

## RESEARCH ARTICLE

# Urbanization and Income Inequality in Post-Reform China: A Causal Analysis Based on Time Series Data

Guo Chen<sup>1\*</sup>, Amy K. Glasmeier<sup>2</sup>, Min Zhang<sup>3</sup>, Yang Shao<sup>4</sup>

**1** Department of Geography, Environment, and Spatial Sciences & Global Urban Studies Program, Michigan State University, East Lansing, Michigan, United States of America, **2** Department of Urban Studies and Planning, MIT, Cambridge, Massachusetts, United States of America, **3** School of Architecture and Urban Planning, Nanjing University, Nanjing, Jiangsu, China, **4** Geography Department, College of Natural Resources and Environment, Virginia Polytechnic Institute and State University, Blacksburg, VA, United States of America

\* [guochen@msu.edu](mailto:guochen@msu.edu)



CrossMark  
click for updates

## OPEN ACCESS

**Citation:** Chen G, Glasmeier AK, Zhang M, Shao Y (2016) Urbanization and Income Inequality in Post-Reform China: A Causal Analysis Based on Time Series Data. PLOS ONE 11(7): e0158826. doi:10.1371/journal.pone.0158826

**Editor:** Cathy W.S. Chen, Feng Chia University, TAIWAN

**Received:** September 29, 2015

**Accepted:** June 22, 2016

**Published:** July 19, 2016

**Copyright:** © 2016 Chen et al. This is an open access article distributed under the terms of the

[Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files. Data are also accessible through the National Bureau of Statistics of China (<http://www.stats.gov.cn/english/>) and the China Data Center (<http://chinadatacenter.org>).

**Funding:** The authors have no support or funding to report.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

This paper investigates the potential causal relationship(s) between China's urbanization and income inequality since the start of the economic reform. Based on the economic theory of urbanization and income distribution, we analyze the annual time series of China's urbanization rate and Gini index from 1978 to 2014. The results show that urbanization has an immediate alleviating effect on income inequality, as indicated by the negative relationship between the two time series at the same year (lag = 0). However, urbanization also seems to have a lagged aggravating effect on income inequality, as indicated by positive relationship between urbanization and the Gini index series at lag 1. Although the link between urbanization and income inequality is not surprising, the lagged aggravating effect of urbanization on the Gini index challenges the popular belief that urbanization in post-reform China generally helps reduce income inequality. At deeper levels, our results suggest an urgent need to focus on the social dimension of urbanization as China transitions to the next stage of modernization. Comprehensive social reforms must be prioritized to avoid a long-term economic dichotomy and permanent social segregation.

## Introduction

Rapid urbanization and rising inequality are two main characteristics of China's development over the past three plus decades. Since the start of the open and reform policy in the late 1970s, the number of urban inhabitants in China has ballooned from 172 million (or 18% of the total population) in 1978 to 749 million (or 55% of the total population) in 2014 [1]. The massive urban transition has not only fueled China's burgeoning economy but also helped modernize Chinese society as a whole. In parallel to the country's rapid urbanization is widening income inequality. Although the post-reform economic growth has been praised for lifting millions of (rural) Chinese out of absolute poverty [2–4], it has also contributed to income stratification

due to uneven developmental policies [5–8] and the persistent disparity in the distribution system [8]. According to official statistics, China's income Gini index increased from 0.3 in 1978 to around 0.5 in 2014 [1], and a recent survey estimates that the national wealth Gini index could be as high as 0.73 [9].

Whether rising urbanization and inequality are causally related becomes a natural question given their co-occurrence. Classical developmental theories, e.g., those popularized by W. A. Lewis and by S. Kuznets, consider urbanization a key step in reshaping emerging economies dichotomized by a subsistence rural sector and an industrializing urban sector. The rural-to-urban population shift is an aspect of the processes behind the famous inverse-U-shaped curve hypothesized by Kuznets, who argues that as more people move from the lower-income rural sector to the higher-income urban sector, the overall income inequality will first increase and then decrease [10]. It is believed that China has already passed the turning point and, therefore, urbanization should help alleviate income inequality [11]. This assessment seems to be consistent with the fact that the overall income inequality in China can be attributed in large part to the rural–urban income gap [12], and urbanization can potentially reduce the impact of the gap (if not narrow it) by absorbing more people into the higher-income urban sector [13].

However, to date this theoretical assessment has only been empirically tested in a limited way. After all, co-occurrence (and correlation) does not equal causation and no previous studies have empirically established or refuted the causal relationship between urbanization and income inequality in China. In this paper, we address this blank by analyzing the annual time-series of China's urbanization rate and Gini indexes since 1978. We adopt the inference framework of predictive causality widely used in econometrics and employ autoregressive methods to model the potential causal relationships. The purpose is to establish a parsimonious yet predictively verifiable baseline causal link between urbanization and income inequality, which offers important insights for China's urban studies and policy-making. The implications of our results are then discussed in the context of post-reform China.

## Theoretical Framework

### Causal Effect of Urbanization on Income Inequality

We measure income inequality using the popular Gini index, which is mathematically based on the Lorenz curve [14]. Suppose the cumulative income-to-population distribution function of a country is  $F(t)$ , where  $t$  is a random variable of personal (or household) income and  $F(t)$  is the share of population (or households) earning less than or equal to  $t$ . The Lorenz curve is depicted in Fig 1, where the vertical axis  $L$  denotes the cumulative income of the population  $F(t)$ . The Gini index is the area between the Lorenz curve and the straight line  $L = F(t)$  (the shaded gray area in Fig 1) multiplied by 2.

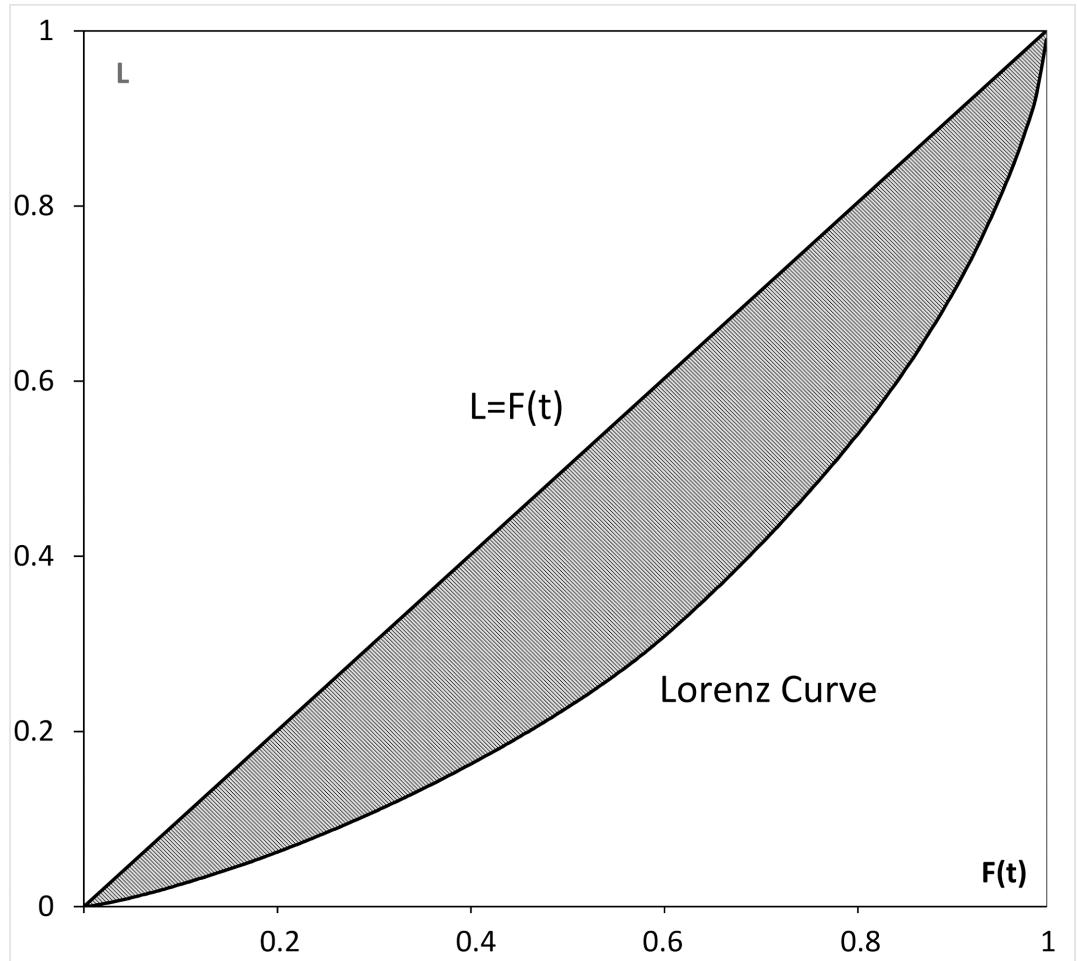
First, we follow steps illustrated in previous studies [15–17] to derive the function from the income variable  $t$  to the Gini index  $G$ . Based on the definition of the Lorenz curve, we have

$$L(F(x)) = \frac{1}{V} \int_{t=-\infty}^x t f(t) dt = \frac{1}{V} \int_{t=-\infty}^x t dF(t) \quad (1)$$

and

$$dL = \frac{1}{V} t dF \quad (2)$$

where  $V$  denotes the average income and  $f(t)$  denotes the income probability density function.



**Fig 1. Lorenz curve and the Gini index.** The value of the Gini index equals the area of the shaded gray region.

doi:10.1371/journal.pone.0158826.g001

Let  $B$  denote the area between the Lorenz curve and the horizontal axis such that we have

$$B = \int_{F=0}^1 L dF \quad (3)$$

By integrating by parts and replacing the integral bounds, we have

$$B = [LF]_{F=0}^1 - \int_{L=0}^1 F dL = 1 - \int_{L=0}^1 F dL \quad (4)$$

After substituting (2) into (4), changing the integral bounds to those of  $t$ , and integrating by parts, we have

$$B = 1 - \frac{1}{V} \int_{t=-\infty}^{+\infty} tF(t)dF(t) = 1 - \frac{1}{2V} \{ [tF^2(t)]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} F^2(t) dt \} \quad (5)$$

To further simplify (5), it is reasonable to assume that in reality the income variable  $t$  is within the range  $[0, T]$ , where  $T$  denotes the highest income. Then, we can rewrite (5) as

$$B = 1 - \frac{1}{2V} \left\{ [tF^2(t)]_0^T - \int_0^T F^2(t) dt \right\} \quad (6)$$

Based on the definition of  $F(t)$ , we have  $F(0) = N_0/N$ , where  $N$  denotes the total population size and  $N_0$  is the number of people (households) with zero income. As  $N_0$  is usually small relative to  $N$ , we can further assume that  $F(0) \rightarrow 0$  when  $N$  is large. Similarly, we can assume that  $F(T)$  approximates 1 when  $N$  is large. Then, (6) becomes

$$B = 1 - \frac{1}{2V} \left\{ T - \int_0^T F^2(t) dt \right\} \quad (7)$$

Under the same assumptions, we have

$$V = \int_{t=0}^T t dF(t) = T - \int_0^T F(t) dt \quad (8)$$

Let  $G$  denote the Gini index and it is straightforward that

$$G = 1 - 2B = \frac{\int_0^T F(t) dt - \int_0^T F^2(t) dt}{T - \int_0^T F(t) dt} \quad (9)$$

Next, we derive the relation between urbanization and the Gini index. Suppose the cumulative population-to-income distribution functions in the urban and rural sectors can be represented by  $F_u(t)$  and  $F_r(t)$ , respectively. Let  $x \in [0,1]$  denote the urbanization rate (i.e., the percentage of urban population), and let  $y = 1-x$  denote the percentage of rural population. We have

$$F(t) = xF_u(t) + yF_r(t) \quad (10)$$

Let  $U$  denote the average income of the urban population,  $R$  denote the average income of the rural population,  $G_u$  denote the urban Gini index, and  $G_r$  denote the rural Gini index. Based on (8) and (9), we have

$$U = T - \int_0^T F_u(t) dt \quad (11)$$

$$R = T - \int_0^T F_r(t) dt \quad (12)$$

$$G_u = \frac{\int_0^T F_u(t) dt - \int_0^T F_u^2(t) dt}{U} \quad (13)$$

$$G_r = \frac{\int_0^T F_r(t) dt - \int_0^T F_r^2(t) dt}{R} \quad (14)$$

Moreover, we define

$$D = \int_0^T \{F_u(t) - F_r(t)\}^2 dt \quad (15)$$

By substituting (10)–(15) into (9) and letting  $G' = G_u - G_r$  denote the difference between the urban and rural Gini indexes, we have

$$G = G_r + \frac{UG/x + Dxy}{Ux + Ry} = G_r + \frac{-Dx^2 + (UG' + D)x}{(U - R)x + R} \quad (16)$$

To study the monotonicity of (16), we consider the first derivative of  $G$ :

$$\frac{dG}{dx} = \frac{-(U - R)Dx^2 - 2RDx + R(UG' + D)}{\{(U - R)x + R\}^2} \quad (17)$$

As  $\{(U - R)x + R\}^2 > 0$ , setting (17) to 0 requires the nominator to be 0. By solving the quadratic equation

$$-(U - R)Dx^2 - 2RDx + R(UG' + D) = 0 \quad (18)$$

we have

$$x = \frac{\pm \sqrt{UR\{1 + \frac{G'}{D}(U - R)\}} - R}{U - R} \quad (19)$$

Note that by definition  $0 \leq x \leq 1$  and we assume  $U > R$ . Therefore,  $\frac{-\sqrt{UR\{1 + \frac{G'}{D}(U - R)\}} - R}{U - R < 0}$  and the turning point of (16) should be at

$$X = \frac{\sqrt{UR\{1 + \frac{G'}{D}(U - R)\}} - R}{U - R} \quad (20)$$

As the opening of the quadratic curve is downward and  $X$  is the larger root of (18), we can see that when  $0 \leq x < X$ ,  $dG/dx > 0$  and the Gini index increases with urbanization. Similarly, when  $X < x \leq 1$ ,  $dG/dx < 0$  and the Gini index decreases with urbanization. This is consistent with the inverse-U-shaped curve depicted by Kuznets. It is noteworthy that  $X$  is not always on  $[0, 1]$ . If  $G' > D/R$ , i.e., if the urban Gini index exceeds the rural Gini index plus  $D/R$ , then  $X > 1$  and urbanization will always cause  $G$  to increase. On the other hand, if  $G' < -D/U$ , i.e., if the urban Gini index is smaller than the rural Gini index minus  $D/U$ , then  $X < 0$  and urbanization will always cause  $G$  to decrease.

## Predictive Causal Model

The advantage of (20) is that it only depends on the rural and urban income distributions, as  $U, R, G_u$ , and  $G_r$  are all functions of  $F_u(t)$  and  $F_r(t)$ . A main assumption is that both  $F_u(t)$  and  $F_r(t)$  remain the same after the population shift. In reality, however, this condition is unlikely to hold. Newly arriving rural migrants are concentrated in the lower-income strata and/or informal sectors of the urban economy, which tends to downwardly skew  $F_u(t)$ . It takes time for  $F_u(t)$  to recover and the extent to which  $F_u(t)$  can recover depends on a range of factors, including the structure of the urban economy, the extent of social mobility, and access to opportunities for the migrant groups. Moreover, some of the impact of urbanization could be long-lasting and/or lag over time. For example, adding needed labor to a fast-growing urban

sector can accelerate economic expansion, whereas barriers to opportunities can trap rural migrants in persistent poverty [18, 19]. In both situations, income distribution will shift and income inequality may change over time.

For these reasons, in reality the impact of urbanization on income inequality is a complex process that may span multiple years. Under the traditional causal inference framework based on *ceteris paribus* (i.e., all other factors equal), addressing such dynamics requires highly complex structural models (or systems of models) with a large number of parameters. In most situations, however, such models are difficult to correctly specify and/or efficiently estimate due to limitations in data availability, measurement quality, and model statistical power. Moreover, such models are highly sensitive to misspecification. It is not uncommon that a dynamic structural model with many parameters is not immediately identifiable and additional assumptions are needed to restrict the parameter space. As argued by C. Sims, these requirements not only obfuscate the logic of causal inference but they can also produce highly misleading results, as the numerous assumptions imposed by the structural model are not necessarily warranted in real-world analysis [20].

The concept of predictive causality, largely pioneered by Clive Granger [21], offers an alternative approach to causal inference in dynamic domains. In this approach, causality between two processes is defined by whether observations of one can help predict the other. Compared with *ceteris paribus* this is a weaker causality framework, as it merely describes the temporal correlation between two time-series without controlling for other factors. In other words, the predictive causality model focuses on prediction but may not offer the same explanatory power as a traditional dynamic structural model would. However, it provides a simple yet powerful tool in practice for domains such as macroeconomic analysis, where data are often limited and detailed structural relationships are usually too complex to correctly specify and/or identify. In this study, we hypothesize that predictive causal relationships exist between urbanization and income inequality. We first use vector autoregression (VAR) to explore such relationships. Since we have a relatively small number of time points, the results of the VAR model are not intended to provide inferential evidence; instead they are interpreted exploratorily to guide the specification of a more parsimonious autoregressive model. The final inference is based on a simple autoregressive model, which requires fewer parameters and thus can be more convincingly supported by the short time series.

From the perspective of traditional structural regression, the omission of variables could be a serious problem for our analyses. For example, one may argue that any relationships between urbanization and Gini index are spurious because the omission of economic development, which affects both urbanization and income inequality. This is indeed an issue if we take the model at face value in policy-making and attempt to regulate inequality by boosting or controlling urbanization without checking the conditions of other relevant factors such as economic growth [20]. However, it is less a concern if we stick with the logic of predictive causality and treat the results only as evidence of two temporally correlated processes observed in the study time frame. In other words, our goal is to reveal the temporal relationships between urbanization and income inequality as suggested by the time series, which should not be interpreted explanatorily but could offer an informative baseline to elicit more detailed explanatory studies.

## Data and Methods

The official counts of urban population in China have historically been subject to inconsistencies in regard to both the definition of urban resident and the protocols used to collect data [22]. Since 2000, Chinese authorities have adopted a more sensible and internationally

comparable definition of urban population, which counts all permanent residents (i.e., those who have lived locally for at least 6 months) in urban areas defined by socioeconomic activities rather than administrative boundaries [22]. The annual urbanization data before 2000 have since been revised using the new definition and based on the results of the population censuses. In this study, we use the latest revised urbanization data series, which is now included as the official data in the NBSC's most recent publications.

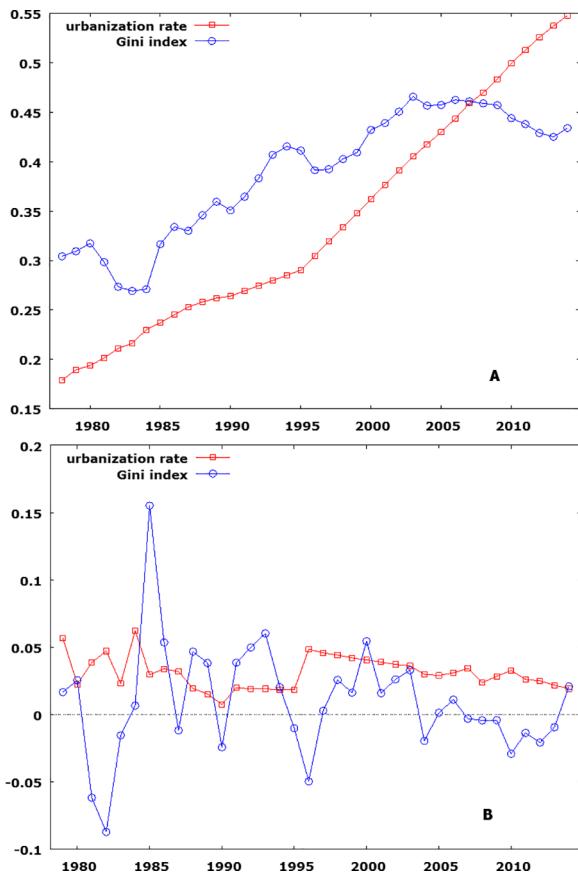
The Chinese authorities did not consistently publish official estimates of the national Gini index before 2003, although the NBSC has a history of reporting the average disposable (or net) incomes of different urban and rural income strata estimated based on household surveys. In addition to the NBSC's surveys, income data are also available from other national surveys, e.g., those administered under the China Household Income Project (CHIP) [23] and the Chinese Family Panel Studies (CFPS) [24]. Unfortunately, only the NBSC surveys provide a relatively complete data series since 1978. In the related literature, a common method used to obtain the annual Gini index series is to fit the urban and rural income distributions based on the NBSC data. In the present study, the Gini indexes from 1978 to 2006 were taken directly from Chen (2010), who used the lognormal functions to fit the grouped income data from the NBSC's urban and rural surveys [25]. Estimates for the years after 2006 were calculated based on lognormal functions fitted by the authors using the NBSC data. We also used Gini index estimates from other studies to cross-check our results. It turned out that although different estimation methods produced different Gini index values for individual years, the methods generated similar trend patterns for the time series.

We followed the standard procedure outlined in the related literature to conduct the time series analyses [26]. First, we explored trends in the time series and performed unit root tests to confirm the (non-)stationarity in the data. After the expected non-stationarity in all three variables was confirmed, we log-transformed the original variables to linearize their relationships and took the first differences of the log-transformed data. The VAR model was built using the first differences of the log-transformed variables. After that, orthogonalized impulse-response functions were obtained based on Cholesky decomposition. It is reasonable to assume that (1) urbanization may have immediate effects on income inequality and (2) the feedback from income inequality to urbanization may occur at a later time. Therefore, the variable ordering for Cholesky decomposition should be urbanization followed by inequality.

Based on the results of the VAR model, Granger causality tests were conducted as F-tests (or equivalence) on the coefficients of relevant autoregressors. After exploring the plots of impulse response functions produced by the VAR model, we built an autoregressive model with a minimal number of parameters—which only includes variables at time lags that were found to be relevant by the VAR model. The ACF and PACF plots were used to determine the necessity of the autoregressive (AR) and moving average (MA) terms. The final model was estimated using the Cochrane-Orcutt method. All the data treatments were performed in the software packages EViews [27], R [28] and gretl [29].

## Results

The time series of China's urbanization rate and Gini indexes from 1978 to 2014 are plotted in Fig 2. Both processes appear to be non-stationary and the first differences of their log-transformed values appear to be stationary. These impressions are confirmed by unit root tests (Table 1). The first lull of the Gini index series in the early 1980s can be attributed to the rural decollectivization during the early stage of the reform [8, 30, 31]. But as the process of urban reform picked up pace, the Gini index's upward trend resumed. Meanwhile, it is likely that the decrease in income inequality in the mid-1990s is related to the wage reform [32], which



**Fig 2. Urbanization rate and Gini index in China from 1978 to 2014.** Fig 2A shows the time series in levels. Fig 2B shows the first differences of log-transformed data.

doi:10.1371/journal.pone.0158826.g002

implemented a more merit-based reward structure and increased the wage levels in most work units [33]. However, as state-owned enterprises (SOEs) started to face drastic reform measures taken by the government that pushed many of SOE employees out of work [34–36], income inequality climbed again in the latter half of the 1990s.

The Gini index's dip after 2006, unfortunately, is less clearly understood. On the one hand, China's national income inequality may have been alleviated by recent economic and policy changes, including the shift of manufacturing jobs from the coastal areas to the inland

**Table 1. Results of unit tests.**

	Urbanization rate	Gini index
ADF	> 0.1	> 0.1
ADF-GLS	> 0.1	> 0.1
KPSS	< 0.01	< 0.01
ADF	< 0.01	< 0.01
ADF-GLS	< 0.01	< 0.01
KPSS	> 0.1	> 0.1

For ADF and ADF-GLS tests, the null hypothesis is that the series has a unit root. For KPSS tests, the null hypothesis is that the series is stationary. The ADF and ADF-GLS tests followed a “test-down” procedure starting with a maximal lag order of 8, whereas the KPSS tests used a lag order of 3.

doi:10.1371/journal.pone.0158826.t001

provinces [37], the abolition of agricultural taxes [38], and the continued development of a social safety net together with income redistribution [39–42]. After all, one cannot deny that given China's rapid economic development, there is a possibility that the downhill part of the Kuznets curve may have been reached and that income inequality has started to decline. On the other hand, researchers remain highly suspicious of the official data, as almost all non-governmental data sources reported higher estimates of the Gini index for the same period of time than the official data sources suggest. The same kind of decline, for example, is not seen in estimates based on CFPS data [12]. To ensure the consistency of our analysis, we proceeded with the estimates based on NBSC, but used sub-samples and alternative estimates of the Gini index to cross-validate the results.

## Results of VAR Estimation

It is noteworthy that trend changes around the mid-1990s can also be observed in the urbanization rate series, which appears to climb steeply after 1995 (Fig 2A). This could be the cumulative effects of relevant transformations in China's economic system, which started following a more city-oriented strategy to accelerate economic growth in the 1990s [8]. The full-scale housing marketization, which jump-started a fast-growing urban real-estate industry, also occurred in the second half of the 1990s [43]. Consequently, we created a dummy exogenous variable, which was assigned as 1 for the years after 1995 and 0 otherwise. A second dummy variable was created for year 1989 (i.e., 0 for all but year 1989), as observations in this year turned out to be outliers. Eventually, a VAR(2) model was selected based on information criteria, and its estimates are summarized in Table 2. Despite the relative short time series, the VAR model passed all tests for model adequacy. Consequently, the VAR model provides good exploratory insights over the data.

Fig 3 depicts the impacts of an unexpected unit (one standard deviation) shock in one variable on the other over time. For the Gini index, the immediate response to urbanization is

**Table 2. Summary of VAR estimates.**

	Urbanization rate	Gini index
For individual equation		
R-squared	0.912437	0.458810
Adjusted R-squared	0.892978	0.338546
Sum of squared residuals	0.003164	0.033568
F-statistic (7, 27)	40.19268	3.270014
P-value	< 0.001***	0.012**
Mean dependent	0.030539	0.009202
S.E. equation	0.010825	0.035260
S.D. dependent	0.011581	0.042336
F-tests of zero restrictions (Granger causality tests)		
All lags of urbanization rate	< 0.001***	0.007***
All lags of Gini index	0.6261	0.006***
For the system as a whole		
OLS estimator	T = 34 (1981–2014), Log-likelihood = 180.19023	

Significance codes:

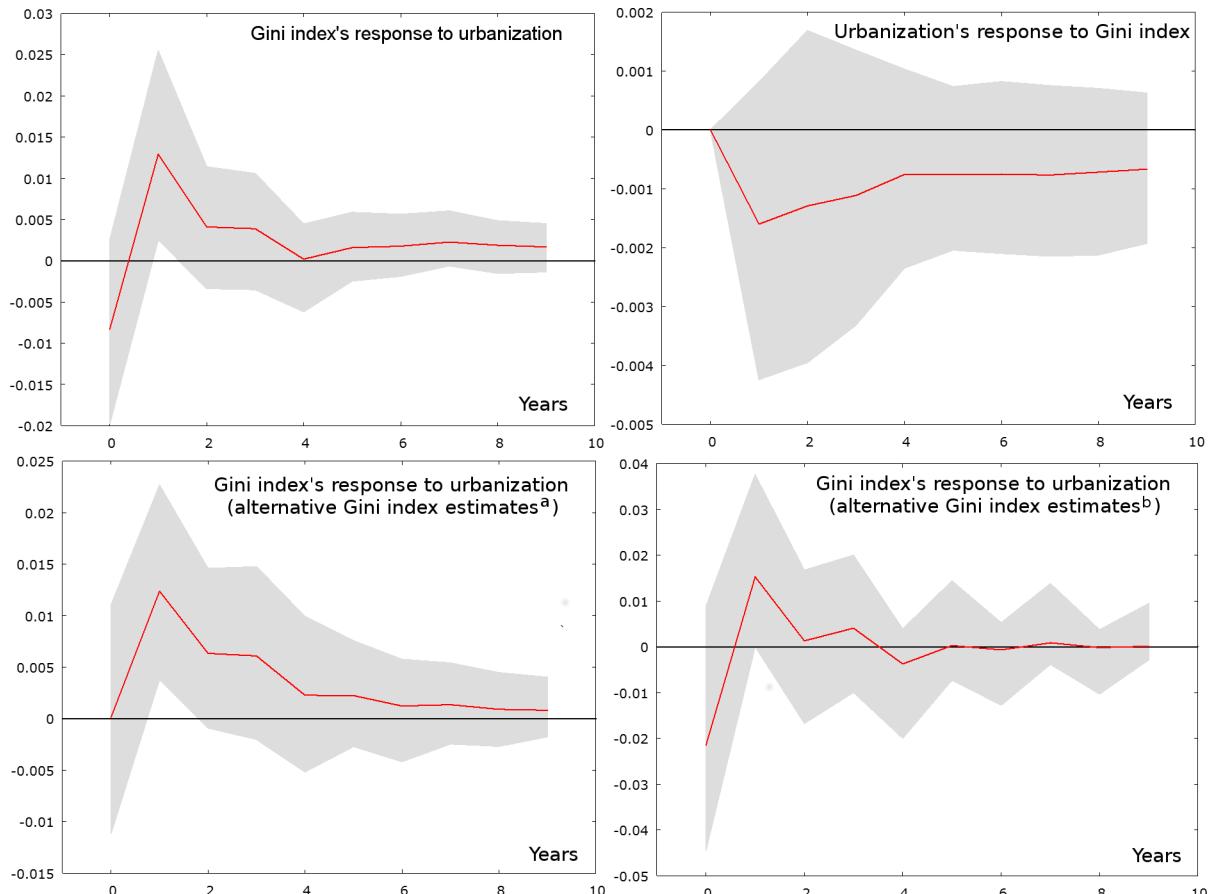
\* < 0.1

\*\* < 0.05

\*\*\* < 0.01

Model estimator: OLS with T = 34 (1981–2014)

doi:10.1371/journal.pone.0158826.t002



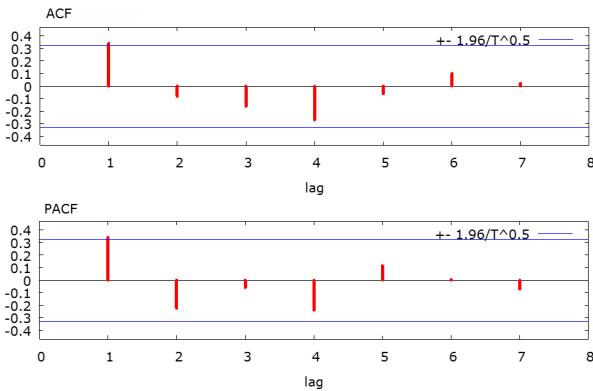
**Fig 3. Impulse response functions with 90% intervals.** <sup>a</sup> Using Gini index estimates by Cheng (2010) (Years: 1981–2004). <sup>b</sup> Using Gini index estimates by Wu and Perloff (2006) (Years: 1985–2001).

doi:10.1371/journal.pone.0158826.g003

negative, suggesting that the urbanization shock leads to reduced income inequality in year 0. However, a significant positive hike occurs in year 1, which means that income inequality's response tends to increase significantly one year after the initial urbanization shock. On the other hand, urbanization's response to the Gini index shock is not significantly different from zero. In other words, the impulse response functions suggest a one-way causal relationship from urbanization to the Gini index, which is consistent with the results of the Granger causality tests in [Table 2](#). To show the robustness of the Gini index's response, [Fig 3](#) also presents impulse response functions obtained using alternative Gini index estimates [44, 45] under the same model specification. These data series contain even fewer time points than our data, and the responses they produce barely cross the 90% confidence intervals. But it is important to note that all these impulse response functions exhibit similar trend patterns. From the perspective of explorative analysis, this means that the Gini index's response to urbanization revealed by the VAR model is unlikely to be a random artifact.

## Results of AR(1) Estimation

Based on the exploratory interpretation of the VAR results, an AR(1) model in first differences with no constant and no MA term was built for the Gini index. The correlograms for ACF and PACF for the first differences of the log-transformed Gini index series were depicted in [Fig 4](#),



**Fig 4. Correlogram for the first differences of the log-transformed Gini index series.** ACF: autocorrelation function; PACF: partial autocorrelation function.

doi:10.1371/journal.pone.0158826.g004

which justified the inclusion of the AR(1) term and the exclusion of the MA term. According to the results of the Granger tests in [Table 2](#), urbanization Granger-causes the Gini index but not vice versa. So we include the urbanization rates in lag 0 and lag 1 as exogenous variables. The Gini index at lag 1 was also included as a control. The results of the 3-parameter AR(1) model were reported in [Table 3](#). All three terms were highly significant, suggesting the Gini index in any year is dependent on its value in the previous year as well as the urbanization rates in both the same year and the previous year. The signs of coefficients for the two urbanization rate variables are consistent with what the VAR model depicts, i.e., negative in year 0 and positive in year 1. Consequently, there is high statistical confidence that urbanization's prediative causal effects on income inequality are not artifacts.

## Discussion

The Gini index's lagged positive response indicates that urbanization may aggravate income inequality over time—a point that should be addressed as China's marketization continues to move forward. It is possible to identify three general directions for more detailed explanatory research from the perspectives of regional inequality, the rural–urban gap, and the persistent marginalization of rural migrants.

## Potential Aggravating Factors of Inequality

First, urbanization may have worsened regional inequality. Regional development in China is known to follow a laddered geographic pattern, where the fast-growing coastal areas and the

**Table 3. Summary of AR(1) estimates.**

	Coefficient	Std. Error	p-Value
Urbanization rate in the same year (lag 0)	-1.24793	0.586921	0.0416**
Urbanization rate in the previous year (lag 1)	1.55939	0.501417	0.0040***
Gini index in the previous year (lag 1)	0.19676	0.366520	0.5952

Significance codes:

\* < 0.1

\*\* < 0.05

\*\*\* < 0.01

Dependent variable: Gini index (first difference of log-transformed values); Model estimator: Cochrane-Orcutt with T = 34 (1981–2014); Normality test of residuals: p-value = 0.32 (null hypothesis of normal distribution)

doi:10.1371/journal.pone.0158826.t003

less-developed inland provinces (or the eastern, central, and western regions in a tri-partite view) form distinct growth “clubs” [46]. It has been shown that economic development has converged within each club, which helps suppress regional inequality at the national level [47]. But the inter-club gap seems to be persistent despite the central government’s recent efforts to accelerate inland development [48]. As migrants from the inland rural areas have been an important force for both the urbanization and economic development of the coastal regions, it is reasonable to hypothesize that urbanization may have causally increased the income gap between the origin and destination of the migration due to the widening difference in economic growth. Note that this process is probably more subtle than coarse scale statistics (e.g., average GDP and/or income at the provincial level) can indicate, because it is highly dependent on the structure of local urbanization, the conditions of the local labor market, and specific migration flows.

Second, the lagged positive response in the Gini index may reflect a growing difference in productivity between the rural and urban sectors. Despite the success of the rural decollectivization in the early stage of China’s economic reform, the incentives for farming have gradually languished due in part to high taxes and fees and relatively low returns [49]. Moreover, crop-land in rural China is tenured through the household responsibility system (HRS), under which land is owned collectively but controlled by individual households. Although HRS boosts household-level productivity and protects land security for peasants [50], it limits technological advancements and economies of scale due to ambiguous property rights [51], the under-developed land rental market [52], and the lack of adequate financial services for individual farmers [53]. As the urban economy continues to outpace its rural counterpart, the opportunity cost for farming rises and the most productive group of the rural population (e.g., young, healthy, and educated adults) continue to leave for the cities. Consequently, a structural shortage of human resources has emerged in some rural areas [54] and the rural sector lags further behind the booming urban sector in terms of productivity. Again, this process may be difficult to identify using aggregated data given the variation in population flows and local labor markets.

Third, urbanization’s lagged effect on income inequality could be the result of the persistent marginalization of the rural migrants. Rural migrants in China are known to face additional hurdles set up by discriminatory institutions, especially those related to the Hukou household registration system [55], which ties people’s access to services to their residential status. Due to the institutionalized measures whereby migrants are deprived of job opportunities, educational rights, and social welfare benefits, the marginalization of a rural migrant usually lasts for an extended period of time [56]. Consequently, urbanization can lead to a ballooning low-income group that further widens urban income inequality, and the lack of integration with the mainstream urban society can trap migrant groups in segregated spaces and long-term poverty [19, 57]. Although a well-recognized problem in the related literature, the persistent poverty of rural migrants is poorly documented in many official and unofficial data sources, as most general-purpose surveys tend to undersample the migrant population [58]. Future research on this subject, therefore, must draw extensively on more representative and preferably firsthand data sources.

## The Importance of Social Reform

It is noteworthy that some of the above problems have been addressed to some extent in recent years. The relocation of the manufacturing industry from the coastal areas to the inland provinces [37] has the potential to reduce urbanization’s aggravating effects on regional inequality, as many rural migrants can now find jobs in nearby cities. The abolition of the agricultural tax

[38] and the continued development of the cropland rental market help boost agricultural productivity [52]. There are also advances in health care and social welfare reforms, which have greatly expanded the population protected by the emerging social safety net [39, 42, 59]. A main challenge at this moment, however, is probably the Hukou reform. Although restrictions on employment and residential mobility have gradually withered as economic marketization has deepened, the role of Hukou as a social divider continues to prevail. The ongoing effort to extend social and welfare services to non-local Hukou holders has been beleaguered by the lack of an affordable and accountable solution. Currently, local governments at the receiving end of urbanization are asked to carry out the reform and absorb its enormous cost, which has created significant budgetary pressure on the cities. As the recent slowdown in China's real estate sector has worsened widespread local debt situations [60], real progress in regard to the Hukou reform is at a standstill in most parts of China.

From a broader policy perspective, the problems of assimilation and integration caused by long-standing discriminatory institutions have resulted in a permanent dichotomy in the social and economic system, which have already obstructed China's progress toward the second stage of the Lewis model. For example, many rural migrants have to leave dependents, including women, children, and aging parents, in the countryside because of the cost and restrictions imposed by the Hukou system on housing, education, and retirement arrangements. The resulting separation of families, which is estimated as affecting more than 60 million children, is creating profound social, psychological, and human development problems that may haunt Chinese society for generations to come [61]. Due to the high stakes nature of this issue, it is important for both central and local governments to design and effect a more comprehensive, cross-region, and cross-administrative-level reform strategy, as problems such as those related to Hukou cannot be solved by the central or local government alone.

In summary, a broad message from this study is the importance of comprehensive social reforms. Currently, the discussion on China's urbanization is preoccupied by two dominant theses: the pro-growth thesis that considers urbanization an economic engine, and the pro-sustainability thesis fueled by reactions to problems created or exposed by urbanization. In our opinion, both theses, in their current forms, lack a social dimension, and as a result, the ongoing policy debate is often out of touch with the people directly affected by the urban transition. Pro-growth urbanization policies embraced by GDP-chasing local governments, for example, have led to land-oriented rather than people-centered urban expansion, creating problems ranging from urban sprawl to environmental degradation [62]. On the other hand, the pro-sustainability argument often mistakenly places the blame for environmental woes on the growing urban population, ignoring the fact that expanding land development and increasing consumption among the existing urban residents are probably more culpable than any effect wrought by newcomers. Addressing the income-inequality problem helps reintroduce social justice into the discussion, which constitutes an important first step toward a more socially balanced urbanization policy.

## Conclusion

In summary, this study shows that the rising income inequality in post-reform China can be causally associated with urbanization. Time series analyses indicate that urbanization has a non-linear and time lagged effect on income inequality, as it causes the Gini index to initially decrease but eventually increase. The lagged aggravating effect of urbanization on income inequality points to an important direction for future research, which can be approached from the perspectives of regional inequality, the rural–urban income gap, and the persistent marginalization of rural migrants. These results highlight the lack of a social dimension in China's

urbanization policy, particularly the stagnation of the Hukou reform. Due to the profound impact this problem may have on China's economic and social system, more comprehensive social reforms must be prioritized in order to redress the widening income inequality during urbanization and address the discriminatory institutions behind it.

## Supporting Information

S1 File. Urbanization rates and Gini index estimates in China from 1978 to 2014.  
(CSV)

## Author Contributions

Conceived and designed the experiments: GC. Performed the experiments: GC. Analyzed the data: GC. Contributed reagents/materials/analysis tools: GC AKG MZ. Wrote the paper: GC AKG MZ. Contributed to revising the manuscript: YS.

## References

1. National Bureau of Statistics of China. China Statistical Yearbook. Beijing, China: China Statistics Press; 2015.
2. Ho P, Eyferth JJK, Vermeer EB. Rural development in transitional China: the new agriculture. London; Portland, OR: Frank Cass; 2004. 309 p. p.
3. Yao SJ. Economic development and poverty reduction in China over 20 years of reforms. *Economic Development and Cultural Change*. 2000; 48(3):447–74. doi: [10.1086/452606](https://doi.org/10.1086/452606) PMID: [WOS:000086492600001](https://pubmed.ncbi.nlm.nih.gov/1086452606/).
4. Ravallion M, Chen S. China's (uneven) progress against poverty. *Journal of development economics*. 2007; 82(1):1–42.
5. Yang DT. Urban-Biased Policies and Rising Income Inequality in China. *The American Economic Review*. 1999; 89(2):306–10. doi: [10.2307/117126](https://doi.org/10.2307/117126)
6. Bao SM, Chang GH, Sachs JD, Woo WT. Geographic factors and China's regional development under market reforms, 1978–1998. *China Economic Review*. 2002; 13(1):89–111. doi: [10.1016/s1043-951x\(02\)00055-x](https://doi.org/10.1016/s1043-951x(02)00055-x) PMID: [WOS:000176994100005](https://pubmed.ncbi.nlm.nih.gov/1200176994100005/).
7. Fan CC. Uneven development and beyond: Regional development theory in post-Mao China. *International Journal of Urban and Regional Research*. 1997; 21(4):620–+. doi: [10.1111/1468-2427.00105](https://doi.org/10.1111/1468-2427.00105) PMID: [WOS:000071062500007](https://pubmed.ncbi.nlm.nih.gov/9300071062500007/).
8. Huang Y. Capitalism with Chinese characteristics: entrepreneurship and the state. Cambridge; New York: Cambridge University Press; 2008. xviii, 348 p. p.
9. Xie Y, Zhang X, Li J, Yu X, Ren Q. China Minsheng Development Report 2014. Beijing, China: Peking University Press; 2014.
10. Kuznets S. Economic growth and income inequality. *The American economic review*. 1955;1–28.
11. Kanbur SMR. Inequality in Asia and the Pacific: trends, drivers, and policy implications2014. xxiii, 412 pages p.
12. Xie Y, Zhou X. Income inequality in today's China. *Proceedings of the National Academy of Sciences*. 2014; 111(19):6928–33. doi: [10.1073/pnas.1403158111](https://doi.org/10.1073/pnas.1403158111)
13. Lin J, Chen B. Urbanization and urban-rural inequality in china: a new perspective from the government's development strategy. *Front Econ China*. 2011; 6(1):1–21. doi: [10.1007/s11459-011-0119-3](https://doi.org/10.1007/s11459-011-0119-3)
14. Gini C. Concentration and Dependency Ratios. *Rivista di politica economica*. 1997; 87(8–9):769–90.
15. Cheng Y. Calculation and decomposition of the overall gini coefficient in dual economies. *Economic Research Journal*. 2006; 1(1):109–20.
16. Brown JAC, Mazzarino G. PRACTITIONER'S CORNER\*: Drawing the Lorenz Curve and Calculating the Gini Concentration Index from Grouped Data by Computer. *Oxford Bulletin of Economics and Statistics*. 1984; 46(3):273–8. doi: [10.1111/j.1468-0084.1984.mp46003005.x](https://doi.org/10.1111/j.1468-0084.1984.mp46003005.x)
17. Dorfman R. A Formula for the Gini Coefficient. *The Review of Economics and Statistics*. 1979; 61(1):146–9. doi: [10.2307/1924845](https://doi.org/10.2307/1924845)
18. Chen G, Gu C, Wu F. Urban poverty in the transitional economy: A case of Nanjing, China. *Habitat International*. 2006; 30(1):1–26.

19. Liu Y, W F., H S. The making of the new urban poor in transitional China: Market versus institutionally based exclusion. *Urban Geography*. 2008; 29(8):811–34.
20. Sims CA. Macroeconomics and Reality. *Econometrica*. 1980; 48(1):1–48. doi: [10.2307/1912017](https://doi.org/10.2307/1912017)
21. Granger CWJ. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*. 1969; 37(3):424–38. doi: [10.2307/1912791](https://doi.org/10.2307/1912791)
22. Chan KW, Hu Y. Urbanization in China in the 1990s: New Definition, Different Series, and Revised Trends. *China Review*. 2003; 3(2):49–71. doi: [10.2307/23461904](https://doi.org/10.2307/23461904)
23. Shi L. Chinese Household Income Project, 2002. Inter-university Consortium for Political and Social Research (ICPSR) [distributor]; 2009.
24. Xie Y, Hu J. An Introduction to the China Family Panel Studies (CFPS). *Chinese Sociological Review*. 2014; 47(1):3–29. doi: [10.2753/CSA2162-0555470101](https://doi.org/10.2753/CSA2162-0555470101)
25. Chen J, Dai D, Pu M, Hou W, Feng Q. The trend of the Gini coefficient of China. *BWPI*, The University of Manchester, 2010.
26. tkepohl H. New introduction to multiple time series analysis. Berlin;; New York: Springer; 2005. xxi, 764 p. p.
27. McKenzie CR, Takaoka S. EViews 5.1. *Journal of Applied Econometrics*. 2007; 22(6):1145–52. doi: [10.1002/jae.947](https://doi.org/10.1002/jae.947)
28. R Core Team. R: A language and environment for statistical computing. 2013. Available from: <http://www.R-project.org/>.
29. Baiocchi G, Distaso W. GRETL: Econometric software for the GNU generation. *Journal of Applied Econometrics*. 2003; 18(1):105–10. doi: [10.1002/jae.704](https://doi.org/10.1002/jae.704)
30. Knight J, Song L. The rural-urban divide: Economic disparities and interactions in China. OUP Catalogue. 1999.
31. Du R. Many people, little land: China's rural economic reform. Foreign Languages Press; 1989.
32. Knight J, Song L. China's emerging urban wage structure, 1995–2002. *Inequality and public policy in China*. 2008:221–42.
33. Yueh LY. Wage Reforms in China During the 1990s. *Asian Economic Journal*. 2004; 18(2):149–64. doi: [10.1111/j.1467-8381.2004.00187.x](https://doi.org/10.1111/j.1467-8381.2004.00187.x)
34. Appleton S, Song L, Xia Q. Has China crossed the river? The evolution of wage structure in urban China during reform and retrenchment. *Journal of Comparative Economics*. 2005; 33(4):644–63. <http://dx.doi.org/10.1016/j.jce.2005.08.005>
35. Garnaut R, Song L, Yao Y. Impact and Significance of State-Owned Enterprise Restructuring in China. *The China Journal*. 2006;(55):35–63. doi: [10.2307/20066119](https://doi.org/10.2307/20066119)
36. Lin JY, Cai F, Li Z. State-owned enterprise reform in China. Chinese University Press; 2001.
37. Lemoine F, Poncet S, Ünal D. Spatial rebalancing and industrial convergence in China. *China Economic Review*. 2015; 34:39–63. <http://dx.doi.org/10.1016/j.chieco.2015.03.007>.
38. Chen S, Wang J. Rural Tax Reform: Agricultural Productivity and Local Public Finance. Available at SSRN 2374923. 2014.
39. Cai Y, Cheng Y. Pension reform in China: challenges and opportunities. *Journal of Economic Surveys*. 2014; 28(4):636–51.
40. Yip W, Hsiao WC. What Drove the Cycles of Chinese Health System Reforms? *Health Systems & Reform*. 2015; 1(1):52–61.
41. Ravallion M. An emerging new form of social protection in twenty-first-century China. *The Oxford Companion to the Economics of China*. 2014: 441.
42. Gao Q, Yang S, Li S. Welfare, targeting, and anti-poverty effectiveness: The case of urban China. *The Quarterly Review of Economics and Finance*. 2015; 56:30–42.
43. Wu FL. China's recent urban development in the process of land and housing marketisation and economic globalisation. *Habitat International*. 2001; 25(3):273–89. PMID: [000169995500001](https://pubmed.ncbi.nlm.nih.gov/11750001/).
44. Cheng Y. Evolution of overall national Gini coefficient and its urban-rural decomposition (in Chinese). *Social Sciences in China*. 2007;(4):.
45. Wu X, Perloff JM. China's income distribution, 1985–2001. *Review of Economics and Statistics*. 2005; 87(4):763–75.
46. Yao S, Zhang Z. On regional inequality and diverging clubs: a case study of contemporary China. *Journal of Comparative Economics*. 2001; 29(3):466–84.
47. Zhang Q, Zou H-f. Regional inequality in contemporary China. *Annals of Economics and Finance*. 2012; 13(1):113–37.

48. Candelaria C, Daly M, Hale G. Persistence of regional inequality in China. Citeseer; 2013.
49. Yep R. Can “Tax-for-Fee” reform reduce rural tension in China? The process, progress and limitations. *The China Quarterly*. 2004; 177:42–70.
50. Zhang QF, Donaldson JA. China's Agrarian Reform and the Privatization of Land: a contrarian view. *Journal of contemporary China*. 2013; 22(80):255–72.
51. Ho P. Who owns China's land? Policies, property rights and deliberate institutional ambiguity. *The China Quarterly*. 2001; 166:394–421.
52. Jin S, Deininger K. Land rental markets in the process of rural structural transformation: Productivity and equity impacts from China. *Journal of Comparative Economics*. 2009; 37(4):629–46.
53. Zhao Y. China's rural financial system: households' demand for credit and recent reforms. 1st ed. London; New York, NY: Routledge; 2011. xviii, 168 p. p.
54. Xie H, Wang P, Yao G. Exploring the Dynamic Mechanisms of Farmland Abandonment Based on a Spatially Explicit Economic Model for Environmental Sustainability: A Case Study in Jiangxi Province, China. *Sustainability*. 2014; 6(3):1260–82.
55. Chan KW. Cities with invisible walls: reinterpreting urbanization in post-1949 China. Hong Kong; New York: Oxford University Press; 1994. ix, 193 p. p.
56. Fan CC. The elite, the natives, and the outsiders: Migration and labor market segmentation in urban China. *Annals of the Association of American Geographers*. 2002; 92(1):103–24.
57. Chen G. Privatization, marketization, and deprivation: Interpreting the homeownership paradox in post-reform urban China. *Environ Plann A*. 2011; 43(5):1135–53. PMID: [ISI:000292993500010](#).
58. Chen G. The heterogeneity of housing-tenure choice in urban China: A case study based in Guangzhou. *Urban Studies*. 2015;0042098015571822.
59. Yip WC-M, Hsiao WC, Chen W, Hu S, Ma J, Maynard A. Early appraisal of China's huge and complex health-care reforms. *The Lancet*. 2012; 379(9818):833–42.
60. Ambrose BW, Deng Y, Wu J. Understanding the Risk of China's Local Government Debts and its Linkage with Property Markets. Available at SSRN 2557031. 2015.
61. Ye J, He C, Wu H. Left-behind population in rural China. Beijing: Social Sciences Academic Press; 2010.
62. Bai X, Chen J, Shi P. Landscape Urbanization and Economic Growth in China: Positive Feedbacks and Sustainability Dilemmas. *Environmental Science & Technology*. 2012; 46(1):132–9. doi: [10.1021/es202329f](#)