Empirical Economics (2024) 66:539–585
https://doi.org/10.1007/s00181-023-02472-0

Inequality and growth in China

Haiyan Lin[1] · **Markus Brueckner**[1]

Received: 7 June 2022 / Accepted: 1 July 2023 / Published online: 3 August 2023

Abstract
We provide estimates of the effects that income inequality has on economic growth in China. Our empirical analysis is at the county level. Using data provided by the China Health and Nutrition Survey, we construct measures of inequality and the growth rates of household incomes per capita for 72 Chinese counties during the period 1989–2015. System-GMM estimates of panel models show that the within-county effect of inequality on economic growth is significantly decreasing in initial average income. For the relatively low levels of initial average incomes that were prevalent in China during the 1980s and 1990s, our model estimates imply that the increase in inequality that occurred in China during the 1980s and 1990s had a significant positive effect on economic growth. However, for current levels of average income, our panel model predicts that inequality has a negative effect on economic growth: a 1 percentage point increase in the Gini would reduce the per annum growth rate by around 1 percentage point.

Keywords Inequality · Growth · Instrument strength of sys-GMM · China

JEL Classification O40 · O11 · O53 · D30

Deng Xiaoping (1985): “Let some people get rich first."
Xi Jinping (2015): “We must ensure that the fruits of development benefit all people."

✉ Haiyan Lin
haiyan.lin@anu.edu.au

Markus Brueckner
markus.brueckner@anu.edu.au

1 Research School of Economics, Australian National University, Canberra, Australia

1 Introduction

In this paper, we seek to gain an understanding of how, specifically, in China inequality affects economic growth. The economic theory proposed by Galor and Zeira (1993) predicts that inequality affects growth differently depending on initial wealth. In the presence of fixed costs of human capital investment and imperfect credit markets, the theory predicts that inequality reduces growth in a country with high average income while the opposite is the case for a poor country. Based on this seminal theoretical paper, we would expect that the effects of inequality on growth in China differ across time and space. At the early stage, when Chinese counties were poor, a more unequal income distribution might have enhanced growth. However, average incomes have grown substantially in China over the past four decades: for current levels of average income, it might be the case that inequality has a negative effect on economic growth as suggested by the theory.

Estimates of an econometric model are required to gain a precise understanding of the quantitative effects that inequality has on economic growth in China. The main contribution of this paper is to provide such estimates. For a panel of 72 Chinese counties during the period 1989–2015, we estimate dynamic panel models that account for county and time fixed effects. We estimate the dynamic models using system-GMM (sys-GMM). Motivated by the economic theory, discussed above, we include in our econometric models not only measures of inequality but also the interaction between inequality and initial average income. Our county-level data set enables a much more rigorous econometric analysis of the effects that inequality has on economic growth in China, relative to using a country-level data set where the number of observations, for China, would be quite small.

Our panel model estimates show that the effect of inequality on economic growth is significantly decreasing in initial average income. There is a threshold of average income below which inequality has a positive effect on economic growth. Above that threshold inequality has a negative effect on economic growth. For our panel of Chinese counties, the estimated threshold of initial average income is about 6000 yuan in 2015 prices.

In the late 1980s, all counties in our panel data set had a household income per capita below 6000 yuan. Thus, for the late 1980s, our model estimates suggest that inequality in China had a positive effect on economic growth. In contrast, by 2015, all counties in our panel data set had an average income above 6000 yuan. Hence, for 2015, our model estimates suggest that inequality reduced growth in China. For the average county in the year 2015 (that has a household income per capita of about 22,000 yuan) our model estimates show that a 1 percentage point increase in the Gini reduces the per annum growth rate of household income per capita by around 0.5 percentage points. For a county at the top 10$^{\text{th}}$ percentile (that in the year 2015 has a household income per capita of about 36,000 yuan), our model estimates show that a 1 percentage point increase in the Gini reduces the growth rate of household income per capita by around 0.7 percentage points. For current, i.e. year 2021, levels of average income in China our panel model estimates predict that a 1 percentage point increase in the Gini would reduce the growth rate by around 1 percentage point. (The result is symmetric: i.e. for current levels of average income, a decrease in the Gini by 1 percentage point would increase the per annum growth rate by around 1 percentage point.)

We further document that the impacts of inequality on human capital are decreasing in initial income levels, aligning with the channel modelled in Galor and Zeira (1993). In the relatively poor counties, the higher initial income distribution has a positive effect on the share of individuals with high school and above degree, as well as on the share of skilled labour while in the relatively rich counties, the reduction in inequality has a beneficial effect on the educational attainment rate and the rate of skilled workers.

Our paper makes several important contributions to the literature. The first and foremost contribution is that we provide a rigorous econometric analysis of the effect that inequality has on economic growth specifically for China. This specific focus on China is warranted. In PPP terms, China is the world's largest economy. For a macroeconomist who is interested in the relationship between economic growth and income inequality, a China-specific study is very much of interest. Although there is a vast literature exploring the impact of inequality on economic growth (e.g. Forbes 2000; Easterly 2007; Brueckner and Lederman 2018), the literature focusing on China is very limited. Compared with two influential studies by Ravallion and Chen (2007) and Benjamin et al. (2011), our paper investigates the mediating effect of initial income, extends the analyses to a longer period, and controls for county-invariant heterogeneities.

Another important contribution of our paper is with regard to data: we provide a county-level data set on income inequality for China. Using household survey data, we construct various measures of the income distribution for Chinese counties during 1989–2015. The data from the China Health and Nutrition Survey (CHNS) are available for up to 72 counties of 12 provinces in China. The reason why we chose the CHNS as our main data source for constructing county-level measures of inequality is that the CHNS yields the largest number of county-survey year observations.

Importantly, the CHNS survey data yield a time series for country-level measures of inequality in China that is similar to the time series of country-level measures of inequality for China that were constructed in previous research papers, e.g. Chen and Ravallion (2007) and Piketty et al. (2019). We also document that the CHNS survey data yield a time series of country-level measures of inequality for China that is similar to that of other potential data sources, such as the CHIP, the CFPS, and the NBS. In that regard, we are confident that the CHNS provides an accurate picture of inequality trends in China. The CHNS data yield comparable trends for country-level measures of inequality in China to that of other data sets, which for the purpose of our county-level econometric analysis are not ideal however. Ultimately, we want to use county-level data because this enables a much more rigorous econometric analysis of the relationship between inequality and growth in China than country-level data.

Our county-level data set reveals several interesting stylized facts about inequality. With regard to trends: the median county-level Gini in China increased between 1989 and 2015 by about 12 percentage points–from 0.35 in 1989 to 0.47 in 2015. This is about as large of an increase in inequality as what country-level inequality data show for China during that time period (discussed in detail in Sect. 2). The main value added of our county-level data set is, of course, not that it enables to provide an understanding of the median inequality trend for Chinese counties. The main value added of our subnational panel data set is that it enables to gain an understanding of how within-county inequality differs across counties in China.

There is remarkable variation of within-county inequality in the cross section of Chinese counties. For example, in 1989, which is the first survey year of the CHNS, our data show that the 10 per cent most unequal counties had a Gini coefficient of 0.48 or more. In contrast, the 10 per cent most equal counties in 1989 had a Gini coefficient of 0.26 or less. This means that, in 1989, there is a more than 22 percentage points difference between the bottom and top 10th percentile of the county-level Gini. In 2015, that difference is even larger, amounting to around 26 percentage points. The cross-sectional differences in the income Ginis of Chinese counties are almost as large as the cross-sectional differences in the income Ginis of countries. For example, according to the Standardized World Income Inequality Database of Solt (2020), in 1989 the difference in the Gini between a country at the bottom 10th percentile and a country at the top 10th percentile is around 16 percentage points; in 2015 that difference was around 17 percentage points.

We also make a methodological contribution to the empirical literature that has estimated effects of inequality on economic growth: we carefully address the issue of instrument relevance in our sys-GMM regressions. As in all instrumental variables regressions, the strength of the instruments has to be carefully assessed. A weak statistical relationship between instruments and endogenous variables leads to biased estimates. Testing for weak instruments in diff-GMM is straightforward, but less so, for sys-GMM. Bazzi and Clemens (2013) unbundle the system into the level and the first-difference equation, and then investigate the strength of instruments separately for each equation. One issue, however, with the Bazzi and Clemens (2013) approach is that it does not assess the joint instrument strength for the whole system. We apply a new method for assessing joint instrument strength in sys-GMM, which was developed by Kripfganz (2019). The basic idea is to transform the moment conditions for the level and difference equations in sys-GMM into that for one level equation, and then obtain standard diagnostics statistics of joint instrument strength.

The mainstream approach in the applied IV literature has been to assess instrument strength based on the F-statistic on the null hypothesis that the instruments have jointly a zero effect on the endogenous variables. For our baseline estimates, the Kleibergen Paap F-statistic on the null of a zero joint effect of the instruments in the system is around 40. This means that one can rule out that our sys-GMM regressions are subject to a large weak instrument bias according to the criteria developed in Stock and Yogo (2005). One issue, however, is that the critical values in Stock and Yogo (2005) were developed under the assumption of i.i.d. errors. Olea and Pflueger (2013) develop critical values when errors are not conditionally homoskedastic and serially uncorrelated. Olea and Pflueger (2013) consider a model with only one endogenous variable; while Stock and Yogo (2005) develop critical values also for the case (which is the relevant one for our paper) when the model has multiple endogenous variables.

Young (2021) discusses the issue of leverage and non-i.i.d errors for IV regressions in detail. He suggests that when errors are non-i.i.d. the bootstrap is preferable with regard to inference. Following this suggestion, we use a wild-restricted-efficient cluster bootstrap over the t-statistic. The wild-restricted-efficient cluster bootstrap yields confidence intervals for the coefficients of interest that are similar to the GMM asymptotic 95% confidence intervals. We also show that our results are robust to a 99% winsorization. This suggests that extreme observations are unlikely the cause for the significant relationships between inequality and growth that we uncover.

That the instruments in our sys-GMM regressions are relevant is an improvement of our paper relative to previous literature. GMM is a widely used method in the empirical literature on inequality and economic growth.[1] Previous empirical literature that used GMM did not address instrument relevance. This is a major shortcoming. Examples of papers are Forbes (2000), Halter et al. (2014), and Berg et al. (2018), among others. For the data sets and econometric model specifications of papers that Kraay (2015) reviews, there is clear evidence that the instruments are weak–and hence the results in the papers reviewed by Kraay (2015) are subject to weak instrument bias.

Diff- and sys-GMM were used in the literature to address, primarily, two types of biases: the Nickel bias in dynamic models with fixed effects; and endogeneity bias arising from reverse causality, i.e. an observed correlation could simply reflect causation running from growth to inequality. In order for diff- and sys-GMM estimates to not be subject to those biases, a necessary condition is that the instruments are relevant.

As is well known, the sys-GMM estimator is not well suited to deal with omitted variables bias. To reduce concerns that our estimates are subject to omitted variables bias, we include in our panel model time fixed effects.[2] The control for time fixed effects is particularly advantageous for our aim of estimating a causal effect of inequality on growth in our panel of Chinese counties. The time fixed effects account for any China-wide variable that is time varying. During the 1980s and 1990s there were a multitude of China-wide reforms: e.g. the de-collectivization of agriculture, the opening up of the country to foreign investment, the permission for entrepreneurs to start businesses, the privatization and contracting out of state-owned industry, and the lifting of price controls. By including time fixed effects in the model, we effectively control for these China-wide reforms.

The remainder of this paper is organized as follows. Section 2 discusses related empirical literature. Section 3 provides a detailed description of the survey data that we use to estimate the econometric model. Section 4 discusses some stylized facts about inequality and growth in China. Section 5 describes the econometric model and estimation strategy. In Sect. 6, we present the empirical results from the estimates of the econometric model. Section 7 provides several robustness checks. Section 8 concludes.

[1] Notable exceptions are Easterly (2007), Galor et al. (2009), and Brueckner and Lederman (2018). These papers do not rely on internal instruments, but rather use external instruments. Another approach for identifying causal effects is based on using industry-level data; see Erman and te Kaat (2019).

[2] Our GMM estimates also account for county-fixed effects. For the level equation, the assumption is that lags of the first-differenced variables have zero correlation with the county fixed effects, and for the difference equation the county fixed effects that appear in a level equation drop out.

2 Related empirical literature

In this section, we discuss empirical papers that have estimated effects of inequality on economic growth using subnational data. We begin by discussing two papers that have provided estimates, specifically, for China on the relationship inequality and growth. The two papers are: Ravallion and Chen (2007), and Benjamin et al. (2011). We then discuss two papers–Panizza (2002), and Litschig and Lombardi (2019)–that use subnational data for the USA and Brazil, respectively, to estimate effects of inequality on growth. We do not discuss empirical papers that use cross-country data; we refer the interested reader to Kraay (2015) who reviews some important papers in the cross-country empirical literature on inequality and growth.

For a sample of 28 Chinese provinces, Ravallion and Chen (2007) provide cross-sectional estimates of the relationship between inequality and growth. The cross-sectional estimates reported in Ravallion and Chen show a significant negative correlation between the initial Gini and growth during 1980–2001. One difference between our paper and Ravallion and Chen is that we emphasize a within- relationship between inequality and growth; Ravallion and Chen only report estimates for a cross-sectional relationship. The question that we seek to address in this paper is: what happens to economic growth if inequality in China increases? To answer that question, one needs estimates of a within-relationship; estimates of a cross-sectional relationship are not informative for answering the above question. Another difference between our paper and Ravallion and Chen is that our estimates are based on a much larger sample: our panel of Chinese counties has more than 400 observations; the sample size of the estimates reported in Ravallion and Chen is much much smaller than that – Ravallion and Chen estimate their model on 28 observations. A third difference between our paper and Chen and Ravallion is that we examine how initial income affects the relationship between inequality and growth while Ravallion and Chen do not examine how initial income affects this relationship.

Benjamin et al. (2011) provide estimates of the effects of inequality on income growth in China. These authors use village-level data. Their panel includes 82 villages in 9 provinces during the period 1987–2002. Benjamin et al.'s main measure of inequality is the mean log deviation, i.e. the difference between the log of mean income and the mean of log income. Based on cross-section regressions, the authors find a negative relationship between the mean log deviation and the change in the mean of households' log incomes per capita.

With regard to model specification and identification, there are three main differences between Benjamin et al. and our paper. First, in their published paper, Benjamin et al. (2011) do not report instrumental variables estimates; these authors only report least squares estimates. The results in Benjamin et al. therefore cannot be interpreted as causal. A negative cross-village relationship between inequality and growth could also be interpreted as follows: an increase in average income reduces inequality. Indeed, the Galor and Zeira (1993) model predicts that as average incomes rise inequality decreases. We provide instrumental variables estimates to ensure that there is no such reverse causality bias. Second, in their baseline model Benjamin et al. (2011) do not include fixed effects. We, on the other hand, use as a baseline a model that includes fixed effects. The third difference is that Benjamin et al. (2011) do not include in their model an interaction term between inequality and initial average income.

Using the publicly available panel data set of Benjamin et al. (2011), we show that sys-GMM estimates of an interaction model yield insignificant coefficients on inequality and on the interaction between inequality and initial average income.[3] In our view, the reason for why sys-GMM estimates of the interaction model are insignificant when the model is estimated on the Benjamin et al. data set is that this data set has only a limited time coverage, i.e. the data set ends in 2002, and hence it does not include relatively more recent years during the 2000s and 2010s when average incomes were much higher than during the 1990s. In contrast, our panel data set that is based on CHNS data covers the time period 1989–2015.

We show, using our baseline CHNS data set, that estimating the model for the period 1989–2004, which is similar to the time span of the Benjamin et al. (2011) panel data set, yields insignificant estimates of the interaction model. Only when the interaction model is estimated based on CHNS data for the longest possible time period, i.e. 1989–2015, are the estimated coefficients on inequality and on the interaction between inequality and initial average income significantly different from zero.

We are only aware of two other papers on income inequality and economic growth that use subnational data: Litschig and Lombardi (2019), and Panizza (2002). Litschig and Lombardi (2019) provide estimates of the effects that inequality has on economic growth in Brazil.[4] These authors use subnational data during the period 1970–2000. Litschig and Lombardi find a significant positive effect of lower-tail inequality on growth; upper-tail inequality has no significant effect on growth in their data set. Litschig and Lombardi show that the significant positive effect of lower-tail inequality on growth is limited to those regions in Brazil that in 1970 were relatively poor. For lower-tail inequality, the results documented by Litschig and Lombardi for Brazil are in line with our results for China.

Panizza (2002) provides estimates of the relationship between inequality and growth for a panel of US states during the period 1940–1980. Controlling for fixed effects, Panizza's GMM regressions show a significant negative effect of the Gini on economic growth. According to data from the Penn World Table, version 10.0: in 1960, i.e. the midpoint of Panizza's sample, constant price PPP GDP per capita in the USA was around 20,000 USD. Hence, in Panizza's sample average income is relatively high. The findings in Panizza (2002) are therefore consistent with the estimates from our interaction model on Chinese county-level data, which showed that, at relatively high levels of initial average income inequality has a negative effect on economic growth.

3 The China health and nutrition survey

Our main data source for constructing county-level measures of inequality and average incomes is the China Health and Nutrition Survey (CHNS). This is one of the most

[3] The relevant estimates are discussed in online appendix.

[4] Litschig and Lombardi (2019) also provide diff-GMM and sys-GMM estimates; however, they do not address the issue of instrument relevance.

comprehensive household survey data sets that exist for China. The CHNS is conducted jointly by the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health at the Chinese Center for Disease Control and Prevention. The website for the CHNS is: https://www.cpc.unc.edu/projects/china.

The survey covers 3428–5812 households in 48–72 counties from 12 provinces during 1989–2015. The 12 provinces are Beijing, Chongqing, Shanghai, Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi, and Guizhou. Seven out of 12 provinces were part of all surveys conducted during 1989–2015. Liaoning was unable to participate in the survey for natural disaster, political and administrative reasons in 1997, and returned in 2000. Heilongjiang took part in the survey since 1997. Beijing, Chongqing and Shanghai were covered by the survey from 2011 onwards.

A multistage, random cluster sampling method is used to collect data. Specifically, within each province, four counties are randomly selected with a weighted sampling technique, stratified by income levels. And two cities, normally the provincial capital and a lower income city, are selected. Then, in the next stage (for a lower administrative level), villages and towns are randomly selected in a county; urban districts are randomly selected in a city. These areas are referred to as communities. In the last stage, households are randomly selected in the selected communities.

The CHNS is the most appropriate data source for our paper's endeavor, which is to provide estimates of the relationship between inequality and growth of Chinese counties. There are several reasons for why we chose the CHNS. First, the CHNS provides the largest number of time-series observations among household survey data that are available for China at the subnational level. Starting from 1989, the CHNS collects data every 2–4 years and is still ongoing. The available waves are: 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011, and 2015. Second, the CHNS tracks the respondents throughout time. One benefit from this is that the CHNS data enable us to construct measures of inequality across counties and time that are based on the same measurement standards and sampling method. Further, we can relate the constructed measures of the income distribution to the average income level, which is based on the same sampled households that are used for computing the income distribution at the county level.

Another advantage of the CHNS is that this data source provides a price index that we can use to deflate nominal income. That is, from the same data source, we have information on the price index, which in turn enables us to compute so-called cost-of-living adjusted measures of inequality. This is important. The purchasing power of one unit of currency differs a lot among Chinese counties. For example, according to the price index of the CHNS, for the year 2015, the cost for a fixed basket of goods ranges from 81 to 133 among counties in the Liaoning province (the cost for the same fixed basket in the urban area in Liaoning province in 2015 was set equal to 100). Across years, costs of the same goods in the Liaoning province in 1989 are 31 in the rural areas and 34 in the urban area. Comparable income across regions and years is important for calculating real income growth and time-varying inequality indices that are based on real incomes of households. For the whole sample, the average Gini based on PPP income is about 1.5 percentage points lower than the Gini based on nominal income.

According to the CHNS, household income is defined as the sum of net income (gross revenue minus expenditures) from business, farming, fishing, gardening, and raising livestock, plus wages (retirement and non-retirement), subsidies, and other income. Note that according to this definition, household income can be negative. When observations with negative incomes are part of the data set, the constructed Gini can be larger than one. To deal with this, we adjust the Gini following the method proposed by Chen (1982). We have also computed the Gini for gross income. We will report estimation results that use the gross Gini as a robustness check.

4 Economic growth and income inequality in China: stylized facts

In this section we discuss some intriguing stylized facts about economic growth and income inequality in China. For China's country-level time-series evolution of inequality and real average incomes, two very interesting stylized facts are as follows:

1. During the 1980s and 1990s, there was a large increase in inequality (i.e. an upward trend in inequality) and a significant increase in real average income (i.e. an upward trend in real average income).
2. From about the mid-2000s to the present, there is no visible upward in inequality – if anything, inequality has been on a slight downward trend – while real average income was on a steep upward trend.

We elaborate on these two stylized facts in Sect. 4.1. We note upfront that we are not the first to point out these two stylized facts: they have been pointed out by several other authors, e.g. Luo et al. (2020) and Kanbur et al. (2021). We point out here these two stylized facts so that the reader is well aware of the big picture. Another reason why we point out these two stylized facts is that they provide a benchmark for comparison to: we will show that with the CHNS data we can match the two stylized facts on the country-level time-series evolution of inequality and real averages income in China. These two stylized facts have been pointed out by previous literature based on different data sources. It is important and good to know that different data sources yield the same two stylized facts (at the country level) for China with regard to the time-series evolution of the income Gini and real average income.

In Sect. 4.2 we elaborate on stylized facts about cross-sectional differences in inequality and real average incomes of Chinese counties. These stylized facts are based on (county-level) measures of inequality and real average incomes that we constructed using data provided by the CHNS. Two very interesting stylized facts about within-county inequality that becomes apparent from our constructed county-level panel data set are:

1. Already in the late 1980s (the first survey year is 1989), there were some, though not many, Chinese counties with a very unequal distribution of income.
2. During 1989–2015, there is a strong positive co-movement between the average (or median) county-level Gini and the Gini for China as a whole. That is, the

time-series behavior of the median county-level Gini is similar to the time-series behavior of the country-level Gini.

4.1 Stylized facts at the country-level

Since the 1980s, there has been a remarkable increase in China's real GDP. According to the Penn World Tables, version 10.0 (Feenstra et al. 2015), between 1980 and 2015, there was a more than ten-fold increase in China's real GDP: China's constant price PPP GDP was around \$170 billion in 1980; 35 years later, in 2015, that number was a remarkable \$18 trillion. However, while China's trend growth in real GDP was positive throughout this entire period, there are two regimes with regards to trends in inequality that are clearly visible (see Fig. 1).

According to the CHNS data, the country-level Gini for China increased between 1989 and 2006 by around 12 percentage points: CHNS data show that the country-level Gini for China was 0.39 in 1989 and 0.51 in 2006. Ravallion and Chen's data show that China's country-level Gini increased during 1981–2001 by around 13 percentage points. A similar upward trend in inequality for the pre-2006 period is visible from the CHIP data. According to CHIP, the country-level Gini in China increased between 1988 and 2007 by around 11 percentage points.

For the post-2006 period, there is no visible upward trend in the country-level Gini for China. According to CHNS data, the Gini for China was around 0.5 in all three waves conducted post-2006. For the surveys conducted in 2009, 2011, 2015, the country-level Ginis for China, according to CHNS data, are 0.50, 0.47, and 0.50, respectively. That there is no upward trend in the country-level Gini for China in the

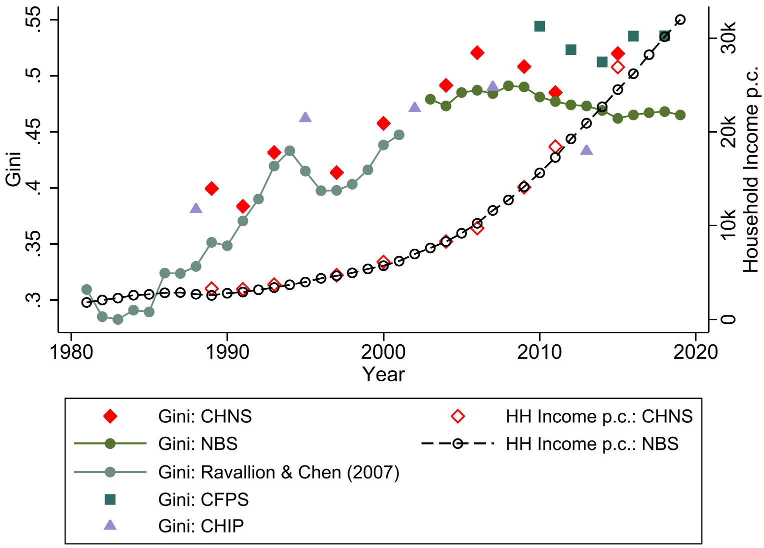

Fig. 1 Inequality and household income per capita in China. The figure plots on the left-hand side of the y-axis the country-level Gini for China. On the right-hand side of the y-axis is household income per capita

post-2006 period is also apparent from other data sources, e.g. the CFPS or the NBS. For example, according to the NBS, the country-level Gini for China was 0.48 in 2007 and 0.47 in 2019. The country-level data suggest that, for the post-2006 period, the trend in inequality for China was either mostly flat, or perhaps, slightly downwards; certainly, not upwards.

Household income per capita in China was on an upward trend throughout the entire period 1980–2020. See Fig. 1 (right-hand side, y-axis) where household income per capita is in yuan. During 1980–2000, the per annum growth rate of constant-price household income per capita was 6.2 per cent. During 2001–2020, growth of constant-price household income per capita was even faster, – around 9.5 per cent. These computed growth rates of household income per capita are based on the NBS data.

CHNS data yield numbers for household income per capita in China that are similar to NBS data. For example, in the year 1997, household income per capita was 4721 yuan according to the CHNS while according to NBS household income per capita was 4653. In 2015, household income per capita was 26,916 yuan according to CHNS while according to NBS household income per capita was 24,512 yuan. In terms of growth rates, these are also similar for CHNS and NBS: for the longest overlapping time-period for which data exist for both CHNS and NBS, i.e. 1989–2015, the CHNS (NBS) yields a per annum growth rate of constant-price household income per capita of 8.4% (9.1%).

The growth rate of household income per capita in China is similar to the growth rate of China's GDP per capita during 1980–2020. Appendix Fig. 10 shows the time series of household income per capita and GDP per capita; both of these variables are in 2017-constant prices, in USD, and PPP adjusted. We use the NBS data on nominal household income per capita, and the PPP conversion rates[5] from the PWT to compute a series of constant price PPP household income per capita. According to NBS data, China's constant-price PPP household income per capita grew at an average rate of 4.0 per cent per annum during 1980–2000; for the period 2000–2019, the per annum growth rate of real household income per capita in China was 6.3 per cent. According to the PWT, during 1980–2000, constant price PPP GDP per capita grew at an average rate of 5.0 per cent per annum; for the period 2000–2019, the per annum growth rate of constant price PPP GDP per capita in China was around 6.5 per cent.

Figure 1 is consistent with Kuznet's (1955) seminal paper on economic growth and income inequality. In that paper, Kuznet noted, for the three major industrial countries at the time – which were the USA, England, and Germany – that, towards the end of the 19th and the beginning of the twentieth century: inequality first increased and then decreased, while average income per capita in those countries was on an upward trend throughout the entire period. The mechanism for the observed inequality-growth relationship in Fig. 1 may, however, not be the same as in Kuznet (1955): Ravallion and Chen (2021) argue that structural transformation, including rural–urban migration, cannot explain the time-series relationship between inequality and growth in China. The theoretical model of Galor and Zeira (1993) that inspires our paper's econometric

[5] The PPP conversion rate used in this paper is equal to the ratio of nominal GDP (from NBS) over the output-side real GDP at chained PPPs in million 2017 US dollar (from PWT).

model specification and empirical analysis, discussed later in the paper, in Sects. 5 and 6, generates an inverted U-shaped relationship between inequality and average income; and the mechanism is distinct from the structural transformation mechanism alluded to in Kuznet (1955). According to the model of Galor and Zeira (1993): when initial income is low, an increase in inequality has a positive effect on growth; when initial income is high an increase in inequality has a negative effect on growth.

Piketty et al. (2019) provide novel estimates of income shares for China during 1978–2015. Piketty et al. carefully construct such estimates by combining tax data on high-income individuals with household surveys and national accounts. According to Piketty et al. (2019) estimates, China's top 10 per cent income share increased by over 14 percentage points during the period 1980–2006: from 27 per cent in 1980 to 42 per cent in 2006. For the period 2007–2015, there is no visible upward trend in the income share of the top 10 per cent. If anything, there was a slight decrease in the income share of the top 10 per cent. According to Piketty et al.'s data, in 2015, the income share of the top 10 per cent in China was 41 per cent.

The CHNS yields a time series of the income share of the top 10 per cent in China that is similar to that of Piketty et al. (2019). See Appendix Fig. 11. In Appendix Fig. 11 we plot the income share of the top 10 per cent, based on CHNS data, and the income share of the top 10 per cent from Piketty et al. (2019). During the time period for which there are overlapping observations, i.e. 1989–2015, there is a strong positive correlation between these two time series: the correlation coefficient between the two series is 0.87.

CHNS data and Piketty et al.'s data show similar trends during 1989–2015 for the income share of the top 10 per cent in China. According to CHNS data, China's top 10 per cent income share increased by 11 percentage points during the period 1989–2006. This is about as large of an increase as what Piketty et al.'s data show for this time period. Again there is a visible trend break in inequality for the top 10 per cent income share around the mid-2000s. For the period 2006–2015, China's top 10 per cent income share increased by 0 percentage points according to both, CHNS data and Piketty et al.'s estimates. That is, the trend in the income share of the top 10 per cent after the mid-2000s is flat, according to both CHNS data and Piketty et al.'s data.

Piketty et al.'s data yield a slightly higher income share of the top 10 per cent, for all years, than the CHNS data. But, the differences in the levels of these two time series, for any given year, are nowhere near as large as the differences that Piketty et al. (2019, page 2471) point out with regard to other survey data. Specifically, Piketty et al. (2019, page 2471) write: "For recent years, we find top 10 percent income shares around 41 percent of total national income (versus 31 percent in surveys) and top 1 percent income shares around 14 percent (versus 6.6 percent in surveys)". That is, according to Piketty et al. (2019, page 2471), there is an about 10 percentage points difference, for the year 2015, between Piketty et al.'s data and that of the (other) survey data. (It is not entirely clear to us, from reading Piketty et al., which specific survey data Piketty et al. are referring to on page 2471 of their published paper.) The difference in the income share of the top 10 per cent based on CHNS data and Piketty et al.'s data is only around 3 percentage points. – This is much smaller than the 10 percentage points difference alluded to by Piketty et al. (2019) on page 2471 of their published paper. In the year 2015, the CHNS data yield an income share of the top 10 per cent in China of 38 per cent, while Piketty et al.'s data show an income share of the top 10 per cent in China of 41 per cent.

We have also computed, for comparability purposes, the income share of the top 1 per cent based on the CHNS data. According to the CHNS data, the income share of the top 1 per cent in China for the year 2015 was 13 per cent. This is only about 1 percentage points less than what Piketty et al.'s data show for the year 2015. According to Piketty et al. (2019, page 2471), other survey data imply a much larger difference–around 6 percentage points.

In sum: the CHNS data yield similar trends in inequality at the country-level for China to other data sets, including the data set by Piketty et al. (2019) which was constructed by these authors very carefully by combining tax data on high-income individuals with household surveys and national accounts.

4.2 Stylized facts at the county-level

Table 1 shows summary statistics for log household income per capita, the growth rate of household income per capita, and various measures of inequality of Chinese counties during 1989–2015. All variables in Table 1 are at the county-level. We computed the county-level variables based on the household data provided by the CHNS.

For the first survey year, 1989, one can see that household income per capita of Chinese counties was around 8.0 logs. By 2015, the household income per capita of Chinese counties was around 10.0 logs. These numbers imply that for the average Chinese county household income per capita increased between 1989 and 2015 by around 2.0 logs.

Household income per capita increased in all Chinese counties between 1989 and 2015. The minimum increase in log household income per capita of a county was 1.1 logs; the maximum was 3.0 logs. Even the poorest county in 2015 had a household income per capita that exceeded the household income per capita of the richest county in 1989: In 1989, the richest Chinese county had a household income per capita of about 8.6 logs; a quarter of a century later, in 2015, the poorest Chinese county had a household income per capita of about 8.9 logs.

Cross-sectional differences in household income per capita of Chinese counties increased substantially between 1989 and 2015. In 1989, the difference in household income per capita between a county at the 10th percentile and a county at the 90th percentile was 0.8 logs. By 2015, that difference was 1.1 logs. In 2015 the cross-county standard deviation of log household income per capita of Chinese counties was 0.42, while in 1989 that standard deviation was about 0.35. These numbers imply that the cross-county standard deviation of log household income per capita of Chinese counties increased between 1989 and 2015 by around 20 per cent. These statistics suggest that inequalities (in terms of household income per capita) *across* Chinese counties increased between 1989 and 2015.

In the next paragraphs, we discuss measures of inequalities *within* Chinese counties. Our baseline measure of within-county inequality is the Gini. For each county-survey year, we computed the Gini based on the household income per capita data provided by the CHNS. As already noted in Sect. 2, the CHNS provides data on the price

Table 1 Descriptive statistics

	Obs	Mean	Percentiles		
			10th	50th	90th
Log (income p.c.)					
1989	48	8.030	7.564	8.023	8.445
2000	54	8.649	8.165	8.658	9.168
2009	54	9.459	9.027	9.481	9.886
2015	72	10.001	9.350	10.056	10.509
1989–2015	552	8.928	7.898	8.868	10.032
Annualized growth rates					
1997–2000	48	0.077	− 0.010	0.086	0.171
2006–2009	54	0.128	0.026	0.137	0.227
2011–2015	72	0.093	0.017	0.084	0.177
1989–2015	474	0.072	− 0.052	0.075	0.177
Gini coefficient					
1989	48	0.353	0.257	0.348	0.475
2000	54	0.406	0.326	0.406	0.484
2009	54	0.457	0.378	0.445	0.548
2015	72	0.462	0.328	0.465	0.585
1989–2015	552	0.408	0.296	0.404	0.512
Other inequality measures					
Mean log deviation	552	0.389	0.173	0.360	0.639
90–10 ratio	552	10.085	4.092	7.593	18.936
75–25 ratio	552	2.966	1.944	2.732	4.214
Income shares of first quintile	552	0.043	0.018	0.043	0.074
Income shares of second quintile	552	0.105	0.075	0.104	0.137
Income shares of third quintile	552	0.159	0.129	0.161	0.185
Income shares of fourth quintile	552	0.231	0.202	0.234	0.257
Income shares of fifth quintile	552	0.462	0.375	0.455	0.548

The statistics are calculated at the county level based on the CHNS data. The whole sample covers waves of 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011 and 2015. All income data, measured in *yuan*, are deflated into 2015 prices

level for each county. When computing the county-level Ginis, and other measures of inequality, to be discussed shortly, we use households' real incomes, which are computed as the nominal income per capita of a household, in yuan, divided by the price level in each county-year.

Inequality of the average Chinese county increased between 1989 and 2015. In 1989, the average Chinese county had a Gini of 0.35. In 2015, the average Chinese county had a Gini of 0.46. This implies a more than 11 percentage point increase in the average Gini between 1989 and 2015. However, there are two regimes with regard

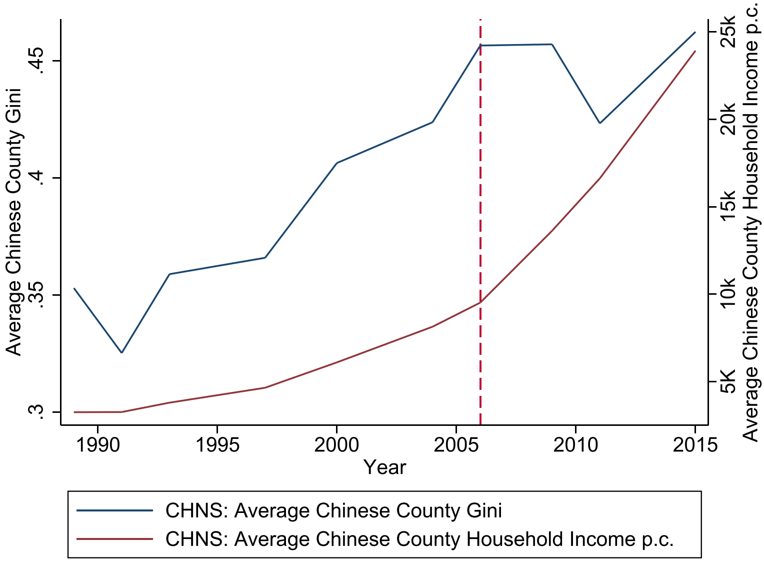

Fig. 2 Inequality and household income per capita in the average Chinese county

to trends in the average Gini: between 1989 and 2006 the average Gini increased by around 11 percentage, while between 2006 and 2015, the change in the average Gini is zero. That is, in 2006 the Gini of the average Chinese county was 0.46, which is about the same as the Gini of the average Chinese county in 2015.

That there are two regimes with regard to trends in inequality of the average Chinese county, while income per capita of the average Chinese county was on an upward trend throughout the entire period of 1989–2015, is similar to what country-level data show for China. In Fig. 2, we plot for all survey years of the CHNS, the Gini and household income per capita of the average Chinese county. Comparing Figs. 2 to 1, it is apparent that for the average Chinese county the relationship between inequality and income per capita is similar to the country-level relationship between inequality and income per capita in China during 1989–2015.

There is a high correlation during 1989–2015 between the country-level Gini for China and the Gini of the average Chinese county. The high correlation between overall inequality in China and inequality of the average Chinese county is apparent from Fig. 3. In that figure, for the period 1989–2015 we plot based on the CHNS data the Gini for China and the average, as well as the median, Gini of Chinese counties.[6] During 1989–2015, there is a strong positive co-movement between the

[6] The formula for the country-level Gini of China in year t is $gini_t = \frac{\sum_{c=1}^{m_t}\sum_{d=1}^{m_t}\sum_{i=1}^{n_{ct}}\sum_{j=1}^{n_{dt}} |y_{ict}-y_{jdt}|}{2\sum_{c=1}^{m_t} n_{ct}\sum_{c=1}^{m_t}\sum_{i=1}^{n_{ct}} y_{ict}}$. The formula for the Gini of a Chinese county c in year t is $gini_{ct} = \frac{\sum_{i=1}^{n_{ct}}\sum_{j=1}^{n_{ct}} |y_{ict}-y_{jct}|}{2n_{ct}\sum_{i=1}^{n_{ct}} y_{ict}}$. In these formulas, y_{ict}

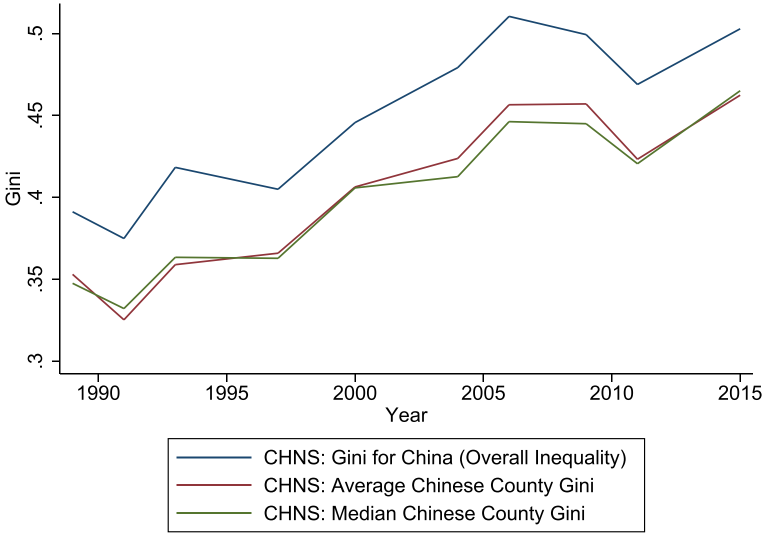

Fig. 3 The country-level gini for China vs. the average and median gini of Chinese counties

country-level Gini for China and the average and the median Gini of Chinese counties. The correlation between the average (median) county-level Gini and the country-level Gini for China during 1989–2015 is 0.988 (0.978).

Already in 1989, there were some–though not many–Chinese counties with considerable income inequality. This is a very interesting stylized fact. This stylized fact is only apparent from the county-level measures of inequality that we have computed based on CHNS data. According to the county-level Gini that we computed based on the CHNS data, in 1989, thirty per cent of counties had a Gini above 0.40; the 10 per cent most unequal Chinese counties had a Gini of 0.48 or above. That already in 1989 there was significant inequality for some, though not many, Chinese counties is also apparent from other measures of inequality that we computed based on the CHNS data. Consider, for example, the income share of the richest quintile, which is one measure for upper-tail inequality. In 1989, in thirty per cent of Chinese counties the income share of the richest quintile was equal to or above 45 per cent. In the 10 per cent most unequal counties, the income share of the richest quintile was equal to or above 51 per cent.

That such large inequalities existed within some, though not many, Chinese counties in 1989 would have been impossible to infer from country-level measures of inequality for China. Recall that, in 1989, overall inequality in China was still relatively low, compared to say, the 2000s. According to the CHNS data, the Gini for China in the survey year 1989 was 0.39. This means that, in the survey year 1989, for the 10 per cent most unequal Chinese counties the county-level Gini exceeded the country-level Gini for China by 10 percentage points.

Footnote 6 continued
denotes the year t income per capita of household i residing in county c; there are m_t counties; and county c_t has n_{ct} households.

Figure 4 shows kernel density estimate plots of the county-level Gini (Panel A) and the county-level log household income per capita (Panel B) for three selected survey years: 1989, 2000, and 2015. From Panel A of Fig. 4, one can see that over time the density function of both the county-level Gini and the county-level log household income per capita shifted to the right. That is, during the time-period 1989–2015 there was an increase in the means and the medians of the county-level Gini and of household income per capita.

Figure 4 also shows that the dispersion in the county-level Gini increased between 1989 and 2015. In 1989, the standard deviation of the county-level Gini was 0.085 while in 2015 it was 0.104. In 1989 the least unequal Chinese county had a Gini of 0.17; the most unequal county had a Gini of 0.51. This means that, in 1989, there was more than 34 percentage points difference in the Gini between the least and the most unequal Chinese county. About a quarter of a century later, in 2015, that difference was even larger – amounting to around 56 percentage points. In 2015, the most equal Chinese county had a Gini of around 0.23 while the most unequal Chinese county had a Gini of 0.79. In 2015, about 25 per cent of Chinese counties had a Gini that was above the Gini of the most unequal county in 1989.

5 Estimation framework

The dynamic panel model is:

$$g_{i,t} = \alpha \ln y_{i,t-1} + \beta \text{Inequality}_{i,t-1} + \gamma \ln y_{i,t-1} * \text{Inequality}_{i,t-1} + u_i + \lambda_t + \epsilon_{i,t} \tag{1}$$

where $g_{i,t}$ stands for the annualized growth rate of per capita household income in county i between survey year t and *t-1*. $\ln y_{i,t-1}$ is the natural logarithm of per capita household income of county i in survey year *t-1*. $Inequality_{i,t-1}$ is a measure of the county i's income distribution in survey year *t-1*. u_i are county fixed effects; λ_t are time fixed effects; $\epsilon_{i,t}$ is an error term. The marginal effect of inequality on growth is: $\beta + \gamma \ln y_{i,t-1}$.

We estimate Eq. (1) using GMM. Our preferred method of estimation is sys-GMM. As a robustness check on our preferred method of estimation, we will report results from diff-GMM. All of our GMM estimations are done using STATA's *xtdpdgmm* command. We use the collapse option to reduce instrument proliferation: In sys-GMM the number of instruments increases quickly with the number of regressors and the number of time periods (see e.g. Roodman (2009)). First-differences magnify gaps in unbalanced panel data, and we deal with this issue by using the orthogonal option, which minimizes such data loss.

Two standard specification tests in GMM estimations are higher-order serial correlation tests and the Hansen test. The second-order serial correlation test is conducted on the first-differenced error terms. If the error term $\epsilon_{i,t}$ is serially uncorrelated, this implies there is no second-order or higher-order serial correlation for $\Delta\epsilon_{i,t}$. Rejection of the null implies that lagged variables are endogenous, and thus, that the internal

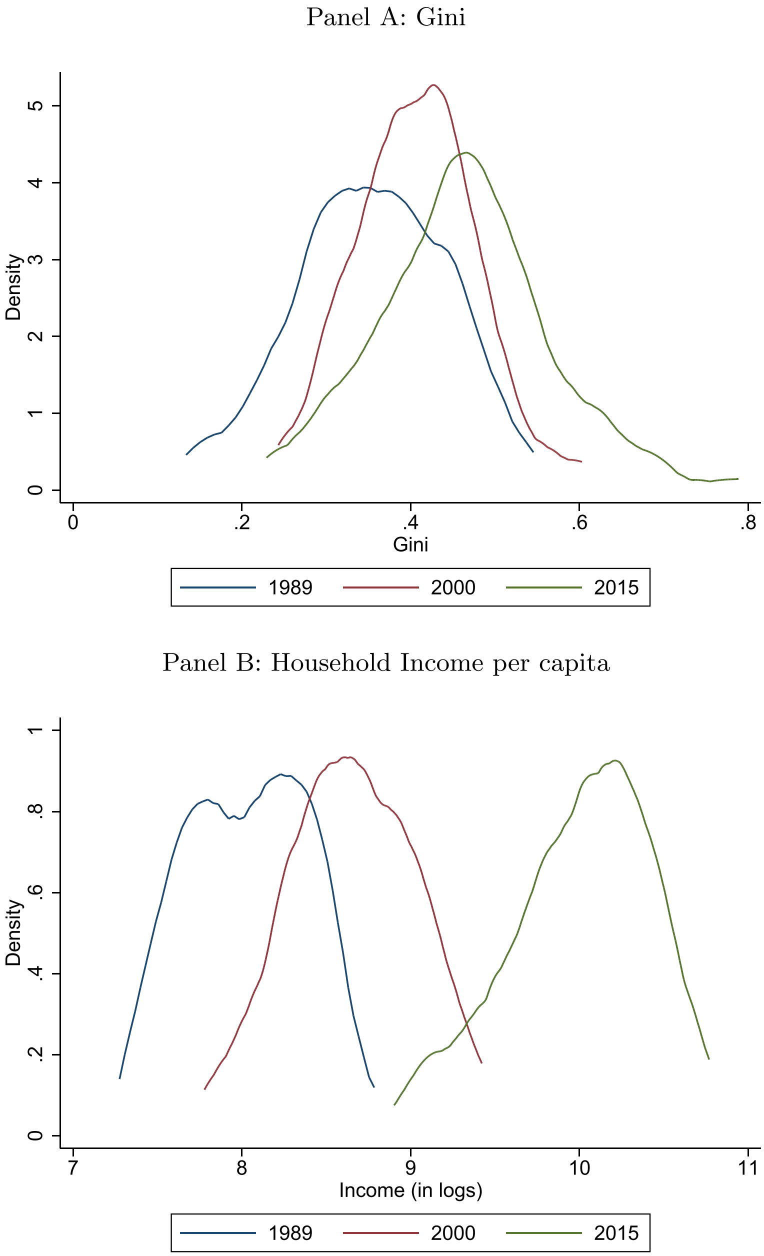

Fig. 4 The distribution of the Gini and household incomes per capita of Chinese counties for three selected survey-years: 1989, 2000, and 2015. Panel A shows a kernel density estimate plot of the Ginis of Chinese counties for three selected survey years: 1989 (blue line), 2000 (red line), 2015 (green line). Panel B shows a kernel density estimate plot of log household income per capita of Chinese counties for three selected survey years: 1989 (blue line), 2000 (red line), 2015 (green line)

instruments are invalid. The other complementary test for assessing instrument validity is the Hansen test, also known as the overidentification test.

We will report two diagnostic statistics for detecting weak instruments: the Kleibergen-Paap LM test and the Kleibergen-Paap F test. For sys-GMM estimation, one would ideally like to have test statistics for the joint hypothesis that the instruments have a zero effect in the level equation and in the first-difference equation. Kripfganz (2019) develops such a test. The basic idea is to transform the system estimator into a level-GMM estimator. And then obtain diagnostics statistics of instrument strength.[7] We compute these statistics after running STATA's command *ivreg2* with the *gmm2s* option on the model specified in Eq. (1). This command produces the same coefficients as two-step sys-GMM when using *xtdpdgmm*.

6 Results

6.1 Baseline GMM estimates

We start by discussing our baseline sys-GMM estimates. Columns (1)–(3) of Table 2 show two-step sys-GMM estimates. One-step sys-GMM estimates are shown in columns (4)–(6). One-step estimators use an arbitrary weighting matrix for moment conditions; two-step estimators use the optimal weighting matrix which is efficient and robust to heteroskedasticity and autocorrelation.[8] Columns (1) and (4) show results for two- and one-step sys-GMM for the largest sample, respectively. Columns (2) and (5) show results for a 99% winsorization, i.e. we estimate the model on a sample where 1 per cent of the lowest Ginis are replaced by the value at the 1st percentile and the 1 per cent of the highest Ginis are replaced by the value at 99th percentile. Columns (3) and (6) show sys-GMM estimates for a limited lag selection.

From Table 2, one can see that the estimated coefficients on the Gini and the interaction term are similar across columns. The estimated coefficients on the Gini are significantly positive, while coefficients on the interaction between the Gini and initial income per capita are significantly negative at the 5% significance level. The negative coefficient on the interaction term means that the effect of inequality on growth is significantly decreasing in initial income per capita.

Figure 5 visualizes how the effect of the Gini on economic growth differs across sample values of initial income. Figure 5 is based on the estimates in column (1) of Table 2. Recall that in the interaction model the marginal effect of the Gini on growth is the sum of the coefficients on the Gini and on the interaction between the Gini and

[7] To put this into the context of previous literature, Bazzi and Clemens (2013) proposed to run separate regressions for the difference and the level equation, using 2SLS and then computing the test diagnostics for each equation separately. Kraay (2015) employs the same method as Bazzi and Clemens (2013) when investigating instrument relevance in previous cross-country papers that used GMM to estimate the effect that inequality has on growth. Although in the AR (1) model illustrated in Blundell et al. (2000) the sys-GMM estimator is a weighted average of the estimators of the difference and level equation, this does not mean that F-statistics on two separate hypotheses (a joint zero effect in the difference equation; a joint zero effect in the level equation), are appropriate for testing instrument relevance in sys-GMM.

[8] In finite samples, the computed standard errors in the two-step GMM estimation are biased downwards. We use the Windmeijer correction in all two-step GMM estimations to address this issue.

Table 2 Baseline Sys-GMM estimates (dependent variable: annualized growth rate)

	Twostep Sys-GMM			Onestep Sys-GMM		
	(1) Baseline	(2) Winsor	(3) Limitedlags	(4) Baseline	(5) Winsor	(6) Limitedlags
L.gini	3.261***	3.250***	2.976**	3.214***	3.223***	3.135***
	(1.018)	(1.043)	(1.130)	(0.996)	(1.003)	(0.998)
L.gini × L.lny	− 0.374***	− 0.371***	− 0.338**	− 0.365***	− 0.366***	− 0.360***
	(0.115)	(0.118)	(0.129)	(0.114)	(0.115)	(0.115)
L.lny	− 0.066	− 0.064	− 0.065	− 0.038	− 0.037	− 0.042
	(0.066)	(0.067)	(0.075)	(0.061)	(0.062)	(0.061)
AR(2) test	0.049	0.066	0.276	0.318	0.331	0.335
p-value	0.961	0.948	0.782	0.751	0.741	0.737
Hansen test	27.949	27.717	17.575	21.392	21.717	8.155
p-value	0.262	0.272	0.129	0.616	0.596	0.227
Instruments						
Number of IVs	36	36	24	36	36	18
IVs for the transformed equation	L(0/7).X†	L(0/7).X	L(0/3).X	L(0/7).X	L(0/7).X	L(0/1).X
IVs for the level equation	D.X	D.X	D.X	D.X	D.X	D.X
Underidentification test						
Kleibergen-Paap LM stat	47.974	48.421	43.018	47.974	48.421	40.491
p-value	0.004	0.003	0.000	0.004	0.003	0.000
Weak identification test						
Kleibergen-Paap Wald stat	40.115	35.840	22.315	40.115	35.840	26.374
Stock-Yogo critical values						
10% maximal IV relative bias	10.74	10.74	10.33	10.74	10.74	9.37
20% maximal IV relative bias	5.89	5.89	5.94	5.89	5.89	5.78
30% maximal IV relative bias	4.24	4.24	4.37	4.24	4.24	4.46
Observations	474	474	474	474	474	474
Number of counties	72	72	72	72	72	72
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 2 (continued)

	Twostep Sys-GMM			Onestep Sys-GMM		
	(1) Baseline	(2) Winsor	(3) Limitedlags	(4) Baseline	(5) Winsor	(6) Limitedlags
	Twostep sys-GMM baseline			Onestep sys-GMM baseline		
Conditional on sample gini at	mean	10th pct	90th pct	mean	10th pct	90th pct
Combined L.lny	– 0.219***	– 0.177***	– 0.258***	– 0.187***	– 0.146***	– 0.225***
	(0.037)	(0.042)	(0.037)	(0.027)	(0.034)	(0.025)

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. $X^{\dagger}$ refers to L.gini, L.gini × L.lny, and L.lny. Columns (2) and (5) show results for a 99% winsorization of gini; Columns (3) and (6) show estimates for a limited lag selection on internal instruments

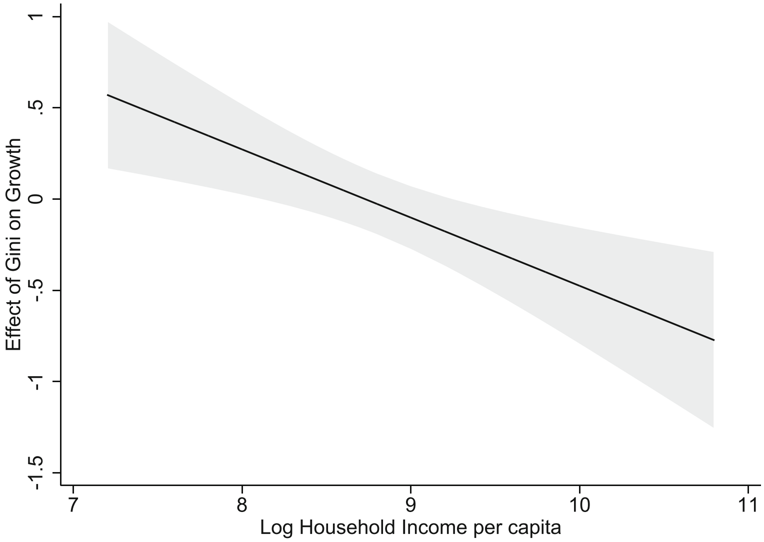

Fig. 5 The effect of the gini on growth. The figure is based on the sys-GMM estimates reported in column (1) of Table 2. On the y-axis of the above figure is the marginal effect of the one-period lagged Gini on the growth rate of household income per capita of Chinese counties. On the x-axis of the above figure is the one-period lagged log household income per capita of Chinese counties. The shaded area is the 95% per cent confidence interval

initial income. Figure 5 plots on the y-axis this sum of coefficients. That is, on the y-axis of Fig. 5 is the marginal effect of the *t-1* Gini on the growth rate of household income per capita of Chinese counties. On the x-axis are the sample values of *t-1* log household income per capita of Chinese counties.

From Fig. 5, one can see that the effect of the Gini on growth is decreasing in initial income. And, that there is a threshold of initial income below which the Gini has a positive effect on growth. Above that threshold the Gini has a negative effect on growth. The threshold is at around 6119 yuan (8.72 logs) in 2015 prices.

Figure 6 shows scatter plots of the relationship between economic growth and the Gini for two sub-samples: rich Chinese counties (right-hand-side panel in Fig. 6), defined as those counties with household income per capita above 8.72 logs; and poor Chinese counties (left-hand side panel in Fig. 6), defined as counties with household income per capita below 8.72 logs. From the left-hand side panel of Fig. 6 one can see that for the relatively poor Chinese counties the relationship between the Gini and growth is positive. In relatively rich Chinese counties (see the right-hand side panel of Fig. 6), the relationship between the Gini and growth is negative.

Recall from the discussion in Sect. 4.2 that household incomes per capita have increased considerably in Chinese counties during the sample period, which is 1989–2015. In 1989, all counties in the sample had a household income per capita below 6119 yuan. Thus, for the late 1980s, our model estimates suggest that inequality in China had a positive effect on economic growth. In contrast, by 2015, all counties

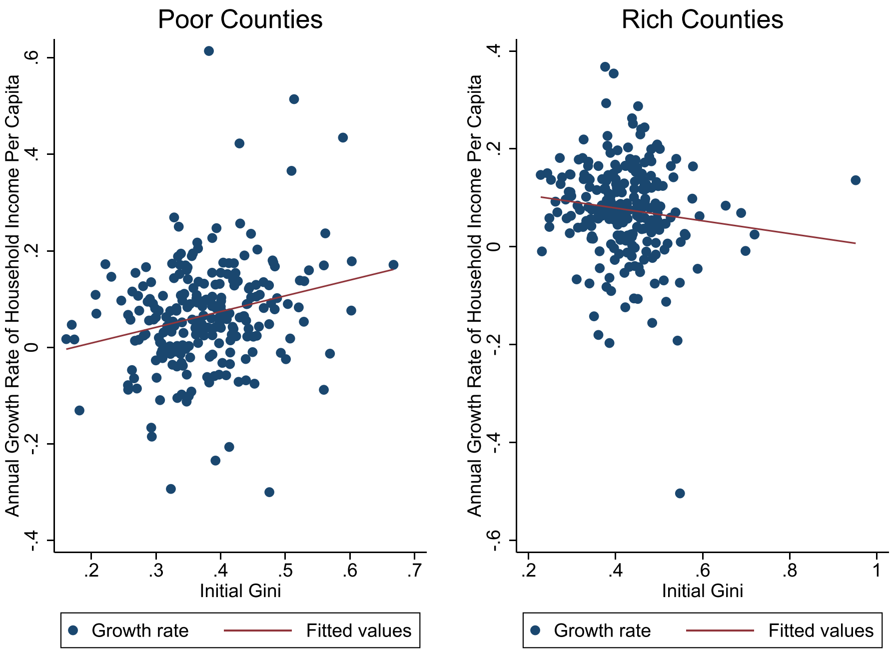

Fig. 6 Scatter plots of the gini and growth of household income per capita in rich and poor Chinese counties. Poor counties are defined as those Chinese counties that in the sample have an initial household income per capita of less than 8.72 logs. Rich counties are defined as those Chinese counties that in the sample have an initial household income per capita of more than 8.72 logs. The threshold of 8.72 logs is determined by our baseline sys-GMM estimates: according to the sys-GMM estimates in column (1) of Table 2, at an initial log household income per capita equal to 8.72 logs the marginal effect of the Gini on the growth rate of household income per capita of Chinese counties is equal to zero; see also Fig. 5

in the sample had a household income per capita above 6119 yuan. Hence, for 2015, our model estimates suggest that inequality reduced growth in China.

Quantitatively, the estimated effects are sizable. Consider, for example, the estimates in column (1) of Table 2. According to the estimates in column (1), for an average county in the year 1989–that has a household income per capita of 3071 yuan (8.03 logs)–a 1 percentage point increase in the Gini increases the growth rate by around 0.26 percentage points. For an average county in the year 2015, that has a household income per capita of 22,048 yuan (10.00 logs), a 1 percentage point increase in the Gini decreases the growth rate by around 0.48 percentage points. The larger household income per capita of a county, the more negative is the effect of inequality on growth. For example, for a Chinese county that in the year 2015 is at the top 10th percentile (household income per capita of around 36,000 yuan), our estimates suggest that a 1 percentage point increase in the Gini reduces the growth rate by around 0.7 percentage points.

Regarding convergence: in Eq. (1) the convergence rate is $\alpha + \gamma * gini_{i,t-1}$. Across all specifications of Table 2 there is significant within-county convergence. The bottom part of Table 2 reports the convergence rates of counties' household incomes per capita for different values of the Gini: at the mean, the 10th percentile, and the 90th percentile. The speed of convergence is significantly increasing in the Gini. At the 10th and 90th percentile of the Gini the per annum convergence rates are around 18 per cent and 26 per cent, respectively. At the mean of the Gini the per annum convergence rate is around 22 per cent.

According to the estimates in Table 2, an increase in the level of the Gini has a long-run effect on the level of household income per capita. The long-run effect of a one unit increase in the Gini on the log of household income per capita, evaluated at the sample mean of the Gini is: $(\beta + \gamma * \ln y_{i,t-1})/0.22$. At the 10th and 90th percentile of the Gini these long-run effects are $(\beta + \gamma * \ln y_{i,t-1})/0.18$ and $(\beta + \gamma * \ln y_{i,t-1})/0.26$, respectively. Hence, in absolute value, the long-run effect of an increase in the Gini on the level of average income is larger in counties that are less unequal.

The following two examples illustrate the above result. Consider, at the one extreme, a poor county with a household income per capita at the 10th percentile: According to the estimates in column (1) of Table 2, for hypothetical values of the Gini equal to the 10th, 50th, and 90th percentile – a 1 percentage point increase in the Gini increases in the long-run household income per capita by around 1.7 per cent, 1.4 per cent, and 1.2 per cent, respectively. Now consider, as an example at the other extreme, a rich county that has a household income per capita in the 90th percentile of the sample: according to the estimates in column (1) of Table 2, for hypothetical values of the Gini equal to the 10th, 50th, and 90th percentile – a 1 percentage point increase in the Gini decreases in the long-run household income per capita by around 2.8 per cent, 2.3 per cent, and 1.9 per cent, respectively.

The instruments in the sys-GMM estimations of Table 2 are both valid and relevant. In Table 2 one can see that the p-values of the AR(2) test and the Hansen test exceed 0.1. The Kleibergen-Paap LM statistic and the Kleibergen-Paap F-statistic suggest that the instruments are relevant: The p-value of the LM statistic is smaller than 0.01, and thus one can reject the null that the model is underidentified at the 1% significance level; and the F-statistic is greater than the critical value provided by Stock and Yogo

(2005) for the null hypothesis that the bias in the IV estimates is greater than 10 per cent of the OLS bias.

Confidence intervals based on asymptotic GMM standard errors are similar to bootstrapped intervals. Consider the two-step sys-GMM estimates in column (1) of Table 2: the asymptotic 95% confidence interval for β (i.e. the estimated coefficient on $Gini_{i,t-1}$) is [1.23, 5.20]; and for γ (i.e. the estimated coefficient on $\ln y_{i,t-1} * Gini_{i,t-1}$) the asymptotic 95% confidence interval is [-0.59, -0.14]. Asymptotic confidence intervals are similar for one-step and two-step sys-GMM. For the one-step sys-GMM in column (4) of Table 2, the asymptotic 95% confidence interval for β is [1.20, 5.20]; and for γ the asymptotic 95% confidence interval is [-0.59, -0.14]. For the one-step sys-GMM estimate in column (4) of Table 2, we computed bootstrapped confidence intervals using STATA's *boottest* command (Roodman et al. 2019). The wild restricted efficient bootstrap over the t-statistic, with 1000 replications clustered at the country level, yields 95% confidence intervals for β and γ of [1.18, 5.28] and [− 0.60, − 0.14], respectively. These bootstrapped confidence intervals are only slightly wider than the sys-GMM confidence intervals. Appendix Fig. 12 shows that confidence plots are centered around the point estimates.

Table 3 reports diff-GMM estimates. Diff-GMM yields estimated coefficients on the Gini that are significantly positive, while the estimated coefficients on the interaction term are significantly negative. Diff-GMM thus yields qualitatively the same result as sys-GMM. However, quantitatively, the (absolute) size of the estimated coefficients on the Gini and interaction term are larger for diff-GMM.

For the diff-GMM estimates, the p-values of the AR(2) test and Hansen test are well above 0.1, which indicates that the instruments in the diff-GMM estimations are valid. However, the joint strength of instruments in the diff-GMM estimations is weak: In all columns of Table 3 the Kleibergen-Paap F-statistic is below the critical values provided by Stock and Yogo (2005).

Because standard test diagnostics indicate that instruments are both valid and relevant in the sys-GMM estimates, while instruments are valid but not strong for the diff-GMM estimates, we prefer the sys-GMM estimates over the diff-GMM estimates. In the discussion of robustness checks and extensions of the baseline model that follows we will, therefore, report results based on sys-GMM.

6.2 Heterogeneous effects by initial income quantiles

Table 4 further illustrates evidence on the heterogeneous effects of inequality on growth by initial incomes. More specifically, we estimate the impacts of inequality on growth at different quantiles of initial incomes. In column (1), the initial incomes are divided into three quantiles, column (2) four quantiles, and column (3) five quantiles. The pattern of combined coefficients of the Gini across three columns are similar, i.e. at the lowest quantile of initial incomes, the Gini has positive impacts on subsequent economic growth, while at the highest quantile of initial incomes, the effects of the Gini on subsequent economic growth become significantly negative. These results are consistent with the prediction of baseline estimates where we employ the interaction

Table 3 Diff-GMM estimates (dependent variable: annualized growth rate)

	Twostep Dif-GMM			Onestep Dif-GMM		
	(1) Baseline	(2) Winsor	(3) Limitedlags	(1) Baseline	(2) Winsor	(3) Limitedlags
L.gini	7.065*** (2.161)	6.905*** (2.164)	7.631*** (2.361)	6.303*** (2.122)	6.287*** (2.080)	7.476*** (2.380)
L.gini × L.lny	− 0.814*** (0.247)	− 0.798*** (0.249)	− 0.889*** (0.270)	− 0.723*** (0.246)	− 0.724*** (0.242)	− 0.866*** (0.278)
L.lny	0.149 (0.121)	0.138 (0.121)	0.210* (0.112)	0.128 (0.109)	0.125 (0.107)	0.213* (0.127)
AR(2) test	− 0.359	− 0.352	− 0.282	− 0.171	− 0.169	− 0.217
p-value	0.719	0.725	0.778	0.864	0.866	0.828
Hansen test	20.642	20.719	4.412	18.766	18.925	3.604
p-value	0.481	0.476	0.220	0.600	0.590	0.308
Instruments						
Number of IVs	33	33	15	33	33	15
IVs for the transformed equation	L(0/7).X†	L(0/7).X	L(0/1).X	L(0/7).X	L(0/7).X	L(0/1).X
Underidentification test						
Kleibergen-Paap LM stat	32.260	32.406	12.273	32.260	32.406	12.273
p-value	0.073	0.071	0.015	0.073	0.071	0.015
Weak identification test:						
Kleibergen-Paap Wald stat	3.830	3.944	2.501	3.830	3.944	2.501
Stock-Yogo critical values						
10% maximal IV relative bias	10.70	10.74	7.77	10.70	10.74	7.77
20% maximal IV relative bias	5.91	5.89	5.35	5.91	5.89	5.35
30% maximal IV relative bias	4.24	4.24	4.40	4.24	4.24	4.40
Observations	474	474	474	474	474	474
Number of counties	72	72	72	72	72	72
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. X† refers to L.gini, L.gini × L.lny, and L.lny. Columns (2) and (5) show results for a 99% winsorization of gini; Columns (3) and (6) show estimates for a limited lag selection on internal instruments

Table 4 Sys-GMM estimates for different initial income quantiles (Dependent Variable: annualized growth rate)

Income quantiles	(1) Tertiles	(2) Quartiles	(3) Quintiles
L.gini	0.295	0.486*	0.660*
	(0.214)	(0.249)	(0.394)
L.gini × Q2	− 0.333	− 0.251	− 0.584**
	(0.233)	(0.292)	(0.287)
L.gini × Q3	− 0.590**	− 0.732**	− 0.681
	(0.224)	(0.311)	(0.435)
L.gini × Q4		− 0.776***	− 0.954*
		(0.289)	(0.560)
L.gini × Q5			− 1.154***
			(0.373)
Q2	0.106	0.025	0.170
	(0.087)	(0.114)	(0.116)
Q3	0.189*	0.177	0.210
	(0.098)	(0.124)	(0.164)
Q4		0.139	0.269
		(0.125)	(0.223)
Q5			0.330**
			(0.144)
Combined effects of the gini by income quantiles			
Combined Q2	− 0.038	0.235	0.076
	(0.167)	(0.220)	(0.223)
Combined Q3	− 0.295**	− 0.245*	− 0.021
	(0.120)	(0.131)	(0.135)
Combined Q4		− 0.290**	− 0.294
		(0.127)	(0.267)
Combined Q5			− 0.494***
			(0.158)
AR(2) test	0.862	− 0.047	0.248
p-value	0.389	0.962	0.804
Hansen test	33.592	40.001	28.754
p-value	0.390	0.297	0.373
Instruments			
Number of IVs	46	52	45

Table 4 (continued)

Income quantiles	(1) Tertiles	(2) Quartiles	(3) Quintiles
IVs for the transformed equation	L(0/7).$X^{\dagger}$	L(1/7).X	L(0/2).X
IVs for the level equation	D.X	D.X	D.X
Underidentification test			
Kleibergen-Paap LM stat	44.622	53.216	38.294
p-value	0.085	0.041	0.093
Weak identification test			
Kleibergen-Paap Wald stat	42.94	40.932	43.958
Observations	474	474	474
Number of counties	72	72	72
Time FE	Yes	Yes	Yes

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. $X^{\dagger}$ refers to L.gini, L.gini × Q, and Q

term between the Gini and initial income levels. All results are estimated by sys-GMM with valid and strong instruments.

6.3 Estimates at different sub-national levels

The baseline analyses use household survey data aggregated at the county level. A county is in the lowest administrative hierarchy that has independent educational, fiscal, and juridical systems. We further explore whether the relationship between inequality and growth also holds at other sub-national levels, such as village level and province level.

Using the same CHNS data set, we construct panel data at the community level, i.e. villages in the rural area and neighborhoods in the urban area. Accordingly, the Gini and average household income per capita are calculated using households residing within the communities. The constructed panel data at the community level contains 307 communities, with 1863 observations in total.

In Table 5, we re-estimate Eq. (1) using the community-level panel data. The estimated coefficients on the Gini are positive while the estimated coefficients on the interaction term are negative, which is consistent with county-level results. One should be aware, though, that in the lower sub-national level the number of households residing within the communities is also smaller: Computed community-level Ginis and average incomes are based on fewer observations than the computed county-level Ginis and average incomes.

Compared with column (1) of Table 5, after replacing extreme Gini above the 1st percentile and 5th percentile of the distribution at the two tails with the percentile values, the magnitude of coefficients drops significantly in columns (2) and (3). We prefer estimates in column (3), which are based on strong instruments at the conventional

Table 5 Sys-GMM estimates for a community-level data set (dependent variable: annualized growth rate)

	(1) Baseline	(2) Winsor (1%)	(3) Winsor (5%)
L.gini	7.795**	2.746**	2.733**
	(3.051)	(1.135)	(1.121)
L.gini × L.lny	− 0.891***	− 0.319**	− 0.320**
	(0.343)	(0.126)	(0.124)
L.lny	0.103	− 0.150**	− 0.153***
	(0.154)	(0.059)	(0.058)
AR(2) test	0.935	1.576	1.651
p-value	0.350	0.115	0.099
Hansen test	4.727	8.279	7.946
p-value	0.579	0.218	0.242
Instruments			
Number of IVs	18	18	18
IVs for the transformed equation	L(6/7).X†	L(6/7).X	L(6/7).X
IVs for the level equation	D.X	D.X	D.X
Underidentification test			
Kleibergen-Paap LM stat	13.666	69.806	75.024
p-value	0.057	0.000	0.000
Weak identification test			
Kleibergen-Paap Wald stat	1.203	5.505	10.081
Stock-Yogo critical values			
10% maximal IV relative bias	9.37	9.37	9.37
20% maximal IV relative bias	5.78	5.78	5.78
30% maximal IV relative bias	4.46	4.46	4.46
Observations	1863	1863	1863
Number of communities	307	307	307
Time FE	Yes	Yes	Yes

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. X† refers to L.gini, L.gini × L.lny, and L.lny

10% significance level. The implied threshold of average household income per capita at the village level is around 8.54 logs, at the 40th percentile of the sample villages, above which the Gini has negative effects on the subsequent economic growth while positive below the threshold.

To conduct province-level analyses, we construct measures of inequality and household income per capita using households residing within the provinces. The province-level panel data has a total of 79 observations, covering 12 provinces during 1989–2015. Figures 7 and 8 provide a map of the province-level Gini and household income per capita of Chinese provinces for three selected waves: 1989, 2000, and

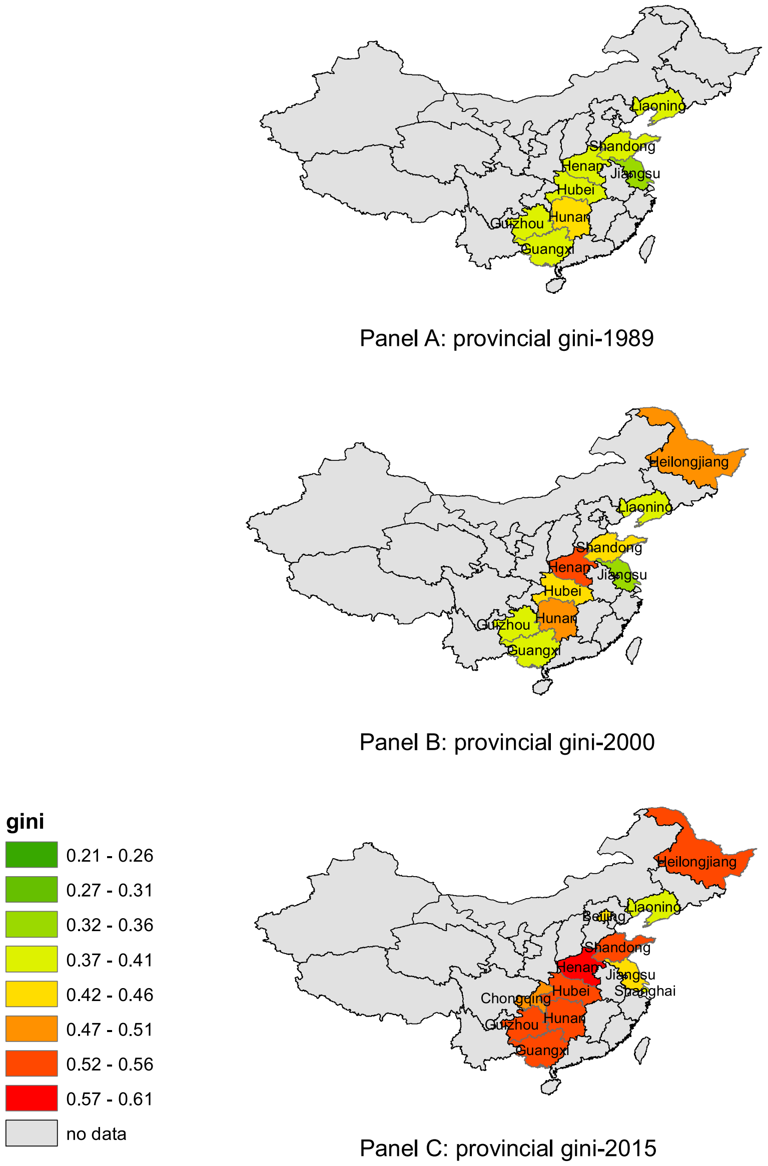

Fig. 7 The ginis of Chinese provinces for three selected survey-years: 1989, 2000, and 2015

2015. From Fig. 7, one can see that inequality increased in all Chinese provinces from 1989 to 2015. And, already in 1989 there were some Chinese provinces with significant inequalities.

Table 6 displays estimates of Eq. (1) using the province-level panel data. Column (1) shows sys-GMM estimates, and column (2) present OLS estimates. The estimated coefficient on the Gini is positive while the estimated coefficient on the interaction

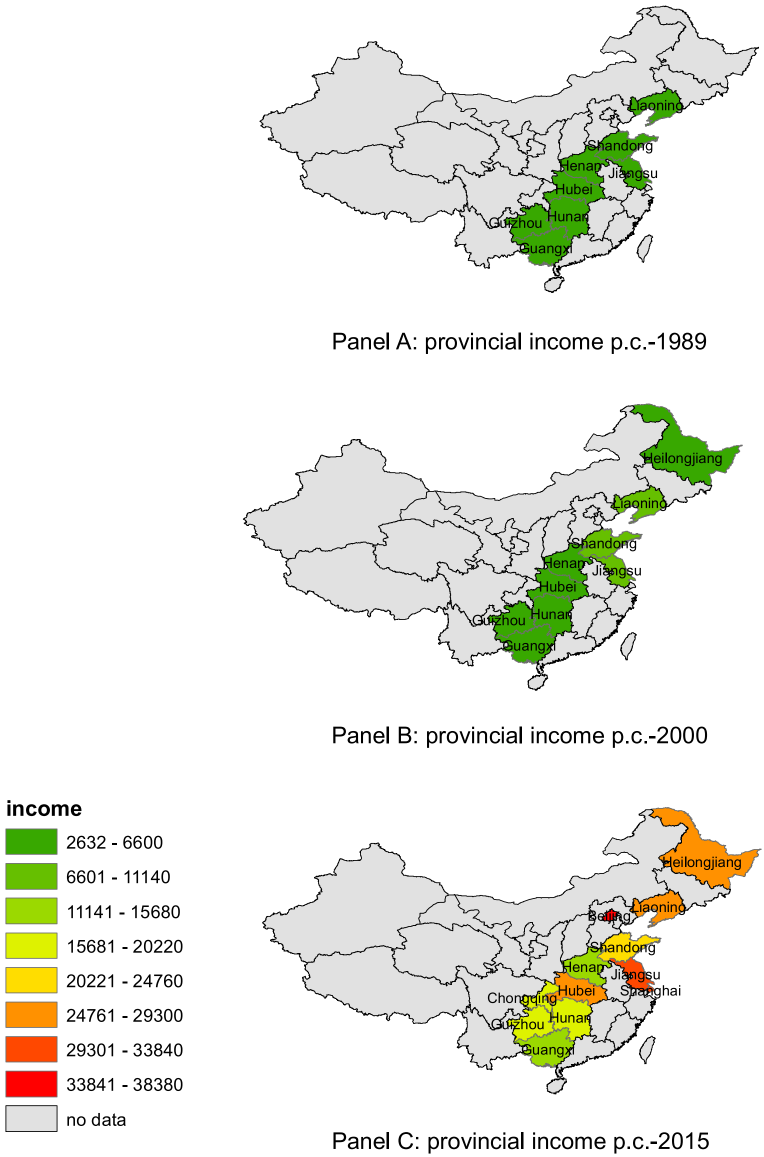

Fig. 8 Average household incomes per capita of Chinese provinces for three selected survey-years: 1989, 2000, and 2015

between the Gini and initial income is negative. Both of these estimated coefficients are significantly different from zero at the 10 per cent level. Quantitatively, the province-level estimates in Table 6 are larger than the estimates obtained from our baseline county-level data set. According to the estimates in column (1) of Table 7, the threshold of average household income per capita above which the effect of the Gini on growth

Table 6 Estimates for a province-level data set (dependent variable: annualized growth rate)

	(1) Sys-GMM	(2) Fe
L.gini	5.977***	7.219***
	(3.168)	(1.599)
L.gini × L.lny	− 0.649***	− 0.774***
	(0.346)	(0.175)
L.lny	0.241***	0.155
	(0.139)	(0.090)
AR(2) test	− 0.612	
p-value	0.540	
Hansen test	3.338	
p-value	0.188	
Instruments		
Number of IVs	5	
IVs for the transformed equation	L(0/0).X†	
IVs for the level equation	D.X	
Underidentification test:		
Kleibergen-Paap LM stat	7.597	
p-value	0.055	
Weak identification test		
Kleibergen-Paap Wald stat	10.632	
Stock-Yogo critical values		
10% maximal IV relative bias	7.56	
20% maximal IV relative bias	5.57	
30% maximal IV relative bias	4.73	
Observations	79	79
Number of provinces	12	12
Time FE	Yes	Yes

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. X† refers to L.gini, L.gini × L.lny

is negative is at about 9.21 logs; below this threshold the effect of the Gini on growth is positive.

6.4 Evidence on mechanisms: educational attainments and skilled workers

According to Galor and Zeira (1993), in the presence of imperfect credit markets, income inequality affects economic growth through investment in human capital. Only those who are rich enough to pay the fixed-costs of education will they invest in education and work as skilled labour later in life. Initial income levels and income

Table 7 Inequality and educational attainment among schooling-age people

Dependent variables	Average schooling years	Share of people by degree attainment			
		Primary	Middle	High	College
Age group	11–25	11–14	14–17	16–20	19–25
L.gini	27.541* (15.082)	0.373 (1.602)	− 1.559 (2.000)	7.547* (4.361)	3.858*** (1.312)
L.gini × L.lny	− 3.149* (1.685)	− 0.050 (0.176)	0.148 (0.227)	− 0.868* (0.495)	− 0.455*** (0.150)
L.lny	1.805** (0.878)	0.031 (0.094)	− 0.004 (0.119)	0.401* (0.215)	0.303*** (0.081)
AR(2) test	1.332	− 2.164	0.101	− 1.454	1.245
p-value	0.183	0.030	0.919	0.146	0.213
Hansen test	23.364	27.339	29.273	24.593	12.163
p-value	0.325	0.289	0.210	0.265	0.204
Instruments					
Number of IVs	33	36	36	33	21
IVs for the transformed equation	L(1/7).X†	L(0/7).X	L(0/7).X	L(1/7).X	L(4/6).X
IVs for the level equation	D.X	D.X	D.X	D.X	D.X
Underidentification test					
Kleibergen-Paap LM stat	46.077	48.467	48.232	47.060	36.766
p-value	0.002	0.003	0.003	0.001	0.000
Weak identification test					
Kleibergen-Paap Wald stat	14.648	13.416	15.125	12.290	9.861
Stock-Yogo critical values					
10% maximal IV relative bias	10.70	10.74	10.74	10.70	10.01
20% maximal IV relative bias	5.91	5.89	5.89	5.91	5.90
30% maximal IV relative bias	4.24	4.20	4.20	4.24	4.42
Observations	480	476	472	469	477
Number of counties	72	72	72	71	72
Time FE	Yes	Yes	Yes	Yes	Yes

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. X† refers to L.gini, L.gini × L.lny, and L.lny distribution affect the proportion of people who can afford for education, and therefore the shares of educational attainment and skilled labour at the aggregate level, which ultimately affect the overall economic growth. For relatively rich counties, an increase in inequality would result in fewer people being educated, leading to lower shares of educational attainment and skilled labour at the aggregate level, and thus slowing down the economic growth. Conversely, for a relatively poor county, an increase in inequality would allow some people to be educated and work as skilled workers, leading to higher rates of educational attainment and skilled labour at the aggregate level, and thus benefiting the economic growth. If these theoretical predictions hold, we should observe that the impacts of inequality on educational attainment rates among schooling-age people, as well as skilled labour rates among working-age people, are decreasing in initial income levels.

Table 7 shows the correlation between inequality and educational attainments of schooling age young people. The dependent variable in column (1) is the average schooling years of people aged between 11 and 25 at the county level. The estimated coefficient on the Gini is positive while the estimated coefficient on the interaction term is negative. The effects of the Gini on the average schooling years are decreasing in initial income, consistent with the prediction by Galor and Zeira (1993). From columns (2)–(5), we separately examine degree attainment rates among the relevant age groups: primary school (11–14 years), middle school (14–17 years), high school (16–20 years), and college (19–25 years).[9] The coefficients in columns (2) and (3) are not significant, which may not violate our findings, given that primary school and middle school are mandated by law.[10] Turning to the impacts of inequality on the share of high school degree attainments in column (4), the same pattern shows up again: the estimated coefficient before the Gini is positive while the estimated coefficient before the interaction term is negative. In column (5) estimated coefficients are qualitatively the same where the share of the college degree attainment is the dependent variable. Consider a county with an initial income equal to the sample mean of Chinese counties' household income per capita in the year 2015 (equivalent to about 10.00 logs), a one percentage increase in the Gini decrease the share of high school attainments by 1.13 percentage points, the share of college attainment by 0.69 percentage points.

Table 8 examines the effects of inequality on the share of skilled workers[11] at the

[9] In general, children start to attend primary school by the age of 7, and the common schooling years for completing primary school is 6 years; middle school, 3 years; technical/high school, 2–3 years; and college/university, 3–5 years. Children might start primary school between 5 to 8 years old, and the corresponding ages they obtain the primary school degree are between 11 to 14 years old, and middle school degree around 14 to 17 years old. After middle school, part of children might attend vocational/technical schools which take 2–3 years, and part of children attend high school. In Table 7, high school degree refers to vocational/technical and high school degrees. The related age group that attains the high school degree is between 16–20 years old. College in Table 7 refers to college or university degrees, and the related age group that attains the college degree is between 19–25 years old.

[10] In 1986, China passed a law requiring a nine-year compulsory schooling attendance (six years of primary school and three years of middle school).

Table 8 Inequality and skilled workers among working-age people (Dependent Variable: share of skilled workers)

Age group	(1) 26–55	(2) 26–35	(3) 36–45	(4) 46–55
L.gini	4.675*** (1.688)	4.620*** (1.353)	4.383** (2.188)	2.632 (2.016)
L.gini × L.lny	− 0.519*** (0.188)	− 0.533*** (0.158)	− 0.500** (0.241)	− 0.278 (0.222)
L.lny	0.297*** (0.088)	0.260*** (0.082)	0.316*** (0.118)	0.206* (0.107)
AR(2) test	0.272	− 0.614	− 0.309	0.777
p-value	0.785	0.539	0.757	0.437
Hansen test	27.205	29.364	27.849	8.256
p-value	0.164	0.207	0.144	0.509
Instruments				
Number of IVs	33	36	33	21
IVs for the transformed equation	L(1/7).X†	L(0/7).X	L(1/7).X	L(5/7).X
IVs for the level equation	D.X	D.X	D.X	D.X
Underidentification test				
Kleibergen-Paap LM stat	46.077	48.216	46.077	32.778
p-value	0.002	0.004	0.002	0.000
Weak identification test				
Kleibergen-Paap Wald stat	14.648	19.951	14.648	10.162
Stock-Yogo critical values				
10% maximal IV relative bias	10.70	10.74	10.70	10.01
20% maximal IV relative bias	5.91	5.89	5.91	5.90
30% maximal IV relative bias	4.24	4.20	4.24	4.42
Observations	480	477	480	480
Number of counties	72	72	72	72
Time FE	Yes	Yes	Yes	Yes

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. X† refers to L.gini, L.gini × L.lny, and L.lny

county level when individuals are at “prime age”, which we define as being between the ages of 26 and 55 years old. Column (1) reports estimates of the Gini on the share of skilled workers over the entire prime age. The estimated coefficient on the

[11] Skilled workers are defined according to their primary occupations. There are 12 types of occupations in total collected by the CHNS survey: senior professional/technical worker; junior professional/technical worker; administrator/executive/manager; office staff; farmer, fisherman, and hunter; skilled worker; army officer and police officer; ordinary soldier and policeman; driver; service worker; athlete, actor, and musician. In this paper, workers are defined as skilled if their primary occupations are senior professional/technical worker, junior professional/technical worker, administrator/executive/manager, skilled worker, athlete, actor, and musician.

Gini is significantly positive and the estimated coefficient on the interaction term is significantly negative. Consider, again as in the discussion of educational attainments, a county with an initial income equal to the sample mean for the year 2015 (10.00 logs), a one percentage increase in the Gini decreases the share of skilled workers by 0.5 percentage points. From column (2) to column (4), we separately estimate the impacts of inequality on the ratio of skilled workers for different age groups. The heterogeneous effects of inequality on the share of skilled workers by initial incomes hold along the life-cycle, but the marginal impacts of the Gini decline. Quantitatively, for a county with an initial income of 10.00 logs, a one percentage increase in the Gini decreases the share of skilled workers by 0.72, 0.62, and 0.15 percentage points for individuals aged between 26 to 35, 36 to 45, and 46 to 55 years old, respectively.

We argue that the results in Tables 7 and 8 are unlikely to be confounded by the general educational expansion or structural changes during the examined period. The nationwide improvements in education or skilled labour are captured by time fixed effects. Moreover, both tables reveal that in recent years, as counties become relatively rich, the rise in the Gini has detrimental impacts on the share of people with high school and above degree, as well as the share of skilled labour, which is at odds with the rising demand for human capital alongside the economic development.

7 Robustness checks

7.1 Sample attrition

Sample attrition, or the loss to follow-up, in the CHNS may be caused by migration, death, or refusal to attend surveys, according to Popkin et al. (2010). The averaged retention rate at the county level of the CHNS is illustrated in Fig. 9, where the retention rate is defined as the ratio of households who participate in the previous survey remaining in the current survey. The average retention rate is 0.95 in 1991 (attrition rate around 5%) and 0.76 in 2015 (attrition rate around 24%).

Selective sample attrition is a common concern when using panel survey data to do analyses. The correlation between initial inequality and subsequent economic growth would be spurious if sample attrition is closely related with initial inequality. More specifically, if poor (rich) households in a county with high inequality are more likely to migrate out, the average household income per capita in the county would be higher (lower) among the rest residents in the following period. Therefore, selective migration would lead to the illusory correlation that high inequality is related with higher (lower) subsequent economic growth.

To prove our findings are not confounded by selective sample attrition, we examine whether the sample attrition is correlated with initial inequality or income levels. In column (1) of Table 9, the estimated coefficients before the initial Gini, initial incomes, and the interaction term between the initial Gini and initial incomes are small and do not significantly differ from 0. We do not find evidence that the sample attrition is selective on the initial inequality or income levels. In column (2) of Table 9, we re-estimate Eq. (1) while controlling attrition rates on the right-hand side of the regression. The estimated coefficients on the Gini and the interaction term of the Gini and the initial

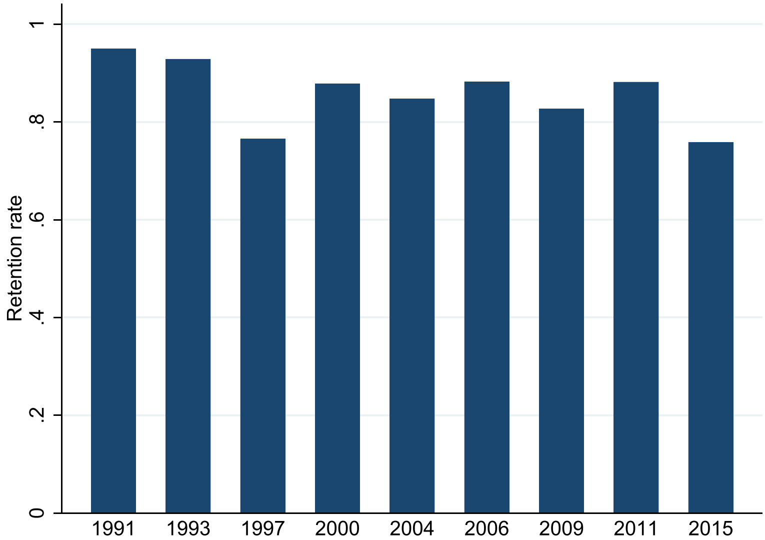

Fig. 9 Mean county retention rate. The retention rate is calculated based on households that participate in previous survey round remaining in the current survey at the county level. The numbers in this figure are the averaged retention rate at the county level

incomes barely have any changes, compared with the baseline estimates in the column (1) of Table 2. Therefore, we argue that our results are not likely to be biased by the sample attrition of CHNS.

7.2 Alternative measures of inequality

Table 10 presents sys-GMM estimates using other alternative measures of inequality, that is, mean log deviation, 75–25 (90–10) income ratio, and the income shares of the 1st, 2nd, 3rd, 4th, and 5th quintile. For comparison purposes, the estimates in column (1) of Table 10 are the same as in column (1) of Table 2 where the Gini is used as a measure of inequality.

From columns (2)–(4), inequality is measured by mean log deviation, 75–25 income ratio, and 90–10 income ratio respectively. The estimated coefficients on the inequality measures are significantly positive while the estimated coefficients on the interaction term are significantly negative. This is qualitatively the same result as what we obtained for the Gini. Since these measures of inequality have different ranges, we therefore discuss quantitative implications of estimates for the case of a one standard deviation increase in the respective inequality measure. Consider a county with an initial household income per capita equal to 10.0 logs, a one standard deviation (0.10) increase in the Gini decreases the per annum growth rate by 4.5 percentage points (column (1)); a one standard deviation (0.19) increase in the MLD decreases the per annum growth rate by 2.8 percentage points (column (2)); a one standard deviation (9.2) increase in the 90–10 ratio decreases the per annum growth rate by 3.7 percentage points (column

Table 9 Sample attrition

Dependent variable	(1) Attrition rate	(1) Annualized growth rate
L.gini	− 0.444 (1.675)	2.953** (1.167)
L.gini × L.lny	0.054 (0.187)	− 0.334** (0.132)
L.lny	− 0.032 (0.101)	− 0.084 (0.078)
Attrition		− 0.084 (0.052)
AR(2) test	0.738	0.128
p-value	0.461	0.898
Hansen test	24.907	38.408
p-value	0.411	0.202
Instruments		
Number of IVs	36	45
IVs for the transformed equation	L(0/7).X†	L(0/7).X
IVs for the level equation	D.X	D.X
Underidentification test		
Kleibergen-Paap LM stat	47.855	50.106
p-value	0.004	0.029
Weak identification test		
Kleibergen-Paap Wald stat	20.711	21.967
Observations	480	474
Number of counties	72	72
Time FE	Yes	Yes

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. X† refers to L.gini, L.gini × L.lny, and L.lny

(3)); and a one standard deviation (1.1) increase in the 75–25 ratio decreases the per annum growth rate by 3.5 percentage points (column (4)).

Estimates of the relationship between the income shares held by the different quintiles along the income distribution and household income growth are shown in columns (5)-(9) of Table 10. For the lower-tail quintiles, the estimated coefficients on the income shares of the first (Q1) and second (Q2) quintiles are significantly negative while the estimated coefficients on the interaction of the first and second quintile income shares with initial average income are significantly positive. For the upper-tail quintiles, oppositely, the estimated coefficients on the income shares of the fourth (Q4) and fifth (Q5)

Table 10 Sys-GMM estimates for alternative measures of inequality (dependent variable: annualized growth rate)

Inequality	(1) Gini	(2) Mld	(3) 90–10	(4) 75–25	(5) Q1	(6) Q2	(7) Q3	(8) Q4	(9) Q5
L.inequality	3.261***	1.127***	0.026*	0.169**	− 11.131**	− 12.417**	− 4.897	7.187	2.681*
	(1.018)	(0.357)	(0.014)	(0.071)	(5.161)	(4.694)	(3.705)	(6.318)	(1.405)
L.inequality × L.lny	− 0.374***	− 0.127***	− 0.003*	− 0.020**	1.294**	1.417**	0.564	− 0.767	− 0.306*
	(0.115)	(0.041)	(0.002)	(0.008)	(0.599)	(0.536)	(0.423)	(0.695)	(0.160)
L.lny	− 0.066	− 0.167***	− 0.192***	− 0.159***	− 0.246***	− 0.364***	− 0.280***	− 0.059	− 0.071
	(0.066)	(0.038)	(0.034)	(0.042)	(0.042)	(0.064)	(0.069)	(0.151)	(0.091)
AR(2) test	0.049	0.029	0.478	0.335	0.476	0.287	0.471	0.467	0.233
p-value	0.961	0.977	0.632	0.737	0.634	0.774	0.638	0.641	0.816
Hansen test	27.949	27.433	30.597	29.096	31.337	29.572	30.428	15.765	28.644
p-value	0.262	0.285	0.166	0.217	0.144	0.101	0.171	0.202	0.234
Instruments									
Number of IVs	36	36	36	36	36	33	36	24	36
IVs for the transformed equation	L(0/7).X†	L(0/7).X	L(0/7).X	L(0/7).X	L(0/7).X	L(1/7).X	L(0/7).X	L(4/7).X	L(0/7).X
IVs for the level equation	D.X	D.X	D.X	D.X	D.X	D.X	D.X	D.X	D.X
Underidentification test									
Kleibergen-Paap LM stat	47.974	48.977	49.590	46.116	42.262	46.813	47.320	39.147	47.495
p-value	0.004	0.003	0.002	0.006	0.017	0.002	0.004	0.000	0.004
Weak identification test									
Kleibergen-Paap Wald stat	40.115	17.860	31.387	19.038	16.733	13.224	61.806	15.466	57.767
Stock-Yogo critical values									

Table 10 (continued)

Inequality	(1) Gini	(2) Mld	(3) 90–10	(4) 75–25	(5) Q1	(6) Q2	(7) Q3	(8) Q4	(9) Q5
10% maximal IV relative bias	10.74	10.74	10.74	10.74	10.74	10.70	10.74	10.33	10.74
20% maximal IV relative bias	5.89	5.89	5.89	5.89	5.89	5.91	5.89	5.94	5.89
30% maximal IV relative bias	4.24	4.24	4.24	4.24	4.24	4.24	4.24	4.37	4.24
Observations	474	474	474	474	474	474	474	474	474
Number of counties	72	72	72	72	72	72	72	72	72
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. $X^{\dagger}$ refers to L.inequality, L.inequality $\times$ L.lny, and L.lny quintiles are positive while the estimated coefficients on the interaction terms are negative. Quantitatively, consider a county with an initial income equal to 10.00 logs, the estimates in columns (5)–(9) of Table 10 imply that: a one standard deviation (0.034) increase in the income share of the first quintile increases the per annum growth rate by 6.1 percentage points; a one standard deviation (0.024) increase in the income share of the second quintile increases the per annum growth rate by 4.2 percentage points; a one standard deviation (0.023) increase in the income share of the third quintile increases the per annum growth rate by around 1.7 percentage points; a one standard deviation (0.023) increase in the income share of the fourth quintile decreases the per annum growth rate by 1.1 percentage points; a one standard deviation (0.078) increase in the income share of the fifth quintile decreases the per annum growth rate by 3.0 percentage points.

7.3 Alternative measures of income

In Table 11 we report estimates of the relationship between gross inequality and gross average incomes of Chinese counties, again for various measures of gross inequality. For the estimates reported in Table 11 we use household gross income per capita to calculate county-level inequality indices and growth rates of household income per capita. In China, for people who live in urban areas gross income is typically equal to net income, since the wage is usually the only income source. But, in rural areas, many people also make an income from farming, fishing, gardening and raising livestock, etc., which often requires input investments. Net income generally is smaller than gross income.

Column (1) of Table 11 shows that the estimated coefficient on the gross Gini is positive and significantly different from zero. The interaction between the gross Gini and initial gross household income per capita is negative and also significantly different from zero at the conventional significance levels. This suggests that gross inequality has a positive effect on the growth rate of gross household income per capita at low levels of development while at high levels of development the effect is negative. A similar result, though statistically somewhat weak, is obtained for other measures of gross inequality (see columns (2)–(9) of Table 11).

Further results and discussions are available in the online appendix, where we use alternative estimation method (OLS estimates) and model specifications (without the interaction term of inequality and income/adding a quadratic term of inequality). In addition, we also discuss sys-GMM estimates using the panel data of Benjamin et al. (2011).

Table 11 Sys-GMM Estimates for Alternative Measure of Income (Dependent Variable: annualized growth rate)

Inequality	(1) Gini	(2) Mld	(3) 90–10	(4) 75–25	(5) Q1	(6) Q2	(7) Q3	(8) Q4	(9) Q5
L.inequality	2.069*	0.940	0.016	0.177**	− 13.639**	− 8.549*	− 4.734	− 0.097	1.954
	(1.195)	(0.688)	(0.022)	(0.083)	(5.413)	(4.393)	(4.614)	(4.249)	(1.379)
L.inequality × L.lny	− 0.237*	− 0.106	− 0.002	− 0.021**	1.591**	1.013**	0.557	0.067	- 0.226
	(0.131)	(0.075)	(0.003)	(0.010)	(0.606)	(0.501)	(0.517)	(0.479)	(0.152)
L.lny	− 0.095	− 0.152***	− 0.176***	− 0.142***	− 0.271***	− 0.288***	− 0.248***	− 0.183*	- 0.077
	(0.075)	(0.052)	(0.045)	(0.044)	(0.050)	(0.070)	(0.081)	(0.101)	(0.089)
AR(2) test	0.447	0.459	0.586	0.280	0.622	0.424	0.494	0.296	0.478
p-value	0.655	0.646	0.558	0.779	0.534	0.671	0.621	0.767	0.633
Hansen test	30.888	31.925	33.028	26.711	28.033	30.216	30.509	30.206	28.997
p-value	0.157	0.129	0.103	0.318	0.259	0.178	0.168	0.178	0.220
Instruments									
Number of IVs	36	36	36	36	36	36	36	36	36
IVs for the transformed equation	L(0/7).X[†]	L(0/7).X	L(0/7).X	L(0/7).X	L(0/7).X	L(0/7).X	L(0/7).X	L(0/7).X	L(0/7).X
IVs for the level equation	D.X	D.X	D.X	D.X	D.X	D.X	D.X	D.X	D.X
Underidentification test									
Kleibergen-Paap LM stat	46.526	48.247	49.410	46.464	44.972	44.392	47.270	46.468	46.938
p-value	0.006	0.003	0.003	0.006	0.008	0.010	0.005	0.006	0.005
Weak identification test									

Table 11 (continued)

Inequality	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Gini	Mld	90–10	75–25	Q1	Q2	Q3	Q4	Q5
Kleibergen-Paap Wald stat	28.869	30.375	42.684	23.325	31.246	38.099	45.297	60.095	27.681
Stock-Yogo critical values									
10% maximal IV relative bias	10.74	10.74	10.74	10.74	10.74	10.74	10.74	10.74	10.74
20% maximal IV relative bias	5.89	5.89	5.89	5.89	5.89	5.89	5.89	5.89	5.89
30% maximal IV relative bias	4.24	4.24	4.24	4.24	4.24	4.24	4.24	4.24	4.24
Observations	474	474	474	474	474	474	474	474	474
Number of counties	72	72	72	72	72	72	72	72	72
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Huber robust standard errors are shown in parentheses, where * indicates statistically different from zero at the 10% significant level, ** 5% significant level, *** 1% significant level. $X^{\dagger}$ refers to L.inequality, L.inequality × L.lny, and L.lny. The county-level inequality indices and growth rates are obtained using household gross income per capita

8 Conclusion

The rhetoric of China's political leaders with regard to income distribution has changed remarkably over the past four decades. Statements by Deng Xiaoping (1985) suggest that, at the early stage of China's economic development, income inequality was tolerated–and perhaps even encouraged by the then political leader of China. In contrast, statements by the current political leader, Xi Jinping (2015), suggest that, nowadays, there is a political desire for income inequality in China to be reduced. The different rhetoric is consistent with both Deng Xiaoping and Xi Jinping having GDP growth of China as a primary objective: our panel model estimates showed that for the relatively low levels of average incomes that were prevalent in China at the time when Deng Xiaoping was China's political leader, inequality had a positive effect on economic growth; for current levels of average income, inequality has a negative effect on growth.

The theoretical model by Galor and Zeira (1993) predict that inequality has a positive effect on economic growth at an early stage of development when average income is low, while at a more advanced stage, when average income is high, inequality reduces growth. The main contribution of our paper was to provide estimates, specifically, for Chinese counties, of the effect that inequality has on economic growth for various levels of initial income. Based on CHNS data, we constructed measures of inequality and household income per capita for a panel of 72 Chinese counties that spans nearly three decades, from the late 1980s to the mid-2010s. Our sys-GMM estimates showed that for an average Chinese county in the year 1989 a one percentage point increase in the Gini increased the per annum growth rate by around 0.3 percentage points. For the poorest county in the year 1989 the effect is even larger amounting to around 0.5 percentage points. For the richest county in the year 1989 the effect of inequality on growth was also positive, though smaller, amounting to around 0.1 percentage points. Hence, in the year 1989 for all levels of average incomes of Chinese counties inequality had a positive effect on growth.

We come back to the statement by Deng Xiaoping (1985): "Let some people get rich first". This statement by Deng Xiaoping was highly appropriate at the time when it was made with regard to the primary objective of maximizing economic growth. According to our panel model estimates, the increase in income inequality that occurred in China during the 1980s had unambigously a positive effect on growth. And, to the extent that there was an intended legacy effect: the increase in inequality that occurred during the 1990s had also very likely a positive effect on economic growth in China.

But it's different for current times: The prediction from our econometric model estimates is that for an average income equal to the GDP per capita of China in the year 2021, a 1 percentage point increase in the Gini would reduce the per annum growth rate by around 1 percentage point. At the current stage of development that China is at inequality has a negative effect on economic growth. Now consider again the statement by Xi Jinping (2015): "We must ensure that the fruits of development benefit all people." This statement is consistent with Xi Jinping, too, having as a primary objective the maximization of China's economic growth. The statement by Xi Jinping is highly appropriate given the current stage of development that China is at.

The relationship between inequality and economic growth found in this paper is supported by further analyses on the underlying mechanism through human capital investment. We show that the impacts of inequality on the share of educational attainment as well as the share of skilled labour are decreasing in initial income levels. Overall, our empirical evidences are consistent with the theoretical predictions by Galor and Zeira (1993).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00181-023-02472-0.

Acknowledgements The authors are grateful to anonymous referees and editor for thoughtful suggestions that significantly improved the paper. The authors also appreciate helpful comments from participants of various conferences and seminars.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Appendix

See Figs. 10, 11, 12.

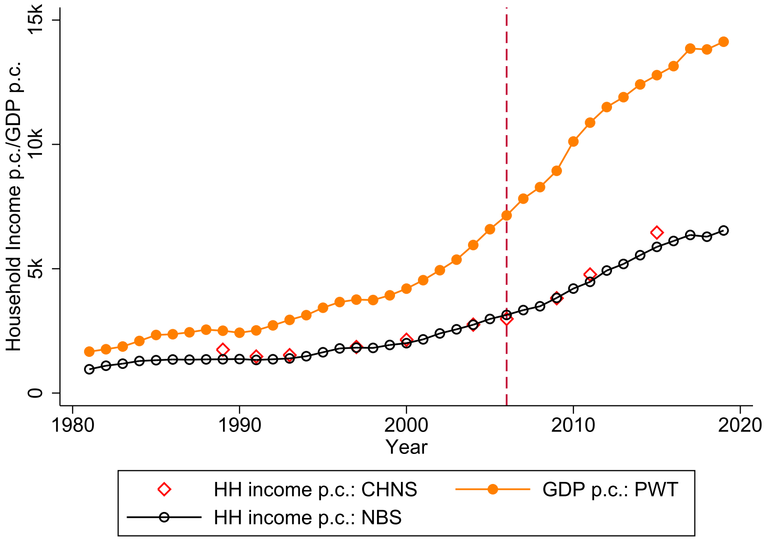

Fig. 10 Household income per capita vs. GDP per capita in China. Household income per capita is in constant prices and converted into PPP USD using data from the PWT on the USD-Yuan nominal exchange rate and the PPP for China relative to the US. GDP per capita is in constant price PPP USD

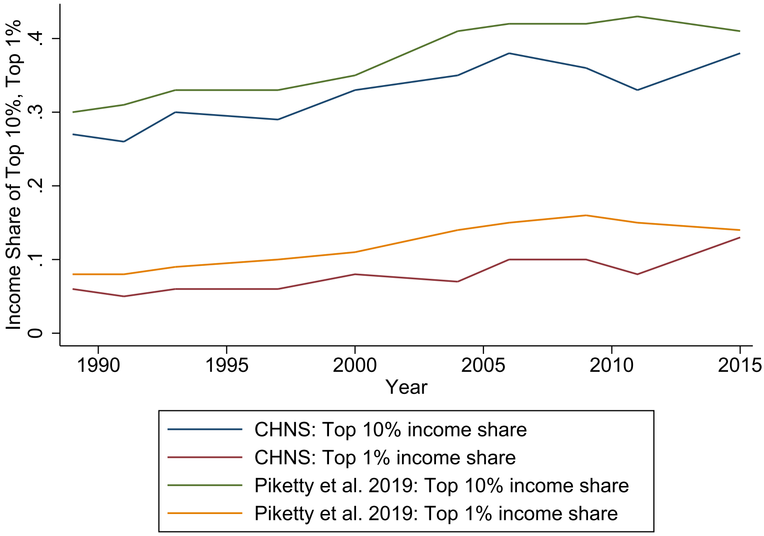

Fig. 11 Income shares in China of the top 10% and Top 1%: CHNS vs. Piketty et al. (2019) data

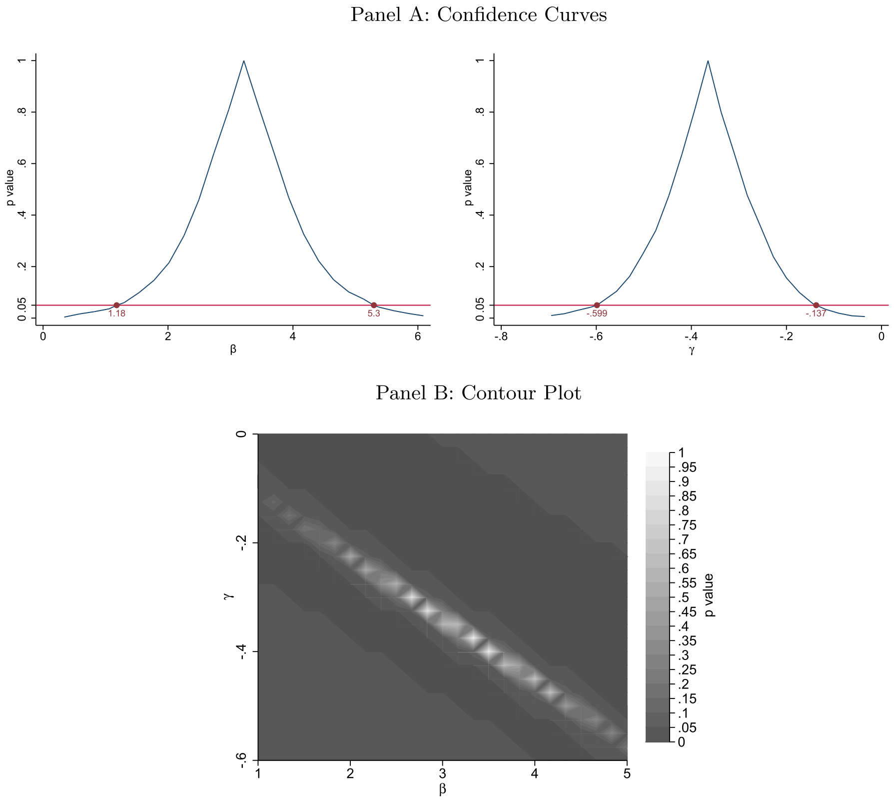

Fig. 12 Wild restricted efficient bootstrap. The confidence curves and the contour plot were generated using STATA's *boottest* command. We used a wild restricted efficient bootstrap (with 1000 draws clustered at the county level) over the t-statistics of the one-step sys-GMM estimates in column (4) of Table 2. The auxiliary random variable for the bootstrapping was drawn from a Rademacher distribution. β refers to the estimated coefficient on L. Gini; γ refers to the estimated coefficient on L.gini $\times$ L.lny

References

Bazzi S, Clemens MA (2013) Blunt instruments: avoiding common pitfalls in identifying the causes of economic growth. Am Econ J Macroecon 5(2):152–186

Benjamin D, Brandt L, Giles J (2011) Did higher inequality impede growth in rural China? Econ J 121(557):1281–1309

Berg A, Ostry JD, Tsangarides CG, Yakhshilikov Y (2018) Redistribution, inequality, and growth: new evidence. J Econ Growth 23(3):259–305

Brueckner M, Lederman D (2018) Inequality and economic growth: the role of initial income. J Econ Growth 23(3):341–366

Chen CN, Tsaur TW, Rhai TS (1982) The gini coefficient and negative income. Oxford Econ Papers 34(3):473–478

Deng XP (1985) Quote extracted from web http://cpc.people.com.cn/GB/34136/2569304.html.

Easterly W (2007) Inequality does cause underdevelopment: insights from a new instrument. J Dev Econ 84:755–776

Erman L, te Kaat DM (2019) Inequality and growth: industry-level evidence. J Econ Growth 24(3):283–308

Feenstra RC, Inklaar R, Timmer MP (2015) The next generation of the penn world table. Am Econ Rev 105(10):3150–3182
Forbes KJ (2000) A reassessment of the relationship between inequality and growth. Am Econ Rev 90(4):869–887
Galor O, Zeira J (1993) Income distribution and macroeconomics. Rev Econ Stud 60(1):35–52
Galor O, Moav O, Vollrath D (2009) Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. Rev Econ Stud 76(1):143–179
Halter D, Oechslin M, Zweimüller J (2014) Inequality and growth: the neglected time dimension. J Econ Growth 19(1):81–104
Kanbur R, Wang Y, Zhang XB (2021) The great Chinese inequality turnaround. J Comp Econ 49:467–482
Kraay A (2015) Weak instruments in growth regressions: implications for recent cross-country evidence on inequality and growth. World bank policy research working paper No.7494.
Kripfganz S (2019) Generalized method of moments estimation of linear dynamic panel data models. London Stata Conference 2019 17, Stata Users Group
Kuznets S (1955) Economic growth and income inequality. Am Econ Rev 45:1–28
Litschig S, Lombadi M (2019) Which tail matters? Inequality and growth in Brazil. J Econ Growth 24:155–187
Luo CL, Li S, Sicular T (2020) The long-term evolution of national income inequality and rural poverty in China. China Econ Rev 62:101465
Olea M, Pflueger C (2013) A robust test for weak instruments. J Bus Econ Statist 31:358–369
Panizza U (2002) Income inequality and economic growth: evidence from American data. J Econ Growth 7(1):25–41
Piketty T, Yang L, Zucman G (2019) Capital accumulation, private property, and rising inequality in china, 1978–2015. Am Econ Rev 109(7):2469–2496
Popkin BM, Du S, Zhai F, Zhang B (2010) Cohort profile: the china health and nutrition survey-monitoring and understanding socio-economic and health change in China, 1989–2011. Int J Epidemiol 39(6):1435–1440
Ravallion M, Chen SH (2021) Is that really a Kuznets curve? Turning points for income inequality in China. NBER working paper 29199
Ravallion M, Chen SH (2007) China's (uneven) progress against poverty. J Dev Econ 82(1):1–42
Roodman D (2009) How to do xtabond2: an introduction to difference and system gmm in stata. Stata J 9(1):86–136
Roodman D, Nielsen MØ, MackKinnon JG, Webb MD (2019) Fast and wild: bootstrap Inference in Stata using Boottest. Stata J 19(1):4–60
Solt F (2020) Measuring income inequality across countries and over time: the standardized world income inequality database. Soc Sci Q 101(3):1183–1199
Stock J, Yogo M (2005) Testing for weak instruments in linear IV regression, 80–108. Cambridge University Press, New York
UNU-WIDER (2021) World Income inequality database (WIID). Version 31 May. https://doi.org/10.35188/UNU-WIDER/WIID-310521
Windmeijer F (2005) A finite sample correction for the variance of linear efficient two-step gmm estimators. J Econ 126(1):25–51
Xi JP (2015) Quote extracted from web http://theory.people.com.cn/n1/2016/0804/c376186-28610221.html
Young A (2021) Leverage, Heteroskedasticity and instrumental variables in practical application. Mimeo, LSE