a negligible role in generating the overall income inequality in the United States, where personal-level and family-level income determinants, especially family structure and race/ethnicity, tend to be more important. This pattern, however, should not be taken as unique in today's China from either a temporal or a comparative perspective. From a temporal perspective, in the prereform era, China's income distribution was already marked by a large urban-rural divide and vast regional variation. In 1980, the Gini coefficient in urban China was as low as 0.16, but the nationwide Gini coefficient was around 0.3, higher than other socialist economies at that time (28, 29). From a comparative perspective, other large developing or middle-income countries, such as Brazil, India, and Indonesia, also exhibit similar patterns of large regional income disparities (30). For example, about 14–15% of the total inequality in Brazil is attributable to regional differences (31).

In conclusion, our results reveal that China's income inequality has grown rapidly in the last three decades, to a very high level around 2010. The rapid rise in income inequality can be partly attributed to long-standing government development policies that effectively favor urban residents over rural residents and favor coastal, more developed regions over inland, less developed regions (13, 25, 32). Given the structural importance in generating China's high income inequality, it seems plausible that income inequality can be reduced by changes in government policies, for example, reducing the rural-urban and regional disparities. We begin to observe early signs of change in this direction between 2010 and 2012 (2).

ACKNOWLEDGMENTS. Financial support for this research was provided by Natural Science Foundation of China Grant 71373012, Peking University, and the Population Studies Center at the University of Michigan, which receives core support from National Institute of Child Health and Human Development Grant R24HD041028.

1. Yue X, Li S (2013) Review of the debate on the Gini coefficient. Available at www.ciidbnu.org/news/201307/20130708161954706.html. Accessed September 1, 2013.
2. Xie Y, Zhang X, Xu Q, Zhang C (2013) *Wellbeing Development Report of China*, eds Xie Y, Zhang X, Li J, Yu X, Ren Q (Peking Univ Press, Beijing), pp 27–53 (in Chinese).
3. Gao Q, Riskin C (2009) *Creating Wealth and Poverty in Postsocialist China* (Stanford University Press, Palo Alto, CA), pp 20–36.
4. Xie Y, Thornton A, Wang G, Lai Q (2012) Societal projection: Beliefs concerning the relationship between development and inequality in China. *Soc Sci Res* 41(5): 1069–1084.
5. Whyte M (2010) *Myth of the Social Volcano: Perceptions of Inequality and Distributive Injustice in Contemporary China* (Stanford Univ Press, Redwood City, CA).
6. Wu X (2009) Income inequality and distributive justice: A comparative analysis of mainland China and Hong Kong. *China Q* 200:1033–1052.
7. Wu Q, Xie Y (2013) *Wellbeing Development Report of China*, eds Xie Y, Zhang X, Li J, Yu X, Ren Q (Peking Univ Press, Beijing), pp 344–367 (in Chinese).
8. Hvistendahl M (2013) The numbers game. *Science* 340(80):1037–1039.
9. Gan L, et al. (2014) *Data You Need to Know About China* (Springer-Verlag, Berlin).
10. Hauser SM, Xie Y (2005) Temporal and regional variation in earnings inequality: Urban China in transition between 1988 and 1995. *Soc Sci Res* 34(1):44–79.
11. Khan AR, Riskin C (1998) Income and inequality in China: Composition, distribution and growth of household income, 1988 to 1995. *China Q* 154:221–253.

12. McCall L, Percheski C (2010) Income inequality: New trends and research directions. *Annu Rev Sociol* 36:329–347.
13. Xie Y (2010) Understanding inequality in China. *Soc* 30(3):1–20.
14. Wang F (2008) *Boundaries and Categories: Rising Inequality in Post-Socialist Urban China* (Stanford Univ Press, Redwood City, CA).
15. Xie Y (2012) User's manual of China Family Panel Studies Data 2010. Available at www.isss.edu.cn/cfps/d/file/wd/jsbg/2010jsbg/a16476476881bc46c90086fa1508880b.pdf. Accessed September 1, 2013.
16. Allison PD (1978) Measures of inequality. *Am Sociol Rev* 43(6):865–880.
17. Li S, Sato H, Sicular T (2013) *Rising Inequality in China* (Cambridge Univ Press, Cambridge, UK).
18. UNU–WIDER (2008) World Income Inequality Database, version 2.0c. Available at www.wider.unu.edu/research/Database/en_GB/wiid/. Accessed October 5, 2013.
19. U.S. Census Bureau (2013) Historical income tables: Income inequality. Available at www.census.gov/hhes/www/income/data/historical/families/2012/F04_2012.xls. Accessed October 5, 2013.
20. Barro RJ (2000) Inequality and growth in a panel of countries. *J Econ Growth* 5(1):5–32.
21. Chiswick BR (1971) Earnings inequality and economic development. *Q J Econ* 85(1):21–39.
22. Kuznets S (1955) Economic growth and income inequality. *Am Econ Rev* 45(1):1–28.
23. Lenski GE (1966) *Power and Privilege: A Theory of Social Stratification* (Univ of North Carolina Press, Chapel Hill, NC).
24. Central Intelligence Agency (2013) The World Factbook: Distribution of family income: Gini index. Available at www.cia.gov/library/publications/the-world-factbook/fields/2172.html. Accessed October 5, 2013.
25. Xie Y, Hannum E (1996) Regional variation in earnings inequality in reform-era urban China. *Am J Sociol* 101(4):950–992.
26. Sicular T, Yue X, Gustafsson B, Li S (2007) The urban–rural income gap and inequality in China. *Rev Income Wealth* 53(1):93–126.
27. Meng X, Shen K, Xue S (2013) Economic reform, education expansion, and earnings inequality for urban males in China, 1988–2009. *J Comp Econ* 41(1):227–244.
28. World Bank (1983) *China: Socialist Economic Development* (The World Bank, Washington, DC), Vol I.
29. Kornai J (1992) *The Socialist System* (Princeton Univ Press, Princeton).
30. Milanovic B (2005) Half a world: Regional inequality in five great federations. *J Asia Pac Econ* 10(4):408–445.
31. Elbers C, Lanjouw JO (2004) *Inequality and Economic Development in Brazil* (World Bank), pp 81–104. Available at www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2004/10/05/000012009_20041005095126/Rendered/PDF/301140PAPER0Inequality0Brazil.pdf. Accessed February 1, 2014.
32. Whyte M (2014) Soaring income gaps: China in comparative perspective. *Daedalus* 143(2):39–52.
33. Brandt L, Holz CA (2006) Spatial price differences in China: Estimates and implications. *Econ Dev Cult Change* 55(1):43–86.
34. National Bureau of Statistics of China (2012) *China Statistical Yearbook 2012* (China Statistics Press, Beijing).

Supporting Information

Xie and Zhou 10.1073/pnas.1403158111

SI TEXT

Data and the Calculation of Gini Coefficients. To measure the size of family-level income inequality in today's China, we calculated estimates of the Gini coefficient using data from seven recent nationally representative surveys. Besides CFPS 2010, which we introduced in detail in the main text, the other data sources are (*i*) the China 2005 1% Population Intercensus Survey (also called the 2005 minicensus, hereafter Mini-Census 2005), (*ii*) the 2010 and 2012 Chinese General Social Surveys (CGSS 2010 and CGSS 2012), (*iii*) the 2011 Chinese Household Finance Survey (CHFS 2011), (*iv*), the 2012 baseline wave of the China Labor Force Dynamic Survey (CLDS 2012), and (*v*) the 2012 wave of the China Family Panel Studies (CFPS 2012). Table 1 summarizes these data sources, with organization name, sample coverage, sample size, measure of family income, and estimated Gini coefficients. Note that all Gini coefficients were calculated for families with a positive annual income.

Mini-Census 2005, as part of the China census program, is an intercensus survey administered by the National Bureau of Statistics in China, which covers 973,159 families (about 2.6 million people) in all 31 provinces and equivalent administrative units of mainland China. It was conducted through a multistage stratified cluster sampling process with a household-based questionnaire collecting each household member's information pertaining to employment, occupation, and income, as well as basic demographic characteristics. For this data set, we define total family income as the sum of all family members' self-reported monthly incomes multiplied by 12. Among the 779,849 families who had a positive annual income, the Gini coefficient is 0.483. For the same sample, we also calculated the Gini coefficient for family income per capita, i.e., total family income divided by family size. In this calculation, we weighted the data by family size such that individuals, rather than families, were treated as the unit of analysis. The resulting estimate is 0.496, slightly higher than the family-level Gini coefficient. This difference may result from a negative correlation between family income per capita and family size in China. Because very poor families (in terms of family income per capita) are more likely to be larger families, relatively more weight is placed at the very low end of the income distribution in the latter calculation, thereby amplifying the measured level of inequality.

The CGSS is a nationwide, repeated, cross-sectional general survey project in China. The project was launched jointly by Renmin University of China and the Hong Kong University of Science and Technology in 2003 (1). We use two recent waves of the study, CGSS 2010 and CGSS 2012, to corroborate our estimates of the Gini coefficient measuring income inequality. Both CGSS 2010 and CGSS 2012 adopted a nationally representative sampling frame. Although the sample of CGSS 2010 covered all 31 provinces and equivalent administrative units of mainland China, the sample of CGSS 2012 did not include Tibet and Hainan. After excluding those families with missing or nonpositive income, the sample size is 10,260 for CGSS 2010 and 10,326 for CGSS 2012. We obtained Gini coefficients from these two samples at 0.545 and 0.539, respectively. Because CGSS 2012 also collected information on family size, we calculated the Gini coefficient for family income per capita. As in the case of Mini-Census 2005, the Gini coefficient at the individual level (0.563) is also higher than that at the family level (0.539).

The CFPS is a large-scale panel survey project conducted by the Institute of Social Science Survey at Peking University. The project was designed to study the long-term dynamics of social transition in China. With the household as the target of sampling, the survey collects comprehensive information on the household as a whole and all individual household members living in the sampled households. The CFPS adopts a nearly nationally representative sampling frame, which includes 25 provinces (excluding Inner Mongolia, Xinjiang, Tibet, Hainan, Ningxia, Qinghai, Hong Kong, Macau, and Taiwan) of mainland China, representing 95% of the Chinese population (2). We use data from the CFPS baseline survey, which was carried out in 2010 (CFPS 2010), and a follow-up survey in 2012 (CFPS 2012). Both CFPS 2010 and CFPS 2012 collected detailed information about households' incomes and expenditures in the previous year. We summed up the household incomes from wages and salaries, agricultural production, investment, and transfers. An earlier analysis compared income measures in CGSS 2010 and CFPS 2010 and found similar distributions of family income and expenditure from the two data sources (3). The baseline survey (CFPS 2010) interviewed 14,798 families, of which 13,837 reported a positive family income. Among these families, we calculated the Gini coefficient to be 0.530 for family income and 0.541 for family income per capita. The follow-up survey in 2012 reinterviewed about 90% of the original sample, including 13,316 families in total and 11,785 families with a positive annual income. From CFPS 2012 data, we obtained the Gini coefficient at 0.532 at the family level and 0.526 at the individual level.

The sixth data set analyzed for this study is the 2011 baseline of the China Household Finance Survey (hereafter CHFS) (website of the project is at www.chfsdata.org/). The CHFS is a survey project aimed at understanding household finance in China. The project is directed by the Survey and Research Center for the China Household Finance Survey at the Southwestern University of Finance and Economics. The baseline survey was carried out in 2011. It collected information about household assets, income, expenditures, and social and commercial insurance in 2010. The baseline survey interviewed 8,438 households, sampled in 80 districts/counties in 25 provinces. We included this data set in our study not only because of its unusual detail in collecting household financial information but also because of the amount of publicity it received. The shockingly high level of the Gini coefficient based on this survey (i.e., the 0.61 estimate) aroused both media attention and an academic debate as soon as it was released (4, 5). The survey and the Gini coefficient based on this survey were prominently featured in a *Science* article on issues concerning data accuracy in China (6). The survey also aroused suspicion of the Chinese government's official statistics, which are believed to either conceal or underestimate the worsening income inequality in China. As shown in Table 1, among the 8,092 families with a positive annual income in the CHFS, the Gini coefficient is 0.611 at the family level and 0.633 at the individual level.

Finally, we use data from the baseline wave of the China Labor Force Dynamic Survey (i.e., CLDS) to corroborate our findings based on other data sets. Launched by the Department of Sociology at Sun Yat-sen University, this survey project was designed especially to monitor labor market dynamics among Chinese adults. The baseline survey (CLDS 2012) interviewed 10,612 households in 28 provinces of mainland China (excluding Tibet, Chongqing, and Hainan). As in CGSS 2012, a single question was asked concerning the total amount of family income in 2011. Among the 9,735 families who reported a positive income, the Gini coefficient

is 0.536, which closely resembles the estimates from CGSS and CFPS.

For all of the data sets except Mini-Census 2005, we also calculated Gini coefficients with purchasing power parity (PPP) adjustment that control for differences in cost of living between rural and urban areas and among provinces. To construct regional price differences, we updated the spatial price deflators in 2000 reported by Brandt and Holz (7) to the year of 2010 using provincial rural and urban consumer price indices published by the National Bureau of Statistics of China (NBS) (8). As expected, the PPP adjustment slightly reduces the measured level of inequality for all of the data sets.*

Table S1 shows income and price differences by province and rural/urban status. All entries are ratios relative to urban families in Beijing. Raw income data are given in the first two columns. Price data are given in the next two columns. Income data adjusted for price variations are given in the last two columns. To compare variability in the three series across the province-rural/urban combinations, we also present the Gini coefficients for the series in the last row of Table S1. The results show that the variation in price by province and rural/urban status (the second series) is much less than that in income, either raw or adjusted (the first or the last series).

Regression Analyses. To compare the structure of income inequality between China and the United States, we use data for regression analyses from CFPS 2010 and the March supplement of the Current Population Survey in 2010 (March CPS 2010). The Current Population Survey (CPS), sponsored jointly by the US Census Bureau and the US Bureau of Labor Statistics, is a monthly survey that provides up-to-date labor force statistics for the US population. Although the CPS is used primarily to provide information on employment and unemployment, it also collects demographic information, including age, sex, race, marital status, educational attainment, and family structure. The March supplement of the CPS, also known as the Annual Social and Economic Supplement, includes supplemental questions on work experience, income, noncash benefits, and migration. A respondent's total income is defined as the sum of his/her money wages and salaries, net income from self-employment, and income other than earnings received in the previous calendar year. The total family income is defined as the sum of the amounts received by all income recipients in the family. The 2010 data for March CPS contain 88,957 family records. In this study, we restricted the sample to those families who had a positive total income in 2009, yielding 85,564 family records.

In CFPS 2010, the total income of a family is defined in the same way as in March CPS 2010. As introduced in the previous section, among the 14,798 families who were interviewed, 13,837 had positive income. To conduct regression analyses, we also excluded those observations with missing data on any of the covariates (i.e., region, area type, education, race/ethnicity, and family structure), a procedure that resulted in 12,523 family records. In fact, most of the sample reduction in this step is due to failure in matching family records with individual records of family heads. This failure occurred because many family heads were not present at the time of the interview. There are no missing values on the covariates in March CPS 2010.

Table S2 provides a detailed description of the covariates that are used in our regression analyses, as well as the associated numbers of degrees of freedom. Specifically, region corresponds to the 25 province-level administrative units in China covered by CFPS 2010 and 51 state-level administrative units in the United States. This variable is used to examine the role of regional variation in generating the overall inequality in China vs. that in the United States. Area type is defined as a dichotomous variable for rural and urban in China and a trichotomous variable for metropolitan, non-metropolitan, and not identified in the United States. Education denotes the educational attainment of the family head. We use the same six-category classification for both China and the United States except that the fifth category corresponds to vocational (3-y) college in China and some college in the United States. Race/ethnicity is also defined as that of the family head. Although 55 ethnic minority groups are recognized in China, only 22 of them are represented in CFPS 2010. Thus, we have 23 ethnic categories in our analytic sample. The CPS questionnaire uses a 21-category classification of race and a separate question asking if the respondent is Spanish, Hispanic, or Latino. To be conservative, we cross-tabulated these two variables and treated each nonempty cross-classification as a separate group, a procedure that yielded 38 categories of race/ethnicity for the US sample. Finally, for both countries, we adopt the same five-category classification to characterize the family structure: (*i*) primary-individual family, (*ii*) single-parent family, (*iii*) married couple with no children, (*iv*) married couple with child(ren), and (*v*) extended family.

In columns 1–3 of Tables S3 and S4, we present the sample percentages and average family income per capita by different population subgroups in China and the United States. Sample percentages are weighted by family size such that the numbers reflect the proportion of the Chinese/US population in different groups. It is notable that the population shares of different types of families differ significantly between China and the United States; whereas more than 80% of the Chinese population live in nuclear families [i.e., married couple with child (ren)] or extended families, family structure is far more diverse in the United States. In particular, nearly 18% of the US population lives in primary-individual families, whereas the corresponding figure is around 2% in China. Average family per capita is defined as the geometric mean of family income per capita weighted by family size. We chose to report the geometric means because they are less sensitive than arithmetic means to extreme values of income. It is apparent that the average income varies greatly across provinces in China but not across states in the United States. For instance, the average incomes in Beijing and Shanghai are several times higher than those in Sichuan and Guizhou. Similarly, we can see a huge urban-rural gap in China (3,674 vs. 7,805 Yuan) but not in the United States. On the other hand, different types of families vary greatly in economic status in the United States but not much in China. For example, average income in families consisting of a married couple with no children is more than four times as large as in single-parent families in the United States, whereas the corresponding ratio in China is 1.79 (6,748/3,771 = 1.79). These descriptive statistics echo our findings reported in Fig. 3.

In the last two columns of Tables S3 and S4, we present the estimated coefficients in simple and multiple regressions, which are denoted by $\beta_{\text{bivariate}}$ and β_{full}, respectively. We also report the *F*-test results corresponding to each set of predictors in both types of regressions. These *F*-test statistics measure the statistical significance of these five sets of predictors in explaining income inequality. They also bear a one-to-one correspondence with the bivariate and partial R^2 that are reported in Fig. 3. We observe that all of the *P* values are <0.001, indicating that all bivariate and partial R^2 plotted in Fig. 3 are significantly different from zero.

*In this study, the definition of family income does not include agricultural products that are self-consumed in rural areas. Considering that rural families may keep a substantial fraction of their agricultural products for self-consumption, ignoring this part of income may overestimate the rural-urban income gap and nationwide inequality. For CFPS 2010 and 2012, we also calculated Gini coefficients that take into account the imputed value of this type of income. The resulting estimates are 0.512 for CFPS 2010 and 0.498 for CFPS 2012, lower than those reported in Table 1.

1. Bian Y, Li L (2012) The Chinese General Social Survey (2003-8): Sample designs and data evaluation. *Chin Sociol Rev* 45(1):70–97.
2. Xie Y (2012) *User's Manual of China Family Panel Studies Data 2010.* Available at www.isss.edu.cn/cfps/d/file/wd/jsbg/2010jsbg/a16476476881bc46c90086fa1508880b.pdf. Accessed September 1, 2013.
3. Xu Q, Zhang C, Zhou X, Xie Y (2012) *Comparison of Income Distribution: CFPS, CGSS, CHIP, and CHFS*. Available at www.isss.edu.cn/cfps/d/file/wd/jsbg/2010jsbg/a29e71f54de2f5dcd71c6284e4eed017.pdf. Accessed September 1, 2013.
4. Anonymous (2012) To each, not according to his needs: A new survey illuminates the extent of Chinese income inequality. Available at www.economist.com/news/finance-and-economics/21568423-new-survey-illuminates-extent-chinese-income-inequality-each-not. Accessed September 1, 2013.
5. Kao E (2012) China wealth gap continues to widen, survey finds. *South China Morning Post.* Available at www.scmp.com/news/china/article/1102047/china-wealth-gap-continues-widen-survey-finds. Accessed September 1, 2013.
6. Hvistendahl M (2013) The numbers game. *Science* 340(80):1037–1039.
7. Brandt L, Holz CA (2006) Spatial price differences in China: Estimates and implications. *Econ Dev Cult Change* 55(1):43–86.
8. National Bureau of Statistics of China (2011) *China Statistical Yearbook 2011* (China Statistics Press, Beijing).

Other Supporting Information Files

Table S1 (DOCX)
Table S2 (DOCX)
Table S3 (DOCX)
Table S4 (DOCX)