Why is interregional inequality in Russia and China not falling?

Thomas F. Remington
Department of Political Science, Emory University, USA

Article info
Article history:
Available online 12 February 2015

Keywords:
Russia
China
Interregional inequality
Convergence
Institutional legacies

Abstract
In large and heterogeneous countries, cross-regional inequality often fuels political conflict over redistributive demands. Standard economic theory holds that initial inequality in incomes across subnational regions of a country should eventually give way to convergence. However, economic liberalization in Russia and China has not produced cross-regional convergence. The paper demonstrates that interregional inequality remains high in both countries and grounds explanation for this finding in their communist institutional legacies. Their common institutional history complements country-specific explanations for the observed trends in cross-regional inequality.

© 2015 The Regents of the University of California. Published by Elsevier Ltd. All rights reserved.

1. Introduction: the problem of interregional inequality

Although the global rise in interpersonal inequality has been the subject of a great deal of public and scholarly debate in the last decade, inequality across the subnational regions of countries is of no less political importance (Piketty, 2014; Milanovic, 2005, 2010; Atkinson and Piketty, 2010; Atkinson et al., 2011; Beramendi, 2012). Like interpersonal inequality, interregional inequality can generate redistributive conflict. Indeed, interregional inequality can often trigger political conflict more readily than can interpersonal inequality. Given the collective action problem inherent in mobilizing diffuse social constituencies across a national society, it is often less costly for elites to mobilize regional demands for redistribution of rights and resources based on a state’s existing territorial-administrative divisions.

Cross-regional tension over the distribution of resources and obligations can originate both in richer and poorer regions. Groups in worse-off regions may demand that the center distribute a greater share of the national pie to them, while richer regions resent their poorer neighbors’ claims on the national purse. Moreover, ethnic divisions and price controls can exacerbate regional conflict. If prices are subject to administrative regulation (for example, if oil or cash crops are procured by the central government at state-set prices in some regions and sold at below-market prices in others), it is easy to rouse public ire over an unfavorable system of exchange. This was a common perception throughout the Soviet Union, Yugoslavia, and Czechoslovakia in the 1980s. Publics in both richer and poorer republics were convinced that the federal government’s control of the economy put their own republic at a net disadvantage. In the USSR, Russians complained that the natural resources extracted from their republic were subsidizing the rest of the union while Uzbeks complained that the center had forced them into a quasi-colonial mono-cultural cotton economy (Roeder, 1991; Remington, 1985; Remington, 1989).

Similarly, if regional boundaries coincide with ethnic group attachments, then regional conflict can take on a primordial character (Horowitz, 1985). The congruence of ethnic-national and territorial-administrative boundaries meant that cross-regional inequality undermined the unity of all three of the communist ethno-federations, the USSR, Yugoslavia and Czechoslovakia (Bunce, 1999; Roeder, 1991). Similar tensions are common in many developed and developing states. The recent
tendency in the literature to emphasize the redistributive implications of democratization applies equally to the problem of cross-regional inequality (Acemoglu and Robinson, 2006; Boix, 2003). Therefore differential long-term rates of growth across different regions have major political implications.

As Pablo Beramendi has argued, large, heterogenous states often face a trade-off between policies that centralize fiscal and social policy in order to carry out redistribution at the individual level and policies that give different regions autonomy to devise their own fiscal and social policies. If richer regions in such a state fear the destabilizing implications of cross-regional migration from poorer to richer regions, they frequently endorse inter-regional redistributive policies in order to preserve their own autonomy. This expedient is likely to be endorsed by the leaders of poorer regions if it allows them to maintain control over their own labor markets and use the interregional transfers to maintain their own power. The more that the country's economic geography is uneven and its institutions of political representation push toward decentralization, the likelier it is that fiscal transfers across regions will substitute for, rather than complement, more centralized fiscal and social policy that aims at interpersonal redistribution (Beramendi, 2012). Although Beramendi's analysis pertains to democratic states, the political-economic logic he outlines also applies to large authoritarian post-communist states as well.

This paper compares trends in interregional inequality for two such post-communist states, Russia and China. Both have very diverse economic geographies and both are undergoing a market transition including integration into global product and capital markets, spread of private property relations, and use of market-clearing prices. These cases therefore lend themselves to evaluating theories about the dynamics of cross-regional differences in economic growth following a market transition. Standard economic theory holds that, following the end of central planning, an initial rise in cross-regional divergence would give way to a long-term convergence of growth rates. However, as the paper will show, we do not observe such convergence in Russia and China when we examine trends in mean incomes, output per capita, and wages at the level of the first-tier subnational administrative units in Russia (the 83 subjects of the federation) and in China (the 31 province-level units). The paper seeks to offer an explanation for this finding based on the shared legacy of the communist economic and political system. That is, while factors specific to the development of Russia and China are important, similar institutional features of the two states that are inherited from their state socialist past help explain barriers to interregional convergence.

What is the basis for the expectation that over time, interregional inequality will fall? The neo-classical assumption is that, given a reasonably well-integrated national market for capital and labor, mobile productive factors will tend to flow to regions where they can command a higher marginal return. Therefore following some initial technological or institutional shock, regional income inequality might rise initially, but would then fall. After the initial shock, inherited cross-regional differences in the distribution of transportation infrastructure, proximity to ports, natural resources, human capital, and administrative status would be expected to generate increasing returns to initial comparative advantages. Regions more richly endowed in such advantages would grow at a faster rate than poorer regions, widening cross-regional differences in incomes, wages, wealth, and output. However, both economic forces—the incentive on the part of the owners of capital and labor to seek a higher return as well as political pressures for redistribution through the central government—would eventually tend to act to redress inter-regional differences.

For example, Jeffrey Williamson argues that the process of industrialization may affect different parts of a country differently, initially increasing interregional differences as capital and labor flow to the higher-growth regions. Over time, however, the gap tends to close as factor prices equalize and political pressures mount for redistribution of public investment (Williamson, 1965). He posits an inverted-U shaped curve in cross-regional inequality over time, as regional differences initially widen and subsequently diminish in the course of national development. He found considerable evidence for such a pattern over a number of cases. In the case of the United States, Gavin Wright has shown that changes in federal labor laws, such as the minimum wage, that undermined the regional isolation of the Southern labor market, reinforced by the effects of World War II, brought about a rapid increase in Southern growth rates after the 1940s as labor and capital gravitated to the South (Wright, 1987). These institutional changes resulted from a shift in the relative power of local political coalitions that sought to maintain traditional rural and racial hierarchies (Olson, 1983).

The long-run tendency for regional growth differentials across and within national economies to fall has been termed “the iron law of convergence” (Barro, 1996a, 1996b; Barro and Sala-i-Martin, 1992; Barro et al., 1991). The theory recognizes that convergence works only over long time spans and can be disrupted by exogenous shocks. If national political institutions are weak, the central government will be unable to enforce laws uniformly, redistribute across regions, or invest in infrastructure, civil and judicial administrative capacity, and human capital. As a result, rates of development would diverge. Institutional restrictions on the mobility of capital or labor would reinforce cross-regional inequality. Therefore the real question is under what conditions institutional obstacles to national integration can be overcome by the pull of market forces. Douglass North's theory of institutional equilibria suggests that low-level equilibrium traps can persist over long periods of time (North, 1990).

The paper proceeds as follows. In the next section, I describe the different trajectories of economic liberalization in Russia and China since they began market reforms. Part 3 analyzes the data on regional-level growth, incomes and wages and graph and table form. To test standard economic theories of convergence, the paper first presents trends in the inter-quartile spread of the wage, income, and output series across regions over time. It then shows the results of OLS regressions of regional
growth rates on initial levels of output and income at the point market transition began. Part 4 proposes a theoretical explanation for the observed trends. Part 5 discusses the implications of the findings.

2. Russia and China in the reform era

Russia dismantled state socialism comprehensively in 1991, China in stages beginning in 1978, and then decisively with the CCP’s adoption of the “socialist market economy” concept in 1993. As is well known, economic growth trajectories differed markedly between the two countries following the reforms. Russia’s economy sank into a severe depression in the 1990s before recovering in the 2000s, while China’s economy grew much more rapidly and steadily. World Development Indicators figures imply that the average annual compound growth rate for GDP per capita (in PPP terms and constant 2005 dollars) was 1.1% for Russia from 1991 to 2013 and 8.8% for China from 1980 to 2013.2

The shift from a centrally planned, state socialist economy to one dominated by private property, market relations, and openness to world trade upset the distribution of economic and political advantage in both countries and contributed to a rapid increase in inequality in incomes.3 Total income inequality for both has reached levels comparable to those of the United States.4 China’s total level of inequality, in fact, is higher than that of Russia, based on official figures; adding in the distribution of gray incomes would make both China’s and Russia’s inequality estimates still higher. (Fig. 1) Geographic inequality rose as well in both countries, although, as we shall see, rising price differentials mask part of the effect.

The difference in the reform strategies followed by the two countries is an oft-told tale (Naughton, 1995; Shleifer and Treisman, 2000). Only a few salient points need to be underlined here. For the most part, China’s reformers proceeded incrementally, both sectorally and territorially. Nearly all major policy reforms were tested locally first before being adopted nationally (Xu, 2011). China’s early reforms liberalized agricultural production through the household responsibility system, which quickly increased output on household-leased farm land and freed labor for local processing industries (Lin, 1992). The center initially established four Special Economic Zones (SEZ) along the southeastern coast, later expanding the number and giving priority to the development of the coastal provinces generally. Subsequently, other cities and provinces won permission to adopt liberalized rules on investment, management and labor in order to attract overseas investment (Gallagher, 2002). The SEZ’s were concentrated along China’s southeastern coast in order to take advantage of the physical proximity, capital, and technological and managerial know-how available from diaspora communities in Hong Kong and Taiwan. As a result, investment, output, and incomes rose much more rapidly in the Pearl River Delta region and Guangdong than in inland provinces. Beijing also benefited from its position as national capital as did Beijing’s coastal neighbor, Tianjin. Regions rich in natural resources, such as Inner Mongolia, also saw a rapid influx of capital for the development of extractive industry. The rapid growth of the coastal provinces shifted China’s industrial center of gravity from the northeastern provinces which had been industrialized in the 1950s under the influence of the Soviet model. “Rustbelt” towns in the north experienced bankruptcies and layoffs in the 1990s, while “sunbelt” cities along the southeastern coast attracted capital investment and migrant labor from inland provinces (Lee, 2007). China’s strategy of “crossing the river by feeling the stones” took advantage of local experimentation to learn from trial and error. This had several advantages. It tended to reduce and divide potential bureaucratic opposition to reform by encouraging those regions and sectors that were net beneficiaries from the reforms to resist pressures to retrench on the part of party conservatives and state-owned industries (Shirk, 1993). It also allowed the central leaders to disseminate successful reforms. As Sebastian Heilmann has argued, Deng’s pragmatic approach to reform reinvented earlier CCP tactics of “point to surface” local experimentation and learning (Heilmann, 2011).

Russia’s reform strategy differed from China’s in many ways. Russia attempted to carry out a comprehensive restructuring of the economy through harsh fiscal and monetary measures, but was unable to implement them consistently (Gaidar, 1999, 2007; Shleifer and Treisman, 2000). The government’s actual control over state institutions, including regional governments, was extremely limited. For example, the central government never fully controlled the Central Bank, which continued to provide soft credits to favored state industries. Regional governments protected local industry. The powerful state industrial sector lobbied successfully for a relaxation of the most restrictive policies, while the reformers themselves refrained from fully liberalizing prices on energy or basic consumer goods. The government had only a limited ability to force industry to restructure production profiles or induce new entrepreneurial start-ups. The fact that Russia’s economy had been much more closely integrated through central planning than was China’s meant that the drastic cuts in state procurement orders and credits left many regions in an extremely difficult position. Poorly developed financial markets made it difficult for industries to retool their production lines to new markets. Many firms therefore used their influence with regional governors and allied ministers in Moscow to lobby for cheap credits and state orders in order to maintain familiar production lines. Many in fact simply ceased selling their output for cash and instead entered elaborate barter chains or settled their bills with “vekselia” (promissory notes, or bills of exchange). Arrears in wages, taxes and pensions mounted. The liquidity crunch

2 Calculated from WDI data at <http://databank.worldbank.org>,

3 In China, by far the greatest source of overall income inequality is associated with the high and rising gap between rural and urban incomes. Only 11–12% of total inequality is associated with differences in mean incomes across regions; the rest derives from the rural-urban gap and the gap in the share of income from private assets (Li, Sato, and Sicular, 2013).

4 Note that estimates of China’s Gini coefficient vary widely. Key issues concern valuation of social benefits, calculation of rural incomes and calculation of unreported cash incomes. According to one estimate, the large size of unreported incomes going to the highest-income strata mean that the actual decile ratio for China’s income is 65:1 rather than the officially-reported figure of 23:1 (Wang and Woo, 2010).
generated an inefficient “virtual economy” (Gaddy and Ickes, 1998, 2005). Only when the 1998 financial crash cut the ruble’s value to less than a third of its former rate did Russian production become competitive in world markets. The recovery was spurred as well by the beginning of a long rise in world oil prices and by Vladimir Putin’s accession to power. Putin succeeded in imposing a greater degree of political control over the bureaucracy than Yeltsin had managed to do.

The result was that Russian regions enjoying competitive advantages, above all in mineral resources, benefitted much more from the recovery in the 2000s than did regions still dependent on the legacy of Soviet manufacturing. Enormous returns flowed to the oil and gas sector and to the financial sector supporting it, driving up wage levels far beyond those in other branches of the economy. Those regions with a heavy concentration of oil and gas extraction in their economic profiles experienced rapid growth in output per capita, which translated into high per capita incomes, wages, and per capita output. Moscow city benefited as well from its administrative status as capital: most of the major national companies kept their headquarters in Moscow in order to remain close to the national political elite and financial institutions. Moscow and the oil and gas producing regions therefore outpaced other regions in the rate of growth of incomes and output.

This history makes it abundantly clear why inter-regional differences in growth and incomes should have risen sharply in the early phases of the market reform era. The question is whether any tendency toward convergence has set in.

3. Trends in interregional wages, incomes, and output: data and analysis

Before presenting the data itself, a word about the problem of adjusting official figures for the effects of price inflation is in order. Because markets in Russia and China are not well integrated, prices can differ dramatically across regions, requiring that we deflate reported statistics on output, incomes and wages by regionally-varying price yardsticks. High-income regions, particularly in Russia, tend to have much higher prices (in part due to Russia’s extreme distances and climatic conditions.) Therefore in the tables and graphs that follow, figures for wages, incomes, and output per capita have been adjusted to take account of differences in relative prices across regions. As a deflator, I have used a year-specific, region-specific measure of the cost in current currency units of a market basket of subsistence goods.

For Russia, this measure is taken from the state statistical agency, Rosstat, which reports the subsistence minimum figure for each region for each year, expressed as the quantity of rubles needed to purchase a fixed basket of subsistence goods. (NB: figures for 1998 and 1999 are not available.) I use this instead of the published deflator because it is available in absolute rubles rather than as a percentage over the previous year and it is measured consistently each year, although there are occasional adjustments in the methodology causing discontinuous movements from one year to the next. I divide figures for gross regional product per capita, income, and wages by the subsistence minimum for the given region in the given year to derive an inflation-adjusted figure.

For China, I use estimates developed by Loren Brandt and Carsten Holz of the per-province price of a weighted basket of goods for each year from 1984 to 2010 (Brandt and Holz, 2006).5 The adjusted figures for wages, incomes, and GDP per capita are simply the reported current figure divided by the subsistence minimum for the given region in the given year. The validity

\[ \text{Fig. 1. China, Russia and US: Household Gini Indexes. Officially Reported Figures.} \]
of these measures depends on the validity of the method used to estimate the cost of a market basket of subsistence goods. This adjustment reduces the distortion associated with high price inflation in the high-income, high-growth regions. The adjustment’s effect is to lower the estimate of cross-regional inequality in each given year. (Below I will also show the results using current prices.)

Table 1 presents measures of interregional dispersion for incomes, wages and output per capita. Table 1 reports the mean, median, standard deviation, coefficient of variation (CV), minimum and maximum values, and number of observations for each of the three series for both countries over the period. Two statistics are of particular note. First is the coefficient of variation (standard deviation divided by the mean), which is a dimensionless statistic that permits a comparison of dispersion in the distribution of some value that is independent of the measurement scale used. The drawback of this measure is that because it is affected by changes in both the mean and the standard deviation, if the mean grows faster than the standard deviation, the CV will decline. Since the mean is affected by the values of extreme outliers, the CV statistic can give a distorted impression of the range of dispersion for the main body of observations. Therefore a second measure is used, the inter-quartile range. This is the gap between the values of the units at the 75th and 25th percentiles. Since half of all observations fall within the inter-quartile range (IQR), the IQR figure describes the spread in the middle of the distribution independent of extreme values. Because the adjusted figures for wages, incomes and output per capita are expressed as multiples of the local subsistence minimum, the IQR is a ratio of the value at the 75th percentile to that of the 25th. A change in the IQR over time indicates how the range of values in the middle of the distribution has risen or fallen.

For China, Fig. 2 shows that the inter-quartile ranges for GDP per capita, incomes and wages have increased enormously over the 1990–2010 period. Even after adjusting for the differential growth of prices, the inter-quartile ranges for all three series rose ten-fold. The coefficients of variation, by contrast, rose and then fell again, indicating that the growth in the standard deviation did not greatly exceed the growth of the mean values. The more informative statistic is therefore the inter-quartile range, which indicates that the divergence in incomes, wages and output across the group of provinces in the mid-range has grown substantially. (Note that the spread in prices has also grown, although less rapidly: the inter-quartile range for the market basket of subsistence goods rose 4.5 times over the same period.)

For Russia, there is no clear trend in inter-regional inequality using the inflation-adjusted measure. The coefficients of variation have risen and fallen, while the inter-quartile ranges have been volatile but have settled into higher territory than in the early 1990s (Fig. 3). In part this effect is due to the enormous increase in price differentials; the inter-quartile range for the value of the subsistence minimum basket of goods rose over ten times. In Russia, therefore, rising prices in the rich regions outpaced the growth in incomes, wages, and output.
Neither the Russian nor Chinese series gives any evidence in favor of convergence. At first glance it might appear that the declining CV’s in both countries accord with the theory, but the trend is an artifact of the statistic. In both countries, mean incomes, wages and output, partly driven by the extreme concentration of incomes in the richest regions, have risen more rapidly than the standard deviations. More telling is the fact that in China the differentiation across regions in the middle of the distribution has grown many times over. For Russia, it may appear that the cross-regional level of differentiation has neither grew nor declined in the 2000s. This result is driven, however, by rapid increase in price differentials in the 2000s. All three series show rapid growth of IQRs in current prices. To the extent that people evaluate their incomes relative to others based on current prices, then inter-regional inequality has risen with no indication of a reverse trend.

Neo-classical theories of convergence distinguish between beta and sigma convergence. Beta convergence refers to the rate at which economic units tend to reach an equilibrium level of output, while sigma convergence is defined as total...
variance (for example, Barro takes the standard deviation of annual cross-sections of logged income and output, by year). However, variance-based measures suffer from some of the same difficulty referred to above, that is, the sensitivity of the measure to extreme values. For that reason it is more informative to use the inter-quartile range as a measure of dispersion.

The convergence theory predicts that the beta values will demonstrate a tendency toward convergence. That is, growth rates of units will be negatively correlated with output levels at a given starting point: over a sufficiently long period of time, richer units will tend to have slower growth than poorer units. The analysis shows that this is the case for Russia. Table 2 presents two simple ordinary least squares (OLS) models. One regresses the compound annual growth rate for GDP per capita on 1995 GDP per capita, and the other adds controls for the share of mineral extraction in the regional economy (to account for resource rents) and for status of the region as a capital (to account for administrative rents). Consistent with the convergence theory, both models find a negative coefficient for initial level of GDP per capita (Table 2).

The negative coefficient on GDP per capita in 1995 suggests that, over a sufficiently long period, beta convergence would produce sigma convergence. However, the convergence theory assumes an underlying trend toward convergence, and holds that external shocks only cause deviations from the trend. If the underlying trend, however, is for interregional differences to remain constant or to grow, then shocks such as state infra-structure investment projects could well induce short term effects boosting growth rates in poorer regions. The negative coefficient on initial GDP per capita, therefore, is not evidence in favor of the theory.

Table 3 presents the equivalent models for China. Here instead of mineral rents as a possible confounding source of growth, I use the geographic location of the province on the coast or not as a dummy control variable, and Beijing and Shanghai (the political and financial capitals) as controls to capture administrative rent streams. The results for both the reduced and full models indicate that the initial starting point level of GDP per capita has a negative relationship with growth, but the effect is well below conventional levels of statistical significance. Only the dummy for Beijing and Shanghai is significant.

For both Russia and China, then, there is weak evidence of beta convergence, although for China the effect is not statistically significant; and there is no evidence for a long-term decline in interregional inequality when examining the range of values for output, incomes, and wages for the majority of regions.

Table 2

| Model 1: | Coefficient (se) | t     | P > |t| |
|---|---|---|---|---|
| GDP per capita (adjusted), 1995 | −.0007 (.00039) | −1.86 | .067 |
| Constant | .029 | | |
| Model 2: Including controls | | | |
| GDP per capita (adjusted), 1995 | −.0011 (.0002) | −4.82 | .000 |
| Moscow and St. Petersburg | .029 (.025) | 1.16 | .25 |
| % of mineral extraction in regional economy, 2012 | .0013 (.0002) | 6.13 | .000 |
| Constant | .031 | | |

r-squared: .12 (Model 1).
r-squared: .51 (Model 2).
N = 79.

Table 3

| Model 1: | Coefficient (se) | t     | P > |t| |
|---|---|---|---|---|
| GDP per capita (adjusted), 1995 | −.002 (.002) | −1.02 | .315 |
| Constant | .108 | | |
| Model 2: Including controls | | | |
| GDP per capita (adjusted), 1990 | −.005 (.003) | −1.50 | .145 |
| Beijing and Shanghai | .012 (.006) | 2.07 | .048 |
| Coastal location | .0018 (.014) | .12 | .90 |
| Constant | .109 | | |

r-squared: .03 (Model 1).
r-squared: .16 (Model 2).
N = 31.
4. Institutional barriers to convergence

Both China and Russia have very high levels of inequality: China’s is the highest in Asia, Russia’s the highest in the post-communist European and post-Soviet region (Li, Sato, and Sicular, 2013, p. 31; Remington, 2011, p. 9). Economists decompose income inequality in various ways: short-term and variable components vs. long-term; inter-regional vs. intra-regional, and so on. Almas Heshmati calculates that about half of overall inequality in China is associated with inter-regional inequality, and half with intra-regional inequality (Heshmati, 2004, p. 5). However, more detailed analysis by Li, Sato and Sicular finds that the rural-urban gap, which cuts across both inter-regional and intra-regional components, is by far the largest source of inequality, accounting for about half of total income inequality in 2007 (Li, Sato, Sicular, 2013, p. 67). They find that the contribution of inter-regional differences is modest, accounting for only about 11–12% of total inequality (Li, Sato, Sicular, 2013, p. 70). Income inequality within urban and rural areas accounts for the rest.

One important reason that the rural-urban gap is so important a source of income inequality in China is the different administrative status of urban and rural residents, a legacy of the socialist system. This is because of the high value of the non-monetary benefits that urban, and especially state sector, workers are entitled to receive. Despite efforts by the central authorities to improve rural living standards, the urban-rural gap has grown in the 1990s and 2000s. The massive migration of rural workers to urban areas for employment helps to raise wages for rural workers in urban areas, but not the difference in non-cash and cash incomes that stems from the residence registration system (hukou). The differential status of workers with urban and rural registrations has been the subject of intense policy debate (Solinger, 1999; Wang, 2008; Wang, 2005; Gallagher et al., 2011). Relevant for our purposes is the fact that rural migrants to the cities are not entitled to receive the full range of urban social benefits. As a result, they represent a low-wage labor pool, enabling those regions with large numbers of rural migrant workers to produce exportable goods at low cost. Export-oriented production has driven the high growth rates of the regions benefitting from low-cost labor, increasing the differences among regions. Generally, private enterprises employ a much higher share of migrant and informal labor than do state enterprises. As private and quasi-private firms enter the market, they increase interpersonal inequality in the high-growth regions as well as interregional income inequality. Thus far from equalizing economic conditions throughout the country, decentralization has enabled some provinces to benefit disproportionately from national policies.

In Russia, the inter-regional gap in incomes accounts for about a third of overall inequality (Yemtsov, 2003). In contrast to China, the engine for the differential growth of regional economies has been the availability of exportable natural resources, above all oil and gas. On average, incomes in regions where at least half of economic output comes from extractive industry are a third higher than in other regions (about 28 thousand rubles per month per person versus 17 thousand). The gap between incomes in those regions and those of other regions has steadily widened as a result of the high wages in those regions. In Russia, in contrast to China, there is a slight negative relationship between the share of private sector employment and average regional incomes (data not shown). As other research has shown, in Russia, income growth both nationally and at the regional level is associated with higher wages and incomes, but also with rising income inequality (Remington, 2011).

It might at first appear, therefore, that different and nationally-specific factors account for the non-convergence of incomes in Russia and China: for China, the hukou system and the preferential policy treatment of coastal provinces, for Russia, the energy-intensive export structure and its associated institutional pathologies (Ross, 1999, 2001; Fish, 2005; but see Luong-Jones and Weinthal, 2010; Treisman, 2010). A closer look, however, reveals some common legacies of the communist economic system that tend to block convergence.

First, in both countries social policy tends to be non-redistributive, even slightly regressive. This may seem paradoxical. In both countries, the revolutionary elimination of private property was thought to abolish class inequality. Therefore, although wage differentiation in the urban industrial sector was relatively low, after the revolution neither country’s social policy was redistributive in intent or effect. Social benefits were universal and categorical, not targeted and means-tested (Haggard and Kaufman, 2008). After liberalization, new tax systems were created to finance much of what had previously been distributed in kind, such as free or discounted public services and administratively allocated access to health care, education, culture, and recreation. Although payroll (social insurance) tax rates are relatively high in both countries, they are capped at relatively low ceilings, imposing a proportionately greater tax burden on low and middle-income individuals in the formal sector.

Moreover, in both Russia and China, social taxes are widely underpaid, often by tacit or explicit consent of local governments, employers and employees. In this connection, the large-scale strike at the Yue Yuen factory in Dongguan, China, in April 2014, was a notable departure. Some 40,000 workers walked out in protest against the company’s long-standing practice of paying social insurance taxes only on base wages, not on overtime and bonuses (Sevastopulo, 2014). This was a highly unusual cause for a strike, because most workers had hitherto preferred to take more of their wages in cash and contribute less to social insurance funds that might not be portable across jobs (Frazier, 2014; Zheng, 2012, pp. 100–1).

---

6 Official National Bureau of Statistics estimates of income and inequality treat rural migrant laborers working in the cities (which accounts for about 20% of all urban labor) as part of the rural population. Because their incomes tend to be higher than of the rural population living and working in rural areas, this measurement procedure in fact yields a lower estimate of urban-rural inequality. The Chinese Household Income Project Series (CHIP) data analyzed by Li, Sato and Sicular treats long-term migrants (but not short-term migrants) to the cities as part of the urban population. Their procedure reduces measured inequality in the urban areas. However, because this group is relatively small compared to the number of short-term rural migrants, it has little impact on the measurement of overall inequality. (Shi et al., 2013, p. 47.) The urban-rural gap would be larger still if all urban labor, including migrants, were treated as part of the urban population.
Likewise in Russia, according to a survey by the state labor inspectorate, most employers pay social taxes only on part of the employee’s wages, often with the consent of the employees (Ponomareva, 2013). In any case, especially in the informal sector, workers have little power to demand that employers pay their payroll taxes (Avshalumova and Gorelova, 2011). A common arrangement in small business is for the employer to pay social insurance contributions on half of an employee’s wage, and to pay the other half in unreported cash.

Both countries still rely heavily on taxes that are administratively less costly to collect but are regressive with respect to income (Gehlbach, 2008; Lin, 2009). In neither country do income tax revenues form a large part of government revenues; both governments rely much more heavily on turnover and consumption taxes (Lin, 2009; Xu and Cui, 2009; Russian State Tax Service). As much as 70% of Russia’s federal budget revenues come from the mineral resource sector. In China, the tax system is even less progressive despite a nominally progressive income tax (Lin, 2009; Xu and Yue, 2013). There is no capital gains tax or property tax. As a result, the richest pay a much lower share of overall income tax than their share in income (Xu and Cui, 2009). In both countries, the relatively low income ceilings on payroll taxes mean that high-end earners pay a much lower share of their total earnings into the social insurance system than lower-income groups, while informal labor pays little or nothing.

The largely non-redistributive system of taxation and social spending, the preservation of categorical and in-kind rather than cash-based, means-tested benefits, the widespread diversion of public resources into private gain by state officials, and rising labor market dualism between formal and informal employment mean that inequality in earnings is translated directly into high post-tax-and-transfer inequality. In some respects, social policy deepens inequality. In China, Gao and Riskin estimate that the net effect of social benefits other than housing is to increase income differentials (Gao and Riskin, 2009). In the urban sector in China, for example, the cumulative effect of educational advantage is slowing intergenerational mobility and deepening inequality (Knight et al., 2013, p. 18). China’s minimum livelihood subsistence guarantee, the so-called dibao, is means-tested but extremely low (Solinger and Hu, 2012). Neither is it redistributive in intent, nor in effect.

The same is true of social policy in Russia, where no more than 5% of total benefits are means-tested (Strategia-2020: 322). Even most of the means-tested subsidies do not reach the poor; half the poor receive no means-tested benefits at all. The cumulative effect is to compound advantages for those well-off enough to benefit from access to better services. As in the United States, the entrenchment of privilege also increases the ability of the rich and powerful to block moves toward more redistributive taxation and spending or toward broader pooling of risks and benefits (Hacker and Pierson, 2010). Therefore the same factors contributing to the increase in income inequality nationally also contribute to divergent growth of incomes across regions.

In China after the tax reform of 1994 and in Russia after Putin took power in 2000, the distribution of tax revenues between central and subcentral governments shifted markedly in favor of the center (Wang and Hu, 2009; Xu, 2011). Fig. 4 illustrates this point. In both cases, fiscal recentralization was aimed in part at giving the central government greater capacity for redistribution of transfer and infra-structure investment funds across regions. In China, for example, the leadership pursued large-scale regional development projects in the western provinces (beginning in 1999), the northeast (2003) and central region (2006) (Li, Sato, and Sicular, 2013, p. 46; Chen, 2010). Putin in Russia used the “national target programs” in part to redress regional imbalances in living conditions. However, in both countries the strategy of cross-regional redistribution tended to slow rather than to reverse the divergence of incomes and output. As Barry Naughton observed of China’s regional development programs, such state efforts may undermine incentives for locally-generated productive investment and therefore fail to reduce disparities across provinces in the longer term (Naughton, 2004). The same is true of Russia’s large-scale efforts to use massive state infra-structure spending, in part financed through the use of pension funds, to stimulate regional development (Fomicheva, 2013).

Perhaps still more important than the fiscal and social policy institutions bequeathed to post-communist Russia and China is the legacy of communist political institutions. The suppression of civil society and dominance of the communist party undermined the capacity of latent collective interests in society (business, labor, farmers and other social groups) to organize. With respect to inequality, the role of the trade unions is particularly important, since research on industrial democracies demonstrates that high rates of centralization and membership in labor unions tend to reduce income inequality (Pontusson, 2005; Rueda and Pontusson, 2000). In communist states, trade union membership was extremely high, but trade unions were instruments of party and government policy, and were not able to defend class-specific interests (Pringle, 2011; Pringle and Clarke, 2011; Tomba, 2002; Ashwin and Clarke, 2003; Hong and Warner, 1998; Clarke, 1999, 2007). The post-communist successors of the communist-era trade unions—FNPR in Russia, ACFTU in China—herited high nominal membership levels but tend to maintain clientelistic relations with local governments. Although both Russia and China have mechanisms for

---

7 When payroll taxes went up from 26 to 34% on January 1, 2011, the director of one small firm in Russia informed his employees that he could only afford to keep on 5 of the 10 workers of the firm. The employees responded by proposing to nominally quit their jobs, apply for unemployment benefits, and continue working at the company on an informal, cash-only, basis. Anecdotal evidence indicates that such arrangements are common (Avshalumova and Gorelova, 2011).

8 Interview with Moscow businessperson, December 2013. The reason for this arrangement, I was told, is that bribes consume a very sizable share of a firm’s revenues.

9 In China and Russia, however, in contrast to the United States, the resistance of powerful organized interests to more redistributive policies comes less at the stage of policy making than in resisting the enforcement of central policy. For example, on the power of local and regional coalitions to block broader pooling of social insurance funds in China, see Frazier, 2010; Zheng, 2007, 2012.
tripartite negotiation over wages and work conditions, these tend to be pro forma and the role of trade union organizations is distinctly inferior in status to that of the government and employers (Ashwin and Clarke, 2003; Pringle and Clarke, 2011; Clarke and Lee, 2002; Liu, 2014; Qiao and Appelbaum, 2011; Warner and Ng, 1999). Needless to say, in neither country may competing political interests propose opposing policy programs to the electorate, so policies over wages, benefits and social policy are not decided through elections. Rather than corporatist or electoral-partisan institutions, policy over the distribution of labor-market income or post-market redistribution is set through bargaining among government actors, with different coalitions of ministries and recognized bureaucratic agencies acting as proxies for latent public interests (Cook, 2007; Heilmann, 2011; Heilmann and Shih, 2013; Lieberthal and Oksenberg, 1988).

It is common to attribute the growth of inequality to the effects of globalization and technological change (Goldin and Katz, 2008). Integration into the global economy has been a fundamental component of the market reforms of both Russia and China and certainly has contributed to the growth of income differentiation across social groups and across regions. However, the character of national institutions determines how globalization shapes the allocation of well-being across social groups and regions (Hall and Soskice, 2001; Garrett, 2000). In neither Russia nor China do labor market institutions foster pooling of the costs and benefits of economic growth across society, and in both, social policy does little to mitigate differentials in labor market earnings. Under these circumstances, with limited government investment in the formation of human capital, national market-regulating institutions still poorly developed, corruption extensive, and political mechanisms facilitating electoral competition over opposing partisan programs blocked, interregional differences persist. The logic is similar to that outlined by Beramendi for political unions under democracy. Where a state’s economic resources are unevenly distributed across regions, and regional political elites command substantial control over fiscal and social policy, they may favor interregional transfers over centralized institutions for extracting and redistributing resources. Under such circumstances, there are few institutional barriers to continued uneven growth across regions. These political factors help explain why there is no “iron law of convergence.”

5. Discussion

It may be that the institutional legacies of the communist system will slowly recede in favor of new institutions that facilitate nationally integrated markets for capital and labor, aggregate social interests, and resolve redistributive conflicts. Labor shortages in some urban areas in China are already producing significant increases in wages, consistent with the theory of the “Lewis turning point” and the Kuznets curve (Chan, 2009). Wage pressure in the coastal regions in turn may increase wages and incomes in inland provinces. Both demographic and institutional forces may therefore bring about a long-term trend toward the convergence of incomes and living standards across regions. Similarly, in Russia the decline in the size of the labor force has tended to raise wages (wages in real terms have risen about 10% per year between 2002 and 2013). To the extent that labor and capital are mobile, labor shortages should drive a general tendency to replace labor with capital and to equalize wage rates across regions.

Before accepting such conclusions, however, we should consider two economic forces countering convergence. One is the unequal accumulation of wealth, the other is the effect of urban agglomerations. As Thomas Piketty has shown, as economic...
and demographic growth in a society slow, there is a tendency for the share of income from capital in total income to increase (Piketty, 2014). Therefore the concentration of the ownership of wealth yields greater concentration of income in a narrower and narrower segment of the population. Already there is evidence that the private ownership of housing in China is growing rapidly as a source of overall inequality (Sato et al., 2013, p. 130). The widespread ownership of housing in the cities tends to reduce intra-urban inequality but exacerbate rural-urban inequality. Therefore, to the extent that regions differ in the level of home ownership, cross-regional inequality is also deepened by the greater share of imputed rental income to homeowners. Likewise, capital income in Russia is likely to rise as a share of total income as growth rates slow. Currently the Russian statistical agency estimates that income from property represents about 5.2% of total income (income from labor comprises about 67%, from entrepreneurship about 8.6%, from social transfers about 18.5%, and “other” about 2%). This figure has been volatile. Before the economic crash in 2008, income from property was as high as 10.3%, and in some regions it is much higher (in Moscow, income from property was almost 25% at its peak in 2005, but fell to 14.6% in 2011). The concentration of property ownership not just in a small segment of the population, but a segment that is concentrated in Moscow and one or two other locations, will certainly deepen cross-regional income inequality over time.

The second factor inhibiting cross-regional convergence has to do with the increasing returns to innovative economic activity. As Enrico Moretti has shown, cities with larger shares of productive innovation, whether in manufacturing or information technology or other tradable sectors, tend to achieve faster growth rates (Moretti, 2013). The growth—increasing effect of innovative clusters spills out into the non-tradable sectors as well, such as personal services. For this reason, there can be long periods in which the growth rates of different regions diverge as centers of innovation attract complementary enterprises. Although innovation in Russia is extremely low, it is entirely possible that some regions in China will grow faster than others because they feature higher concentrations of innovative activity.

The evidence presented here indicates the importance of institutional factors in mediating economic forces, and emphasizes those institutional factors that are the direct legacies of the communist economic model. When the flow of capital and labor faces significant institutional obstacles, regions enjoying a relative advantage following an institutional or technological discontinuity can grow faster than other regions for significant periods of time. As we have seen, barriers to the flow of capital investment and labor exist in both Russia and China. These include political controls over capital investment and household registration in China, and the extreme distances, limited housing market, and poor infrastructural and institutional climate in Russia. As a result, the mechanisms specified in neo-classical convergence theories and competitive federalism models operate only weakly.

The United States offers a telling contrast. Interpersonal inequality in the United States is on a par with that of Russia and China, but its interregional inequality is far lower, and has grown far more slowly. Over a 30 year period in the US during which the Gini index for interpersonal inequality rose from .428 to .477, cross-state inequality in GDP per capita (the Gini index by state) fell from .165 to .144. The inter-quartile range for GDP per capita rose by 44% while the median state GDP per capita rose 63% and the CV fell by 7%. This suggests that in the United States, the integration of markets tends to reduce cross-state differentials even as it deepens within-state inequality. In Russia and China the high barriers to internal mobility of productive factors allows interregional differences to grow over time. In these countries, institutional conditions fostering the growth of interpersonal inequality are translated into interregional inequality in the absence of significant countermeasures on the part of the state.

Acknowledgment

I wish to express appreciation to Leonard Carlson for valuable comments on an earlier draft of this paper. A visiting fellowship at the Davis Center for Russian and Eurasian Studies at Harvard University made possible the research on which this paper was based. I am grateful as well to Andrei Yakovlev, Israel Marques, Brendan McElroy, and other participants in a workshop at the International Center for the Study of Institutions and Development. I am also grateful to Alicia Ju, Kevin Luo and Charlotte Zhang for research assistance.

References


10 Sato, Sicural and Yue include the imputed rental value of housing in their calculations of household income. This is not of course income from capital. However, as real estate and other income-producing assets contribute an increasing percentage of total income, it is inevitable that capital income will contribute to rising inequality, particularly as economic growth declines.