UNIVERSITY OF WAIKATO

Hamilton
New Zealand

Regional Inequality in China allowing for Spatial Cost-of-Living Differences: Evidence from a Hedonic Analysis of Apartment Prices

Chao Li and John Gibson

Working Paper in Economics 12/18

August 2018

Corresponding Author

John Gibson
School of Accounting, Finance and Economics
University of Waikato
Private Bag 3105
Hamilton, 3240
New Zealand

Tel: +64 (7) 838 4289

Email: jkgibson@waikato.ac.nz

Chao Li
School of Accounting, Finance and Economics
University of Waikato
Private Bag 3105
Hamilton, 3240
New Zealand

Email: chao.li@waikato.ac.nz

Abstract

Studies of inequality in China typically ignore cost-of-living differences between areas. Under the Balassa-Samuelson effect, non-tradeables cost more in richer areas, so nominal inequality exceeds real inequality. This especially matters in China, where spatial cost-of-living differences should have increased with recent development of urban housing markets. We use new data on apartment prices in 104 major cities in China to develop housing-related spatial deflators. The level of spatial inequality in 2016 is overstated 27 percent if cost-of-living differences are ignored. A hedonic analysis of 41,000 individual apartment sales shows most price variation is between areas, rather than from features of individual apartments. The dominant trend in the reform era is for regional inequality in China to decline, contrary to common perceptions. In nominal terms, the Theil Index for inter-provincial inequality in 2016 is just 46 percent of its 1978 level, and in real terms the fall in inequality would be even greater.

Keywords

China
housing
population
regional inequality
spatial deflators

JEL Classification

O47, Q56, R11

Acknowledgements

We are grateful to Geua Boe-Gibson for drawing the maps
and to comments from participants at the New Zealand Association of Economists conference.

1. Introduction

The density of economic activity, and the rewards from this density, are unevenly spread over space. Despite advice to policy makers to embrace the benefits that spatial concentration brings (World Bank, 2009), many interventions reveal concerns about spatial inequality and this is especially so in China. A typical view is that spatial inequality rose in the reform era, particularly when policy neglected the rural sector (Fan, Kanbur and Zhang 2011). Some initiatives to help seemingly laggard regions catch up to seemingly advanced regions, such as the West China Development Project (Lai 2002), follow from this concern with inequality. Likewise, infrastructure investments in western China under the Belt & Road Initiative (BRI) can be seen as part of the long-term attempt by policy makers to disperse economic activity to regions of China that are less economically concentrated (Gibson and Li 2018).

These interventions are costly, so it is important that policy makers are fully informed when making decisions to reallocate resources over space. In this regard, studies of inequality in China are hampered by incomplete evidence on price differences across space. Under the Balassa-Samuelson effect, prices for non-traded goods and services should be higher in areas with higher nominal incomes. Thus, welfare indicators like real income or real GDP, that adjust for spatial price differences, will show less inequality than do indicators valued in nominal terms. While this effect is present in all countries, it should especially matter in China because urban housing markets, and the spatial cost of living differentials due to these, have only developed recently. Thus, both the level of, and the trend in, regional inequality may be affected by modelling decisions about spatial deflation.

Despite the importance of spatial deflation for examining trends in inequality, China's National Bureau of Statistics (NBS) does not use a spatial price index that allows cost-of-living comparisons over space. Instead, the focus on the temporal consumer price index, which is reported at both national and provincial level, allows rates of change in consumer prices to be compared between locations but not inter-area comparisons of price levels or living costs. In response, researchers have developed unofficial spatial deflators from two main approaches: pricing a national basket of items in different regions in the country (for example, Brandt and Holz 2006); and, estimating food Engel curves to derive the deflator that equates food shares for otherwise similar households in different regions (for example, Gong and Meng 2008).

Neither approach is likely to be accurate, and there is ongoing need for reliable evidence on spatial differences in the cost-of-living in China, so that understanding of inequality trends has a firmer foundation. The problem with the Brandt and Holz (2006) approach is the focus on traded goods, given they had no prices for non-traded services, rents, and real estate. Despite claims of fragmented markets in China (for example, Young 2000), traded goods markets are highly integrated and follow the law of one price. Lan and Sylwester (2010) find half-life divergences from the law of one price average just 2.4 months, which is twice the speed of adjustment in the United States. Yet, the same is not true of housing, because of the fixity of land supply. Indeed, Moulton (1995, p.181) notes: 'costs of shelter are the single most important component of inter-area differences in the cost-of-living.' The problem with the Engel curve method is that things omitted from the equation that vary systematically over space get treated as regional cost-of-living differences. Moreover, this method compares badly with benchmark cost-of-living measures from price surveys (Gibson, Le and Kim 2017).

In light of these issues with past approaches, we use newly available data on apartment prices in 104 major cities in China, which are home to over 360 million residents, to develop housing-related spatial deflators. According to the Theil Index, spatial inequality in 2016 is overstated by 27 percent if nominal GDP per resident data are used instead of spatially deflated data that recognize cost-of-living differences. Given that our deflator accounts for only one source of cost-of-living variation over space, the reduced level of inequality that we find should be considered an upper bound to the true level; a more comprehensive deflator that covered other non-traded services, whose prices are also likely to vary over space in some proportion to the spatial variation in housing costs, should show even bigger gaps between inequality estimates that use deflated data and those that use nominal data.

In addition to developing these new spatial deflators, and demonstrating their impact on estimates of regional inequality, we show a mismatch between trends in regional inequality and perceptions about these trends. The reform era has seen regional inequality in China fall significantly since 1978. For example, the Theil Index for inter-provincial inequality in 2016 was at just 46 percent of its level in 1978, with the only sustained period of rising inequality from 1990 to 1993.[1] This trend is for inequality in nominal terms, because we lack data for spatial deflation in early years of the reform period; however, the nominal data overstate spatial inequality by at least 27 percent at the end of the period due just to housing cost variation over space. To the extent that much of this variation is recent, as China's real estate market has developed, spatial deflation in earlier periods should make less difference, and so the trend of declining regional inequality should be even stronger in real terms than in nominal terms.

One factor that may have caused a perception of rising regional inequality is that, until recently, the wrong population data were used to calculate per capita GDP of sub-national units. These wrong data counted where people were registered to live rather than where they actually lived (and where their activities added to GDP). The gap between the two population measures (*de jure* and *de facto*) grew as internal migrant numbers rose from just a few million to almost 300 million. This created apparent gaps in GDP per capita between the origin areas and the

[1] We emphasize the Theil Index, which can be decomposed into within and between components – a useful property for studying regional inequality that the widely used Gini index lacks. Note that the trend decline for regional inequality may not apply to inter-personal or inter-household inequality, which we do not study. Yet each of those two types of inequality also needs spatial deflators, so the findings here are still relevant.

destination areas that are much smaller once GDP is properly denominated by the local resident population. We show in Section 3 that a 'V'-shaped time pattern of regional inequality appears if the registered population data are used for calculating GDP per capita. Given that many studies of regional inequality use these wrong population denominators (Li and Gibson 2013), this issue likely contributed to perceptions of rising regional inequality. Thus, the next two sections of the paper, as part of the scene-setting for developing the spatial deflators, provide evidence on perceptions about China's regional inequality and evidence on the effect of these population denominator errors.

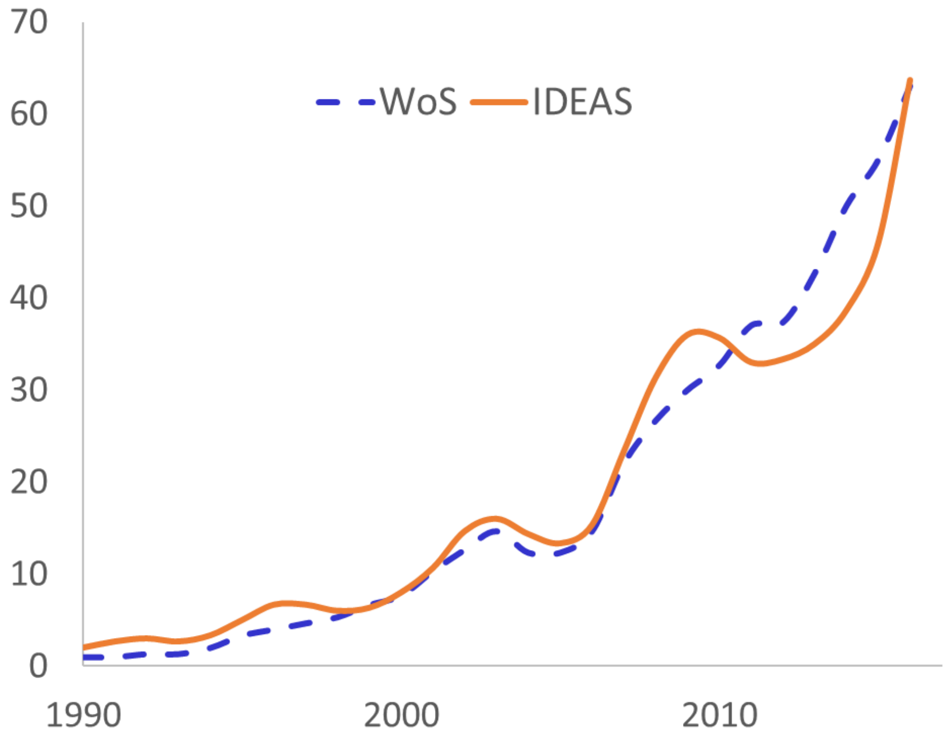
The central contribution of the paper then comes in Section 4, where the housing-related spatial deflators are developed and applied to city-level GDP data to see if spatial cost-of-living differences affect estimates of regional inequality. One concern is that spatial deflators based on city-level average apartment prices may mislead if apartment quality varies systematically over space. In Section 5, then, we report on a hedonic analysis of 41,000 individual apartment sales, in 35 districts of six Chinese cities of various size. There do not appear to be systematic differences in apartment quality over space, so inter-district price gaps in the 'raw' data are similar to the adjusted gaps that control for ten unit and complex characteristics. This hedonic analysis builds on prior work with a very small sample of just 150 apartments in three cities, which also found hedonically-adjusted differences in housing prices between cities to be similar to raw price gaps (Li and Gibson 2014). In line with the current results, that prior study found that not accounting for spatial variation in prices caused regional inequality to be overstated by 30 percent, at the district level, and by 35 percent, at the provincial level.

2. Perceptions about Rising Regional Inequality in China

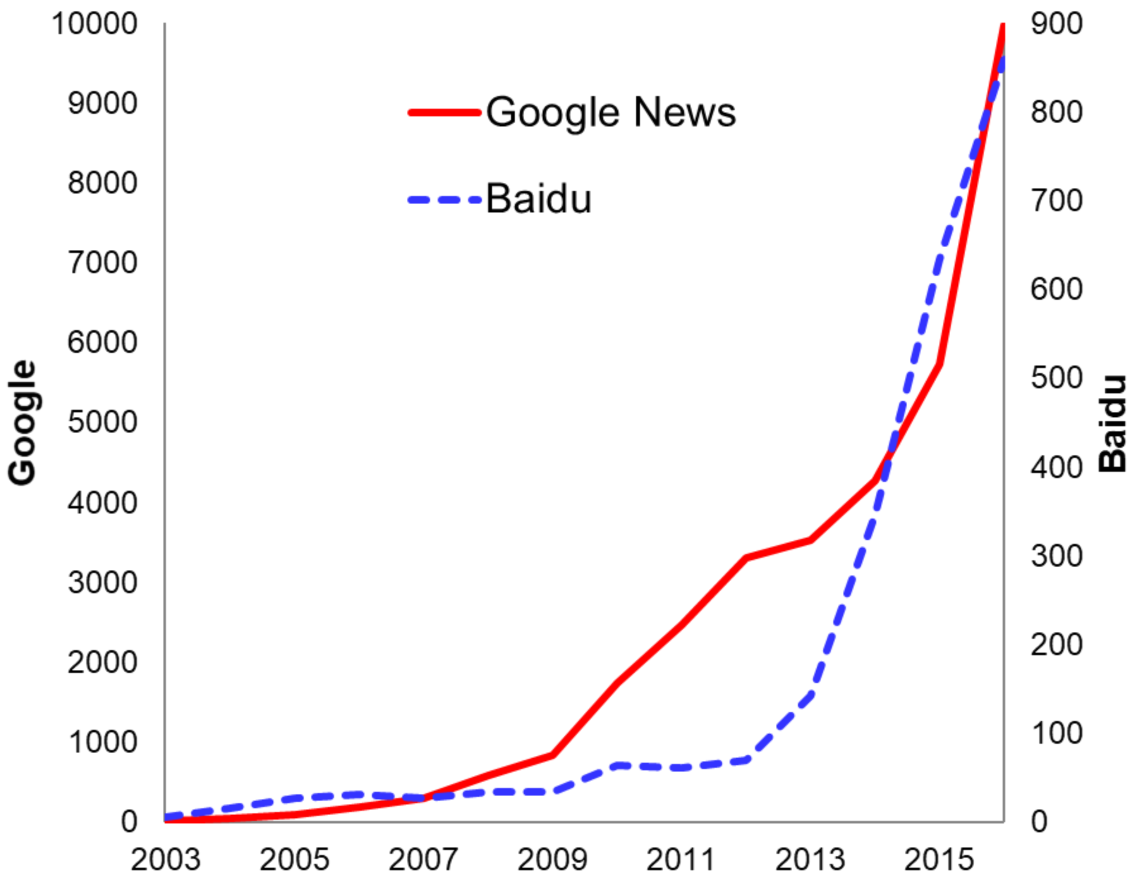
According to Li and Fang (2014), regional inequality has become an important issue in China, attracting considerable attention from all levels of government and from scholars. To quantify growing attention to this issue, we searched two types of literature. First, we counted how many journal articles (in English) have published since 1990 with the terms 'spatial inequality' and 'China'. We use *Web of Science* (WoS) for contributions from a range of disciplines, and *IDEAS/RePEc* for a focus on economics. The results are given in Figure 1a, as moving averages to show underlying trends without short-term volatility. Both databases show a rise in the number of journal articles published on the topic of China's spatial inequality, especially since 2005, when about 15 articles published, compared to about 65 articles published in 2016.

Figure 1: Evidence on Rising Attention to Spatial Inequality in China

(a) Journal Articles (in English)



(b) News articles (in Chinese)



In Section 3 we show one measure of spatial inequality, for inter-provincial gaps in GDP per capita, fell by over 40 percent from 2005 to 2016 (for the Theil Index). The growing attention by scholars was to a 'problem' that by one measure was shrinking. It is possible that growth in the number of articles was to celebrate the fall in spatial inequality, but in our reading the opposite applies; most studies are motivated by, or claim to show, spatial inequality becoming a bigger problem. For example, a review covering inter-household inequality, regional inequality, and urban-rural inequality argued for a similar upward trend in all three types of inequality in China from the early 1980s until about 2010 (Wang *et al.* 2014). Another trend is that recent articles study smaller spatial units, such as counties (for example, Li and Fang 2014), while the earlier studies mostly looked at inequality at province level or for broader regions (west, central and east).

Two modelling problems with focusing on smaller units may contribute to perceptions of rising inequality. First, the population denominator problems described in Section 3 are bigger: there are more unregistered migrants (a migrant from the countryside to their provincial capital city is not an unregistered migrant at the province level but is at the county level); the unregistered migrants are a larger share of source and origin area population for counties; and, there is no backwards correction of resident population estimates for counties, unlike for provinces (Gibson and Li 2017). The second problem is the fundamental difference between urban and rural housing. The right to use rural residential land (*nongcun zhaijidi shiyongquan*) is evenly distributed and free of cost for village collective members, and rural dwellings are self-funded, self-built and self-renovated (Liu 2010), so housing is fairly cheap in rural counties. In contrast, county-level cities and urban districts (the equivalent 3rd sub-national level as counties), have commercial housing, purchased in the market at high cost due to urban land prices. Therefore articles that study county-level inequality, such as Li and Fang (2014) will overstate spatial inequality because they do not adjust for the cost-of-living differences coming from the fundamentally different nature of urban and rural housing.

The second literature we searched was newspaper articles written in Chinese, using the 'News' tab in *Google* and in *Baidu* (the second-largest search engine in the world). The phrase searched was: '中国区域发展不平衡' (Zhongguo quyu fazhan bupinghen), which translates as 'regional inequality in China'. We could only search back to 2003, and *Baidu* only found about one-tenth of what *Google* found, so we use separate Y-axes to show the results from both search engines in the one chart (Figure 1b). The trend is the same, with both search engines showing a big rise in news articles in China related to regional inequality; for *Google* they go from below 1000 articles in 2009 to almost 10,000 articles in 2016. In *Baidu*, it is from below 100 articles in 2012 to almost 900 articles in 2016, with 2013 a turning point that may reflect the timing of President Xi Jinping's two speeches that year to launch the Belt & Road Initiative (BRI). This initiative can be thought of as a form of regional development policy for western China (Gibson and Li 2018) so news about BRI often mentions regional inequality.

The government views on regional inequality can be seen in speeches by President Xi, such as in his report to the 19th National Congress of the Communist Party of China:[2]

> '…Some acute problems caused by unbalanced and inadequate development await solutions; … There are still large disparities in development between rural and urban areas, between regions, and in income distribution; …'
> '…We will devote more energy to speeding up the development of old revolutionary base areas, areas with large ethnic minority populations, border areas, and poor areas. We will strengthen measures to reach a new stage in the large-scale development of the western region; deepen reform to accelerate the revitalization of old industrial bases in the northeast and other parts of the country; … we need to put in place new, effective mechanisms to ensure coordinated development of different regions. …'

This speech, and many others by China's policy makers, reveals a belief that development has been regionally unbalanced. Despite the benefits that spatial concentration of economic activity brings, this speech also shows the intent to direct resources to apparently lagging regions. In the next section we show that, contrary to this view, China's spatial inequality has declined considerably during the reform era. Once account is taken of cost-of-living differences coming from the housing market, there is even less spatial inequality.

3. Trends in Regional Inequality: Population Counting Effects

For China's provinces, prefectures, counties, and districts, widely used annual population data are often for people with household registration (*hukou suozaidi*) from that place, not for people living there. This hardly mattered in the centrally planned era when freedom of movement was limited. Where people registered (a *de jure* criteria) was a good approximation to where they resided. However, in the reform era the number of non-*hukou* migrants residing away from their place of registration has grown from only a few million in the 1980s to over 250 million by 2015 (Guan 2015).

This counting issue affects apparent trends in spatial inequality. For most of the reform era in most sub-national units, local GDP was divided by the count of local *hukou* registrations instead of the count of local residents.[3] The registration count does not match the population producing and consuming local GDP, and the mismatch grew with the rising tide of migration.

[2] For the full report on this speech, delivered on 18 October 2017, see http://www.chinadaily.com.cn/interface/flipboard/1142846/2017-11-06/cd_34188086.html

[3] Some provincial statistics offices (for example, Jiangsu) switched to denominating provincial GDP per capita by estimates of the resident population as early as 1990. More populous provinces, like Guangdong (2001), Henan (2003) and Sichuan (2004) did not switch until much later (Li and Gibson 2013). For sub-provincial units, such as counties, what annual yearbooks label as a resident population is often a *hukou* registered population, as shown by the discrepancy between what is actually counted in census years (on a resident basis) and what is reported as estimates of the resident population for the same place in adjacent years (Gibson and Li 2017).

Coastal provinces have millions more residents than their registered population and the reverse holds for migrant-sending inland provinces (Gibson and Li 2017). Under this counting approach, measures of regional inequality rose as each person moved from China's interior to the coast because the migrant contributed to the numerator of GDP per capita in the coastal destination but was counted in the denominator of the origin, interior, area. Notwithstanding any true divergence between regions, gaps in GDP per capita between provinces (and even more so for the gaps for smaller units) rose in a mechanical way due to these counting and denominator issues.

The need to correctly denominate local GDP was brought to light by Li and Gibson (2013) but claims of rising regional inequality in China had already (and continued to be) published, based on using misleading counts of *hukou* registrations as denominators (Gibson and Li 2017). To show effects of this issue on apparent trends in regional inequality, Figure 2 has three time-series for the Theil Index of inter-provincial inequality from 1978 to 2016. The numerator for each series is based on annual provincial GDP, as reported by NBS (2005) for 1978-99 and in the *China Statistical Yearbook* (NBS 2001-17) each year since then (note these values are not spatially deflated, which we turn to next). The population denominators are:

- Registered population: The year-end registered population for each province, reported in various editions of what is now called the *China Population and Employment Statistical Yearbook* [*CPESY*] (NBS 2007-17). These data are originally from the Ministry of Public Security (the police, because Chinese are meant to register at local police stations when they move but many now do not).

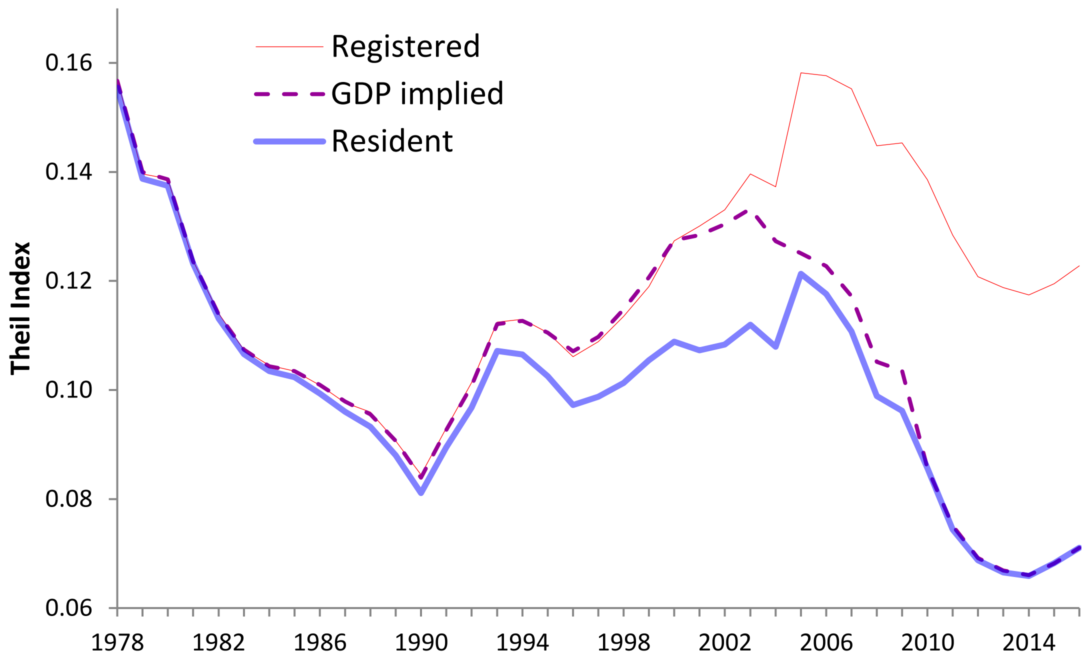
- GDP implied population: The population implied by the reported GDP per capita data, derived by dividing provincial GDP by provincial GDP per capita, where both series are published by the NBS for each province.

- Resident population: In 2011, the NBS used a trend-deviation interpolation with the 2010 population census and 2008 economic census, to backdate estimated GDP, the resident population, and GDP per capita on a resident population denominator basis.[4] This was initially for 1990 to 2010, but has been updated to 2016 in the *CPESY* (NBS, 2012-17) and we use the same approach to extend it back to 1978.

[4] The censuses give corrected endpoints for recalculating trends in GDP and GDP per capita. This 'new trend' helps revise previous annual estimates, with annual deviations from the 'historical trend' (based on prior data that did not benefit from the latest census) applied to the new trend. Wu (2007) provides details for GDP.

Figure 2 shows that initial rural reforms yielded a 12-year decline in inter-provincial inequality. The Theil Index fell to just over 0.08 in 1990, from its starting value of 0.16 in 1978. A key reason for falling inequality was that the Household Responsibility System (the return to family farming) helped to greatly raise rural incomes, and let poorer, mainly rural, provinces close some of the gap with the richer urban provinces (and especially with the province-level municipalities – Beijing, Shanghai and Tianjin). The choice of population denominator has no effect on this initial trend, because there were few non-*hukou* migrants prior to 1990.

Figure 2: Theil Indexes for Inter-Provincial Inequality in GDP Per Capita:
Comparing Three Population Denominators



If GDP is denominated by the number of *hukou* registrations from a province, the fall in inequality due to the rural reforms seems to fully reverse, with Figure 2 showing the Theil Index returning to a value of 0.16 by 2005. One reason inequality seems to rebound when using this denominator is that each migrant from a poorer (for example, interior) to a richer (for example, coastal) province mechanically raised measured inequality because they produced and consumed GDP in the destination but still counted in the denominator of the origin. Until recently, GDP per capita data reported for each province by the NBS used this approach, so the 'GDP implied' line in Figure 2 exactly tracks the line for using the registered population denominator, from 1978 until it starts to diverge in 2001. Analyses of regional inequality that use GDP per *hukou* registered population (and published data on GDP per capita mirrored this before 2001) thus suggest a rapid rise in inequality from 1990 onwards.[5]

[5] See Li and Gibson (2013) for a review of 14 studies, almost all of which took the reported data on per capita GDP at face value and interpreted these as if they were for GDP per resident.

The actual situation, based on denominating provincial GDP by residents – regardless of where registered – is rather different, as seen by the thick blue line in Figure 2. One-third of the initial fall in the Theil Index was reversed over the 1990-93 period, for the next 11 years the trend in inter-provincial inequality was fairly flat, followed by a one-year rise in 2005 and then a sharp fall until 2014. Over the full 38-year period, the Theil Index fell from 0.16 to 0.07, with a trend annual rate of decline of 1.1 percent.[6] Despite attention in the literature and media coverage of rising regional inequality in China (as seen in Section 2) the only sustained episode where inter-provincial inequality rose in the reform era was from 1990 to 1993, representing just three out of 38 years shown in Figure 2.

It is possible that the uncoordinated and unreported transition made by the NBS, from reporting GDP per *hukou* registered population to GDP per resident, creates an impression of rising regional inequality and also may distort understanding of policy effects. The dashed red line in Figure 2 shows the time-series for inequality in usually reported GDP per capita; from 2000 to 2010 this separates from the time-series based on the registered population and moves to the time-series based on the resident population. This movement has two effects: the rebound in inequality seems larger, appearing to reverse two-thirds of the initial fall in inequality; and, a sharp reversal in the inequality trend appears soon after various regional development schemes like the West China Development Project were instigated. In terms of the first effect, the Theil Index of 0.13 in 2003 with the GDP implied denominator is 20 percent above the appropriate value based on GDP per resident. For the second effect, some claimed success of regional development initiatives in reversing rising inequality (Fan *et al.* 2011) may just be from changes in how China's economic statistics used local population data because the actual trend in inter-provincial inequality (based on a resident denominator) was fairly flat when these policies were implemented. The subsequent decline in regional inequality after 2005 may reflect other factors, like rising wages in coastal areas (Li *et al.* 2012) driving mobile production inland, and China reaching the 'Lewis turning point' (Zhang *et al.* 2011) where labour shortages also cause rural wages in the interior to rise.

4. Real Inequality, Using Spatial Deflators from Housing Costs

The trends in inter-provincial inequality shown in Section 3 do not account for cost-of-living differences between areas. We need four types of data to develop and apply housing-related spatial deflators: GDP, resident population, a budget share for housing, and housing costs. Various statistical yearbooks report GDP for provinces, prefectures, counties, and districts. Despite claims, there is little evidence of falsified GDP data (Holz 2014) so we take the GDP data at face value. The implied resident population for merged districts can be derived from the

[6] This time trend is precisely estimated. The *t*-statistic for rejecting the hypothesis of zero trend is 5.1 (or 3.6 if using standard errors robust to unknown forms of heteroscedasticity and autocorrelation with a 1-year lag).

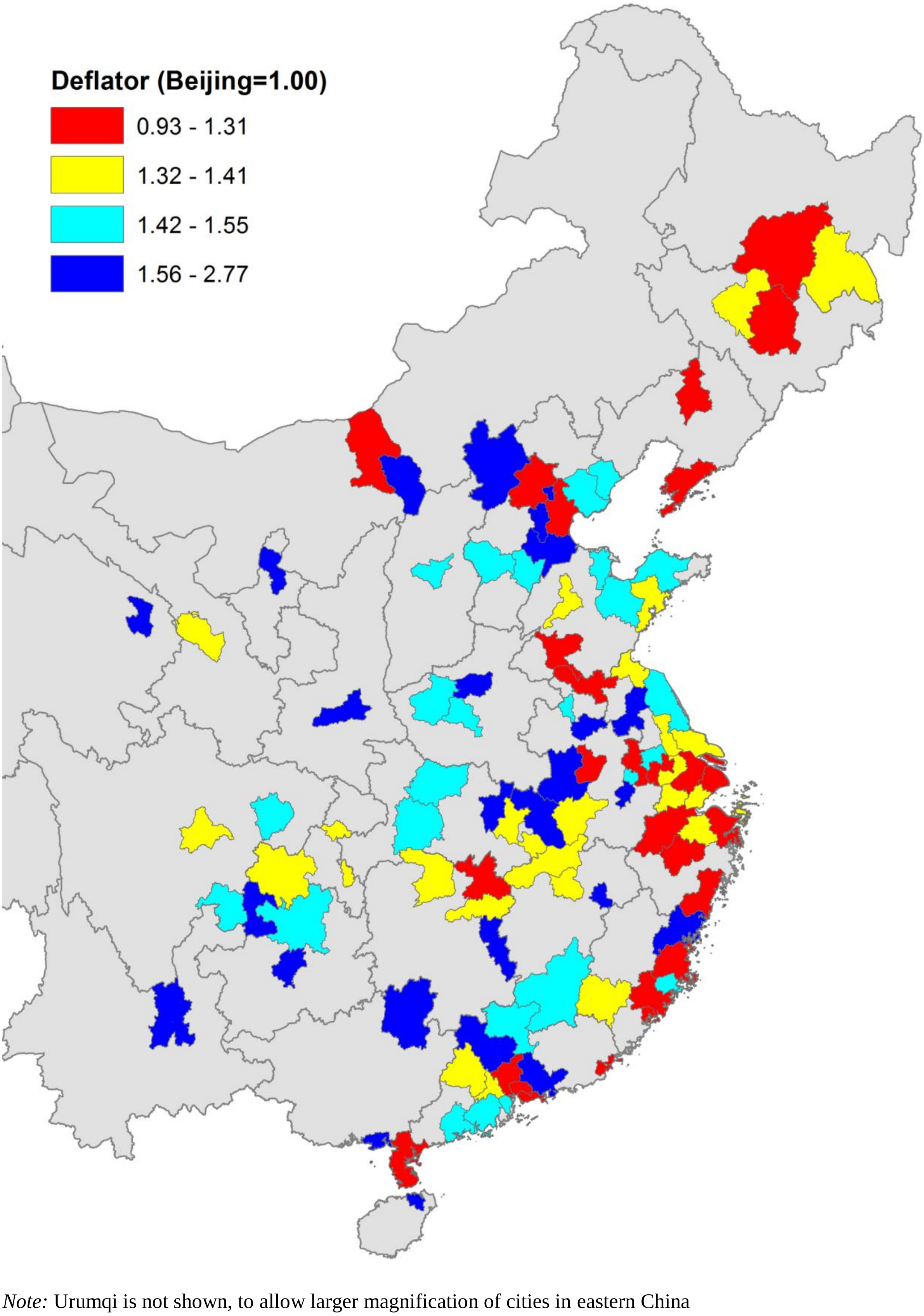
GDP per capita and GDP data reported in the 2017 *China City Statistical Yearbook* and our cross-checks with the census support the reported resident numbers.[7] Big cities like Beijing and Shanghai each have 16 districts, medium-sized ones like Xi'an have 11, and smaller cities like Kaifeng have five, so merged districts are effectively an urban core. We ignore counties; these are mainly rural areas without a residential land market and their reported population is mostly still on a *de jure* rather than a resident basis. The same *Yearbook* gives a budget share for housing, as the ratio of spending on residential real estate to total consumption spending; this averages about one-quarter (the median is 24 percent, the mean is 27 percent).

The housing costs are from the China Real Estate Index System (CREIS); a subscription service whose data are from city housing bureaus and real estate developers (many of whom have national scope) in about 130 major cities. Sales data for each new apartment in a complex are aggregated to district monthly averages of price per square metre. For example, in February 2018 when our data ends, new apartments in Shenzhen sold for RMB 54,200 (USD $8550) per square metre on average while in Shantou, which is a less popular city in Guangdong, the prices were one-sixth as high, at RMB 9038 (USD $1425). The overlap of the CREIS data with the *City Yearbook* gives us 104 cities, whose combined population is 360 million residents in 2016 (compared to 270 million registered for these cities). The location of these cities is shown in Figure 3, but to make it easier to visualize we show the area of the entire prefecture that the city is part of (the median ratio of merged districts area to prefectural land area is just 21 percent).

Our data on housing costs, and on budget shares for housing, are based on new housing units. This is the largest part of the urban housing market in China, partly because rising wealth and rising expectations of required amenities sees even fairly new apartments (say, 20 years old) demolished and rebuilt to meet the needs of richer buyers (Li and Gibson 2014). While actual and imputed rents are used in some countries for measuring housing costs, renters are a small share of the market in China and there are no disaggregated data on rents available from the NBS or from CREIS, unlike the situation for new builds. Also, some other countries (for example, New Zealand) use new housing sales to measure the net increase in the stock of owner-occupied housing, when setting the housing services weight for the CPI. This reflects the change in the price of housing acquired by the owner-occupier segment of the household sector during the reference period, which is analogous to the approach that we use.

[7] One exception is Chongqing where several districts are recently upgraded counties, which matters because county population data still refer to the registered population, outside of census years (Gibson and Li 2017). Chan and Wan (2017) also note that, over time, the merged districts are the best measure of functional urban areas, although recent conversion of rural counties to districts creates some complications.

Figure 3: Housing-Related Cost-of-Living Spatial Deflators for Major Cities in China in 2016



Note: Urumqi is not shown, to allow larger magnification of cities in eastern China

For each city, we calculate the 2016 average of new apartment prices (which is the latest year with GDP data), express this relative to the price in the numeraire city (Beijing) and then weight the logarithm of the price relativities by the average of the housing budget shares in Beijing and the city the under study. This gives a Törnqvist cost-of-living index:

$$T = \exp\left[\sum_{j=1}^{J}\left(\frac{S_{kj}+S_{ij}}{2}\right)\ln\left(\frac{P_{ij}}{P_{kj}}\right)\right]$$

where P is for price, S is for budget shares, and the subscripts are j for the budget item (housing or non-housing), i for the i^{th} city and k for Beijing. The Törnqvist is a superlative price index that closely approximates the true cost-of-living index, for any arbitrary utility function. The value T calculated for each city is the amount by which nominal GDP per capita should be multiplied to express it in Beijing prices. For Shanghai and Shenzhen, the values are less than one because their cost-of-living is even higher than in Beijing, so nominal data overstate how rich they are. For everywhere else, the values are greater than one, due to lower cost-of-living than in Beijing, with a mean (median) of 1.46 (1.41). In other words, nominal GDP per capita outside of Beijing, Shanghai and Shenzhen needs to be inflated by at least 40 percent, on average, to make it comparable to GDP in Beijing prices.

The values of the deflator are mapped in Figure 3, dividing the cities into four quartiles.[8] To help distinguish each city, the map magnification is increased by omitting western China; this region has just one city in the sample, Urumqi in Xinjiang, where the deflator is 1.36. The cities with the highest cost of living, shown in red, are mainly along the coast. Cities in the interior provinces tend to be in the two lowest quartiles, with deflators of 1.42 and above. Thus, conclusions about coastal-inland differences, and inter-provincial differences, will be sensitive to whether or not GDP data are deflated to recognise the large variation in the cost-of-living in different areas of China. Likewise, even if a researcher uses survey data on income (or on total personal or household expenditure as an income proxy) to calculate inequality statistics, these cost-of-living differences across space should be accounted for, especially as they are likely to have increased over time with the development of the urban residential land market.

The analysis in this section is designed to see how much difference is made to estimates of spatial inequality when using deflators derived just from variation in housing costs. The results are summarized in Table 1, for the Theil index, and for the decomposition of this into within-province and between-province components. The nominal values that are deflated are GDP per resident in 2016, taking account of the denominator issues described in Section 3. If

[8] Actual values for each city, together with their GDP and population data, are reported in the Appendix Table 1.

spatial variation in the cost of living is ignored, the level of spatial inequality is overstated by a statistically significant 27 percent (for inter-city analysis). The overstatement is mostly from the between-province component, where not deflating causes an upward bias of 34 percent.

Table 1: Measures of Nominal and Real Inequality in terms of 2016 GDP per Resident

	Nominal (a)	Spatially Deflated (b)	% overstatement [(a÷b)-1]×100
Theil Index for City-level GDP per capita	0.450	0.356	27%
95% Confidence Interval[a]	[0.365-0.548]	[0.277-0.430]	
Within-Province Theil Index	0.119	0.108	10%
Between-Province Theil Index	0.331	0.247	34%

Notes: Author's calculations using GDP per resident from 2017 *China City Statistical Yearbook* and the spatial deflator described in the text.

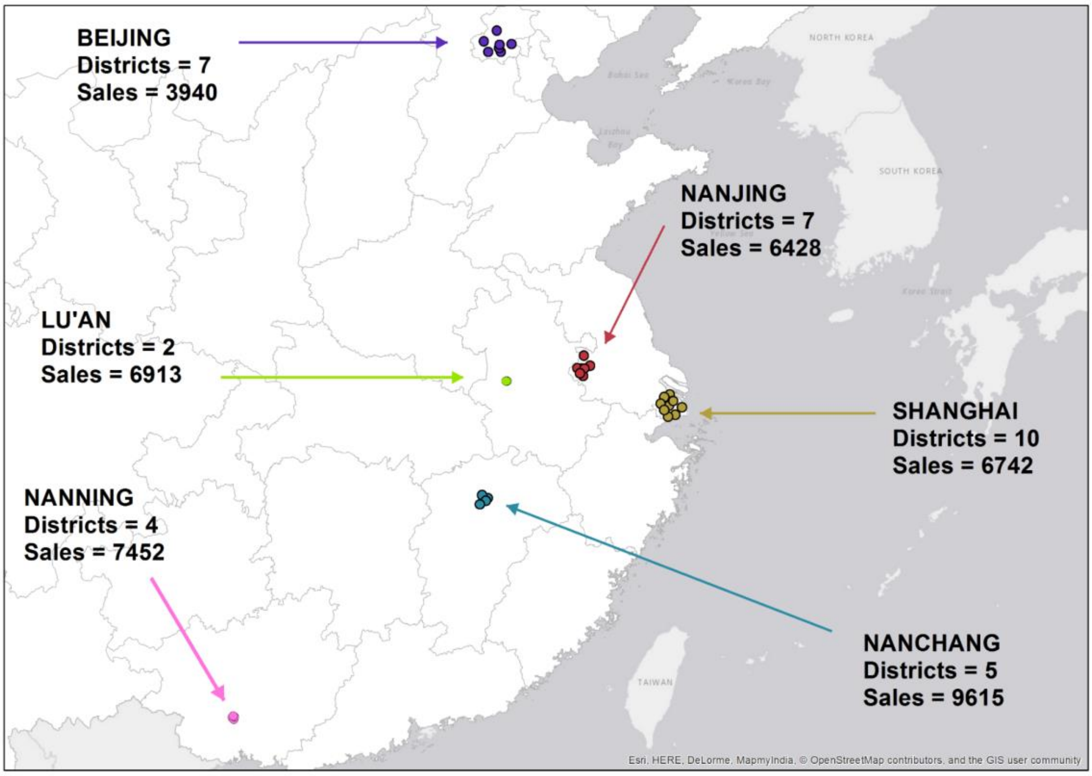
[a]Bias-corrected, based on 100 bootstrap replications. The Theil Index value in column (a) is statistically significantly larger (p <0.01) than the one in column (b). The Lorenz curves for nominal and spatially deflated data are significantly different (at 98 of the 100 percentiles), so any inequality statistic will be significantly different for the nominal and real data.

5. Sensitivity Analysis: Hedonic Adjustment of Housing Costs

A concern with the spatial deflators, and resulting inequality statistics, reported in Section 4 is that city-level average prices may mislead if apartment quality varies systematically between cities, as would happen if cheaper cities have lower quality apartments. This is unlikely because apartment complexes in different cities are developed by the same national-level real estate development companies (sometimes even using the same names for their complexes in each city). While each complex may have dozens of multi-story towers, each containing more than 50 individual units, there is limited within-complex variation in floor plans and in price per square metre. However, there is considerable variation in selling price between complexes in different areas, driven primarily by differences in the price of urban residential land.

In this Section we report on a hedonic analysis of 41,000 individual apartment sales, occurring between April 2016 and February 2018. Data on these sales are from a different part of the CREIS database, and are from 35 districts that we selected from six cities to represent different points in the urban hierarchy. Figure 4 shows the location of these cities, the sample size for number of districts selected and the number of sales for which we have data. Beijing and Shanghai are so-called 1st Tier cities; Nanjing is at the top of the 2nd Tier of cities, and is like cities such as Dalian, Qingdao, and Wuhan that also are capitals of large provinces; Nanchang and Nanning are lower 2nd Tier cities that are capitals of smaller provinces; and, Lu'an is a 3rd tier city that has just two urban districts and a population of about 1.6 million.

Figure 4: Sample Locations and Details for the Hedonic Analysis



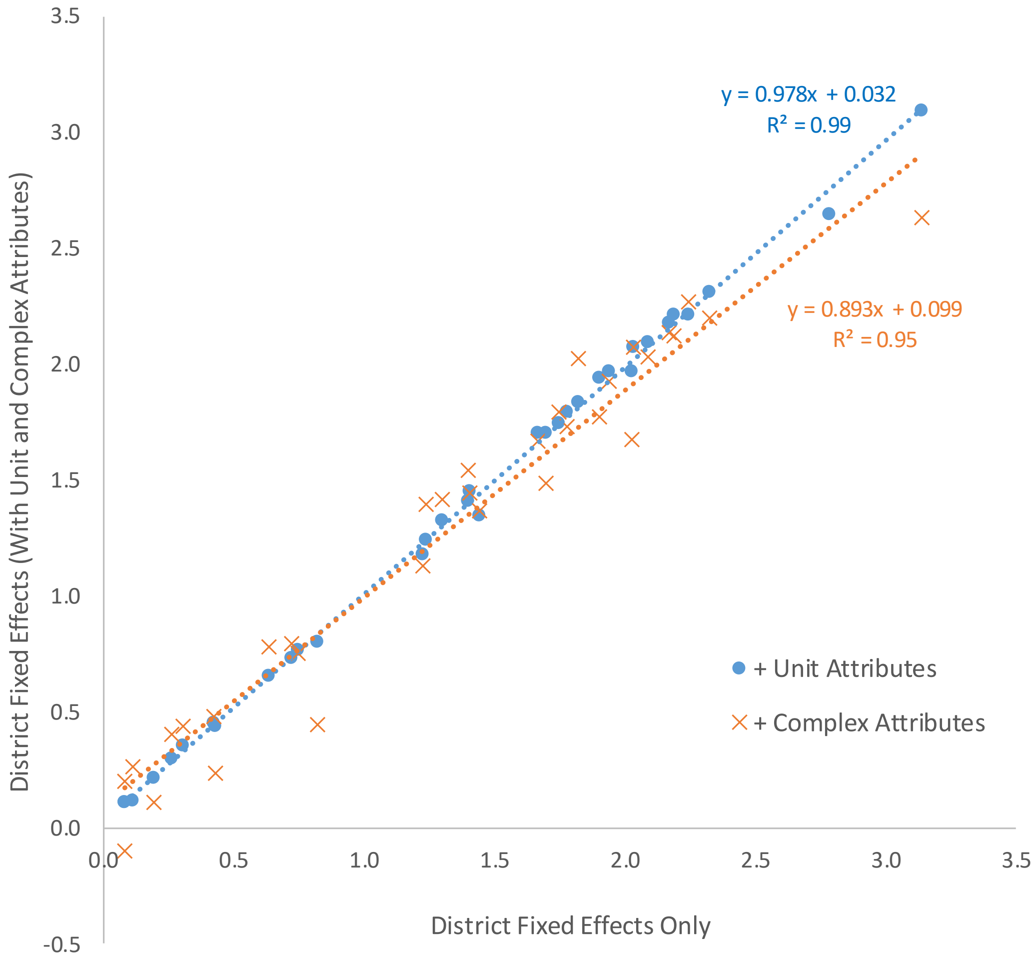
We start with a regression of (log) price (yuan per square metre) on district fixed effects. To adjust for temporal effects, where sales nearer February 2018 would be in nominally higher terms, we also include a quartic in time, as a flexible way to deal with any monthly price differences. The adjusted-R^2 from this regression is 0.957, which indicates that almost all of the variation in selling price for these apartments is due to district-level variation.[9] If we add attributes of individual units, in terms of (log) total area, number of rooms, whether the unit is multi-level, and the floor level (that is, height) in the tower, there is almost no change in either the degree of fit (the adjusted-R^2 is 0.958) or in the pattern of district fixed effects. Specifically, the correlation between the two sets of fixed effects is 0.999 and a scatter plot of the fixed effects from the regression without unit attributes controlled for, and from the regression with them controlled for has a slope of 0.98 (as seen by the blue points and trend line in Figure 5). In other words, the inter-district price relativities are almost the same, with or without characteristics of the individual apartment units controlled for.

If the characteristics of the apartment complex are then added, along with the district fixed effects and the unit attributes, there is a slight increase in fit (the adjusted-R^2 is 0.965) and a slight fall in the gradient of the district effects. Specifically, the trend line for the scatter plot of the district fixed effects for the model without other covariates against those from the model that adds unit and complex attributes has a slope of 0.89 (shown by the orange crosses

[9] Without the time effects the adjusted-R^2 is almost as high, at 0.941.

in Figure 5) and an R^2 of 0.95. The complex attributes included are: whether the complex sells units as empty shells to be fitted out by the owner or instead as pre-fitted; the complex age, in terms of years from when construction first began until sale data; the percentage of green area in the complex grounds; the land area of the complex; the total floor area of the complex; and, the total number of housing units in the complex.

Figure 5: Relationship Between District Fixed Effects for Housing Prices
With and Without Hedonic Adjustment for Attributes of the Unit and Complex



The full regression results, for the regression just with district fixed effects, for the regression that then adds attributes of individual units, and for the one that further adds complex attributes, are reported in Appendix Table 2. The standard errors are clustered at the district level to represent the way that we gathered the sample of individual apartment units. Based on these regressions, we can see that almost all of the variation in the cost of a unit of housing (in terms of one square metre of apartment space) in China reflects where the apartment is located and not the non-location attributes of the apartment. Consequently, the main analyses reported in Section 4, that use average apartment prices by city, should still be valid, given that hedonic adjustment would be unlikely to change the spatial patterns if we had individual apartment sales data from all 104 cities.

6. Conclusions

Many policy makers appear concerned by spatial inequality, despite benefits that come from concentrating rather than dispersing economic activity. These concerns are very apparent in China, where policy makers want to coordinate the development of various regions, and where both academic literature and media have paid increased attention to regional inequality. A typical view in the literature is that, at least until recently, the reform era has seen a rise in regional inequality, along with a rise in inter-household and urban-rural inequality (Wang et al, 2014). It is important to scrutinize evidence influencing such views, because interventions designed to reduce regional inequality – which typically divert resources to places that the market otherwise would not – can be very costly.

In this regard, there are at least two problems with studies of China's regional inequality. First, many studies use sub-national GDP per capita data denominated by the wrong population, being based on the locally registered population not the locally resident population. As the number of internal migrants rose from only a few million at the start of the reforms to almost 300 million now, a spurious gap emerged in reported per capita GDP for source and destination areas. We show in Section 3 that the entire reduction in inter-provincial inequality over the first 12 years of reforms – due largely to effects of the Household Responsibility System – would wrongly seem to be reversed by 2005 if GDP per registered population is used as an indicator. In fact, apart from a short, partial, reversal over 1990-93 the dominant trend in the reform era has been for regional inequality, in terms of GDP per resident, to fall.

The second problem is that many studies do not adjust for cost-of-living differences when they try to measure inequality. Spatial price differences within China are likely to have grown over time, as markets for private urban housing replaced the previous state provision of housing. In this paper we develop housing-related spatial deflators for 104 major cities and we find that spatial inequality in 2016 is overstated by 27% if nominal GDP per resident data are used instead of spatially deflated data that reflect cost-of-living differences. We lack the data for calculating spatial deflators at the beginning of China's reform era, but spatial price differences back should be smaller, given the widespread public provision of housing at the time. Therefore, the trend rate of decline in regional inequality should be even stronger in real terms. This positive outcome from China's reform era should be more widely recognized.

References

Brandt, L., and Holz, C. (2006). Spatial price differences in China: Estimates and implications. *Economic Development and Cultural Change*, *55*(1), 43–86.

Chan, K. W., and Wan, G. (2017). The size distribution and growth pattern of cities in China, 1982–2010: Analysis and policy implications. *Journal of the Asia Pacific Economy*, *22*(1), 136-155.

Fan, S., Kanbur, R., & Zhang, X. (2011). China's regional disparities: Experience and policy. *Review of Development Finance*, *1*(1), 47-56.

Gibson, J., Le, T., and Kim, B. (2017) Prices, Engel curves, and time-space deflation: Impacts on poverty and inequality in Vietnam. *The World Bank Economic Review*, *31*(2), 504-530.

Gibson, J., and Li, C. (2017). The erroneous use of China's population and per capita data: a structured review and critical test. *Journal of Economic Surveys*, *31*(4), 905-922.

Gibson, J., and Li, C. (2018). The 'Belt and Road Initiative' and comparative regional productivity in China. *Asia & the Pacific Policy Studies*, *5*(2), 168-181.

Gong, H., and Meng, X. (2008). Regional price differences in urban China 1986–2001: Estimation and implication. *IZA Discussion Paper* No. 3621.

Guan, X. (2015). How to better support China's migrant population. *Paulson Policy Memorandum* http://www.paulsoninstitute.org/wp-content/uploads/2017/01/PPM_ Migrants_Guan-Xinping_English_R.pdf

Holz, C. (2014). The quality of China's GDP statistics. *China Economic Review, 30*, 309-338.

Lai, H. (2002). China's Western Development Program: Its rationale, implementation, and prospects. *Modern China*, *28*(4), 432–453.

Lan, Y., and Sylwester, K. (2010). Does the Law of One Price hold in China? Testing price convergence using disaggregated data. *China Economic Review*, *21*(2), 224–236.

Li, G., & Fang, C. (2014). Analyzing the multi-mechanism of regional inequality in China. *The Annals of Regional Science*, *52*(1), 155-182.

Li, C., and Gibson, J. (2013). Rising regional inequality in China: Fact or artefact? *World Development*, *47*(1), 16-29.

Li, C., and Gibson, J. (2014). Spatial price differences and inequality in the People's Republic of China: Housing market evidence. *Asian Development Review*, *31*(1), 92-120.

Li, H., Li, L., Wu, B., and Xiong, Y. (2012). The end of cheap Chinese labor. *Journal of Economic Perspectives*, *26*(4), 57-74.

Liu, L. (2010). Development and outlook of farm house building in China. *Urban Studies*, *17*(1), 128–133

Moulton, B. (1995). Inter-area indexes of the cost of shelter using hedonic quality adjustment techniques. *Journal of Econometrics*, *68*(1), 181–204.

National Bureau of Statistics (2005). *China Compendium of Statistics 1949-2004 (Xin Zhongguo Wushiwunian Tongji Ziliao Huibian)*. Beijing: China Statistics Press.

National Bureau of Statistics (2001-2017). *China Statistical Yearbook 2001-2017 (Zhongguo Tongji Nianjian)*. Beijing: China Statistics Press.

National Bureau of Statistics (1991-2006). *China Population Statistical Yearbook 1991-2006 (Zhongguo Renkou Tongji Nianjian)*. Beijing: China Statistics Press.

National Bureau of Statistics (2007-2017). *China Population and Employment Statistical Yearbook 2007-2017 (Zhongguo Renkou he Jiuye Tongji Nianjian)*. Beijing: China Statistics Press.

National Bureau of Statistics (2017). *China City Statistical Yearbook 2017 (Zhongguo Chengshi Tongji Nianjian)*. Beijing: China Statistics Press.

Wang, C., Wan, G., & Yang, D. (2014). Income inequality in the People's Republic of China: Trends, determinants, and remedies. *Journal of Economic Surveys*, *28*(4), 686-708.

World Bank (2009). *World Development Report 2009: Reshaping Economic Geography*. Washington DC: World Bank Publications.

Wu, H. (2007). The Chinese GDP growth rate puzzle: How fast has the Chinese economy grown? *Asian Economic Papers, 6*(1), 1-23.

Young, A. (2000). The razor's edge: Distortions and incremental reform in the People's Republic of China. *Quarterly Journal of Economics, 115*(4), 1091–1135.

Zhang, X., Yang, J., and Wang, S. (2011). China has reached the Lewis turning point. *China Economic Review*, *22*(4), 542-554.

Appendix

Table 1: Details on Spatial Deflator, GDP and Resident Population for each City in 2016

City	Province	Deflator	GDP (RMB billion)	Population (mil)
Beijing	Beijing	1.000	2566.913	21.717
Tianjin	Tianjin	1.249	1788.539	15.545
Shijiazhuang	Hebei	1.476	321.483	4.763
Tangshan	Hebei	1.436	332.383	3.570
Qinhuangdao	Hebei	1.534	93.400	1.644
Zhangjiakou	Hebei	1.560	66.528	1.609
Cangzhou	Hebei	1.791	71.043	0.705
Langfang	Hebei	1.872	90.313	0.857
Hengshui	Hebei	1.544	39.981	0.949
Taiyuan	Shanxi	1.474	275.494	3.537
Hohhot	Inner Mongolia	1.588	236.575	2.168
Baotou	Inner Mongolia	1.306	348.506	2.281
Shenyang	Liaoning	1.291	492.257	6.921
Dalian	Liaoning	1.184	385.569	5.025
Changchun	Jilin	1.350	470.701	4.926
Jinlin	Jilin	1.293	144.278	1.960
Harbin	Heilongjiang	1.238	447.269	6.671
Mudanjiang	Heilongjiang	1.325	34.791	0.968
Shanghai	Shanghai	0.987	2817.865	24.175
Nanjing	Jiangsu	1.134	1050.302	8.253
Wuxi	Jiangsu	1.332	474.902	3.627
Xuzhou	Jiangsu	1.304	307.218	3.254
Changzhou	Jiangsu	1.303	498.566	3.944
Suzhou	Jiangsu	1.194	800.845	5.501
Nantong	Jiangsu	1.346	247.503	2.344
Lianyungang	Jiangsu	1.372	130.759	2.083
Huaian	Jiangsu	1.562	206.233	3.049
Yancheng	Jiangsu	1.446	188.464	2.352
Zhenjiang	Jiangsu	1.444	171.160	1.230
Taizhou	Jiangsu	1.389	171.827	1.624
Hangzhou	Zhejiang	1.190	983.548	7.296
Ningbo	Zhejiang	1.230	557.466	4.106
Wenzhou	Zhejiang	1.114	205.472	4.327
Jiaxing	Zhejiang	1.405	97.041	1.235
Huzhou	Zhejiang	1.335	101.117	1.324
Shaoxing	Zhejiang	1.346	280.184	2.732
Jinhua	Zhejiang	1.279	69.313	1.121
Zhoushan	Zhejiang	1.340	90.817	0.878
Hefei	Anhui	1.307	419.170	3.776
Wuhu	Anhui	1.646	165.978	1.648
Bengbu	Anhui	1.759	74.525	1.143
Maanshan	Anhui	1.488	91.025	0.949
Huaibei	Anhui	1.450	54.581	1.051
Anqing	Anhui	1.401	45.689	0.818
Liuan	Anhui	1.807	48.164	1.967
Fuzhou	Fujian	1.191	314.428	3.095
Xiamen	Fujian	1.071	378.427	3.890
Putian	Fujian	1.547	148.411	2.024
Quanzhou	Fujian	1.194	149.027	1.540
Longyan	Fujian	1.339	90.687	1.081
Ningde	Fujian	1.558	30.019	0.445
Nanchang	Jiangxi	1.341	330.059	3.549
Jiujiang	Jiangxi	1.367	79.147	0.711
Yingtan	Jiangxi	1.705	21.573	0.220
Ganzhou	Jiangxi	1.452	52.079	1.092
Jinan	Shandong	1.324	581.665	5.720
Qingdao	Shandong	1.349	650.596	4.936
Dongying	Shandong	1.491	242.505	1.301
Yantai	Shandong	1.470	294.772	2.319
Weifang	Shandong	1.473	158.610	2.128
Jining	Shandong	1.308	120.958	1.969
Zhengzhou	Henan	1.727	460.971	5.566
Luoyang	Henan	1.513	148.851	2.170
Pingdingshan	Henan	1.541	51.096	1.081
Wuhan	Hubei	1.361	963.060	7.676
Huangshi	Hubei	1.350	61.662	0.888
Yichang	Hubei	1.480	159.302	1.466
Xiangyang	Hubei	1.483	185.882	2.304
Xiaogan	Hubei	1.895	28.774	0.924
Huanggang	Hubei	1.625	20.316	0.389
Changsha	Hunan	1.352	586.720	4.045
Zhuzhou	Hunan	1.869	117.625	1.225
Yueyang	Hunan	1.308	126.641	1.285
Changde	Hunan	1.361	147.195	1.565
Guangzhou	Guangdong	1.127	1954.744	13.772
Shaoguan	Guangdong	1.505	57.935	1.033
Shenzhen	Guangdong	0.930	1949.260	11.644
Shantou	Guangdong	1.286	206.385	5.504
Foshan	Guangdong	1.372	863.000	7.447
Jiangmen	Guangdong	1.501	130.341	1.867
Zhanjiang	Guangdong	1.273	114.562	1.678
Zhaoqing	Guangdong	1.342	110.031	1.537
Huizhou	Guangdong	1.639	208.421	2.445
Yangjiang	Guangdong	1.414	67.110	1.173
Qingyuan	Guangdong	1.760	74.239	1.569
Dongguan	Guangdong	1.180	682.769	8.258
Zhongshan	Guangdong	1.480	320.278	3.220
Guilin	Guangxi	1.573	84.724	1.543
Beihai	Guangxi	2.772	77.427	0.714
Haikou	Hainan	1.631	125.767	2.233
Chongqing	Chongqing	1.406	764.689	8.433
Chengdu	Sichuan	1.365	968.558	10.440
Luzhou	Sichuan	1.635	77.407	1.438
Nanchong	Sichuan	1.538	59.282	1.937
Yibin	Sichuan	1.413	66.159	1.202
Meishan	Sichuan	1.614	51.633	1.159
Guiyang	Guizhou	1.736	240.365	3.312
Zunyi	Guizhou	1.515	98.479	2.194
Kunming	Yunnan	1.564	333.614	3.989
Xian	Shaanxi	1.582	552.766	7.087
Lanzhou	Gansu	1.329	188.225	2.663
Xining	Qinghai	1.710	96.655	1.263
Yinchuan	Ningxia	2.259	97.337	1.396
Urumqi	Xinjiang	1.357	243.808	3.424

Table 2: Details on Hedonic Regression Models
for Sales Prices of Individual Apartments

	Mean [Std Deviation]	District and Time Effects Only	Adding Unit Attributes	Adding Complex Attributes
District Fixed Effects[a]				
Baoshan, Shanghai	0.045	2.087	2.094	2.028
	[0.208]	(65.97)**	(55.48)**	(20.71)**
Changping, Beijing	0.016	2.169	2.180	2.135
	[0.124]	(58.65)**	(50.98)**	(9.73)**
Daxing, Beijing	0.007	2.035	2.075	2.076
	[0.084]	(195.42)**	(60.89)**	(9.25)**
Fangshan, Beijing	0.023	1.402	1.407	1.540
	[0.149]	(108.21)**	(119.54)**	(7.72)**
Fengtai, Beijing	0.017	2.325	2.311	2.201
	[0.128]	(197.61)**	(175.71)**	(12.29)**
Fengxian, Shanghai	0.027	1.669	1.703	1.666
	[0.161]	(32.11)**	(27.29)**	(11.16)**
Gulou, Nanjing	0.006	1.698	1.703	1.489
	[0.074]	(32.79)**	(33.82)**	(8.59)**
Haidian, Beijing	0.017	2.244	2.213	2.272
	[0.129]	(195.75)**	(134.03)**	(11.88)**
Honggutan, Nanchang	0.094	0.422	0.452	0.482
	[0.291]	(14.78)**	(13.58)**	(6.33)**
Huangpu, Shanghai	0.003	3.140	3.095	2.630
	[0.054]	(219.35)**	(143.13)**	(9.38)**
Jiading, Shanghai	0.041	1.406	1.451	1.444
	[0.198]	(200.16)**	(52.99)**	(11.74)**
Jiangnan, Nanning	0.049	0.305	0.351	0.438
	[0.217]	(23.58)**	(10.57)**	(1.63)
Jiangning, Nanjing	0.005	1.226	1.178	1.127
	[0.067]	(71.07)**	(56.19)**	(2.84)**
Jinshan, Shanghai	0.009	0.746	0.766	0.755
	[0.092]	(40.31)**	(32.06)**	(7.69)**
Liangqing, Nanning	0.029	0.191	0.215	0.113
	[0.166]	(3.83)**	(4.13)**	(0.73)
Liuhe, Nanjing	0.086	0.428	0.441	0.234
	[0.280]	(63.29)**	(34.54)**	(0.71)
Mentougou, Beijing	0.008	2.026	1.967	1.674
	[0.090]	(163.87)**	(89.56)**	(3.89)**
Minxing, Shanghai	0.010	1.940	1.967	1.930
	[0.098]	(42.18)**	(33.75)**	(12.11)**
Nanhui, Shanghai	0.006	1.443	1.344	1.370
	[0.079]	(52.40)**	(36.87)**	(4.97)**
Pudong, Shanghai	0.004	2.786	2.648	2.148
	[0.063]	(245.57)**	(43.83)**	(5.04)**
Pukou, Nanjing	0.013	1.235	1.244	1.399
	[0.112]	(275.93)**	(70.61)**	(14.85)**
Qingpu, Shanghai	0.009	2.192	2.211	2.119
	[0.093]	(76.83)**	(59.14)**	(16.69)**
Qingshanhu, Nanchang	0.036	0.724	0.728	0.792
	[0.187]	(19.81)**	(19.78)**	(5.92)**
Qingxiu, Nanning	0.076	0.636	0.658	0.782
	[0.266]	(36.14)**	(32.60)**	(8.99)**
Songjiang, Shanghai	0.011	1.905	1.944	1.774
	[0.105]	(48.62)**	(35.27)**	(10.42)**
Tongzhou, Beijing	0.009	1.823	1.836	2.024
	[0.093]	(162.18)**	(97.36)**	(9.56)**

Table 2, continued

	Mean [Std Deviation]	District and Time Effects Only	Adding Unit Attributes	Adding Complex Attributes
Wanli, Nanchang	0.016	0.079	0.112	0.199
	[0.124]	(1.64)	(2.02)	(1.48)
Xihu, Nanchang	0.005	0.822	0.800	0.443
	[0.069]	(18.21)**	(18.25)**	(2.22)*
Xijian, Nanchang	0.084	0.259	0.297	0.406
	[0.277]	(14.12)**	(10.41)**	(3.47)**
Xixia, Nanjing	0.023	1.298	1.324	1.414
	[0.150]	(53.31)**	(37.02)**	(8.37)**
Xixiangtang, Nanning	0.027	0.081	0.106	-0.101
	[0.162]	(1.60)	(2.02)	(0.53)
Xuanwu, Nanjing	0.016	1.749	1.748	1.793
	[0.126]	(181.15)**	(117.55)**	(7.13)**
Yuan, Liuan	0.047	0.113	0.116	0.264
	[0.212]	(12.08)**	(10.58)**	(1.76)
Yuhuatai, Nanjing	0.009	1.779	1.793	1.734
	[0.092]	(97.46)**	(95.79)**	(9.15)**
Time Effects (Months from 3/16)				
Linear	11.590	0.016	0.018	-0.007
	[6.911]	(0.64)	(0.70)	(0.30)
Squared	182.085	0.006	0.005	0.009
	[173.306]	(1.45)	(1.33)	(2.20)*
Cubic	3275.326	-0.000	-0.000	-0.001
	[3943.375]	(2.20)*	(2.09)*	(2.73)**
Quartic	63188.970	0.000	0.000	0.000
	[88738.980]	(2.44)*	(2.37)*	(2.81)**
Individual Unit Attributes				
(log) floor area, square metres	4.657		0.191	0.145
	[0.242]		(2.19)*	(2.22)*
Floor level (i.e. height in tower)	13.779		-0.001	-0.001
	[9.556]		(0.79)	(0.91)
Number of rooms	2.823		-0.022	-0.038
	[0.710]		(1.03)	(1.99)
Multi-level unit	0.015		0.005	0.046
	[0.120]		(0.10)	(1.00)
Attributes of the Complex				
Units are fitted out, not shells	0.540			0.122
	[0.498]			(1.64)
Green space to total area ratio	0.359			1.203
	[0.062]			(0.64)
(log) complex land area	12.018			-0.138
	[0.785]			(0.55)
(log) complex floor area	13.128			0.230
	[0.829]			(0.99)
(log) total units in complex	7.974			-0.252
	[0.766]			(1.46)
Age of complex (years)	1.932			-0.033
	[1.701]			(0.87)
Constant		8.429	7.602	8.115
		(141.03)**	(18.49)**	(5.62)**
Adjusted-R^2		0.957	0.958	0.965

Notes

Based on sales prices, in yuan per square metre, for 41,094 apartments. The *t* statistics in parentheses are from robust standard errors clustered at district level; * statistically significant at 5%; ** statistically significant at 1% level. [a] Jinan in Liuan is omitted, as the reference category. This is the district with the lowest average prices.