Macroeconomic Dynamics, **25**, 2021, 154–177. Printed in the United States of America.
doi:10.1017/S136510051800069X

REGIONAL CONVERGENCE IN LARGE EMERGING ECONOMIES: A DISTRIBUTION DYNAMICS APPROACH

KIRIL TOCHKOV
Texas Christian University

Rapid growth in large emerging economies has been spread unequally across their regions leading to growing income disparities. This paper examines the distribution of regional output per capita and the evolution of its shape in Brazil, China, India, and Russia from the mid-1990s to the mid-2010s. A comparative analysis of convergence is conducted using a nonparametric methodology. The results reveal different distribution dynamics across the four economies. Chinese regions with below-median output per capita have the highest probability of transitioning toward higher-income levels, while their Brazilian counterparts are trapped at the bottom of the distribution. Although India has the lowest and Russia one of the highest regional income inequalities, they display similar divergence patterns exemplified by high persistence at both ends of the distribution. Our findings indicate that government spending and the rule of law are the major driving forces behind regional convergence, except in Russia where they have the opposite effect. Innovation and property rights also promote convergence in China, but cause regional divergence in India.

Keywords: Regional Income Distribution, Inequality, Convergence, Emerging Economies

1. INTRODUCTION

All large economies are faced with regional income disparities due to a variety of economic, financial, geographical, and institutional factors. But in emerging economies this issue is considerably more urgent because certain regions have been exhibiting disproportionately high growth rates over the past two decades that have led to a rapid polarization of the income distribution. While rich countries have established mechanisms to redistribute income across regions through various grant programs, the fast growth in emerging economies coincided with a decentralization drive that gave regions more control over their resources. This

The author would like to thank Makram El-Shagi, Orn Bodvarsson, Hiro Ito, John Gallup, as well as the participants at the conferences of the Chinese Economists Society in Beijing, Higher School of Economics in Moscow, Henan University INFER Workshop on Applied Macroeconomics in Kaifeng, and the seminar at the Department of Economics at Portland State University for their helpful comments and suggestions. Address correspondence to: Kiril Tochkov, Department of Economics, Texas Christian University, P.O. Box 298510, Fort Worth, TX 76129, USA. Tel.: +1-817-257-7554. Fax: +1-817-257-5058; e-mail: k.tochkov@tcu.edu.

 1469-8056/19

process coupled with an inadequate system of interregional transfers exacerbated existing regional inequalities.

Brazil, China, India, and Russia, which are the focus of this paper, share many similarities in their economic development over the past two decades. Besides being among the largest, most populous, and most rapidly growing economies in the world, the fundamental reforms that started their ascent can be traced back to the early 1990s. Although market-oriented reforms began being implemented in China in the late 1970s, growth picked up dramatically only after the reform drive was revived in 1992–1993 after a period of political turmoil and austerity measures. Following the breakdown of the Soviet Union in 1991, Russia introduced market reforms that initially led to a dramatic decline in output but the booming commodity markets turned Russia's fortunes around in the late 1990s. Faced with a severe economic crisis in 1991, India carried out measures aimed at deregulation and financial as well as trade liberalization, which paid out in the form of rapid economic growth over the next decades. In Brazil, stabilization policies implemented in 1994 brought an end to chronic inflation, budget deficits, and stagnation.

Another common feature of the four emerging economies is the growing income disparity across their regions. As the share of emerging economies in the world economy, and especially in world trade, began to increase over the 1990s, certain regions within these economies were better positioned to benefit from the process of trade and financial liberalization and opening. In China, coastal provinces, such as Guangdong and Zhejiang, turned into manufacturing hubs firmly integrated in the global supply chain. In Russia, regions rich in natural resources and associated with the oil and gas industry, such as Tyumen, profited greatly from growing global demand for energy resources and commodities. Financial hubs, such as Shanghai, Sao Paulo, and Mumbai, have become the wealthiest areas in their respective countries. In general, regions that have attracted foreign direct investment (FDI) and are involved in the export of goods and services in the world market have posted disproportionately higher growth rates than those less integrated in the world trade system.

The goal of this paper is to examine the evolution and determinants of the regional distribution of output per capita in the four major emerging economies from the mid-1990s to the mid-2010s. In particular, we take advantage of a nonparametric methodology to study the shifts and changes in the shape of the distribution as well as the convergence tendencies of regional income. Moreover, we use regression analysis to identify the determinants of distribution dynamics among various economic, demographic, environmental, social, and institutional factors. A number of previous studies have investigated regional inequality and convergence in Brazil [Azzoni (2001), Lima et al. (2010), Mello (2010), and Neto and Azzoni (2011)], China [Hering and Poncet (2010), Lin et al. (2013), and Tian et al. (2016)], India [Barua and Chakraborty (2010), Das et al. (2010), Mallick (2014), and Cherodian and Thirlwall (2015)], and Russia [Fedorov (2002), Kholodilin et al. (2012), Akhmedjonov et al. (2013), and Lehmann and Silvagni

(2013)]. In addition, several works have conducted comparative analysis of these issues across some of the four economies in the sample. For instance, Milanovic (2005) studied the changing patterns of regional inequality in China, India, Brazil, Indonesia, and the United States by focusing on the Gini and Theil indices, while Badunenko and Tochkov (2010) decomposed regional growth in China, Russia, and India and measured the effect of each component on inequality.

In contrast to existing studies, we use a large number of emerging economies and employ a nonparametric methodology that allows us to explore the entire distribution of output per capita and its dynamics over time rather than just the first two moments of the distribution. The shape of the distribution and its evolution are investigated in discreet and continuous space. In particular, we use Markov transition matrices and stochastic kernels to estimate the probability of a region moving from an initial level of output per capita toward or away from the median. A second advantage of the paper over previous works is the comparative nature of the analysis that enables us to study the convergence tendencies across regions in all four emerging economies over the same period of time from the mid-1990s to the mid-2010s using a uniform methodology. Lastly, a unique data set of variables covering various aspects of the economic and institutional framework in which regional economies function helps us identify those factors with the greatest probability of influencing distribution dynamics.

The rest of the paper is organized as follows. The next two sections describe the general methodology and the data, respectively. Section 4 presents the results of the distribution dynamics, while Sections 5 and 6 deal with measuring transition probabilities and mobility. In Section 7 we run robustness tests. Section 8 shows the results of the regression analysis and Section 9 concludes.

2. METHODOLOGY

A variety of methods have been used in the literature to study convergence. Earlier studies tested for the existence of a negative relationship between the average income growth over a period of time and the initial level of income, which became known as β-convergence [Barro and Sala i Martin (1992)]. Other papers emphasized the importance of the decrease in dispersion of per-capita income across countries, termed σ-convergence [Friedman (1992) and Quah (1993a)]. A third group of studies examined whether stochastic shocks that cause income differentials across countries are temporary in nature and would thus have no effect on convergence in the long run [Bernard and Durlauf (1995), Mello (2011), and Holmes et al. (2014)]. The presence of this stochastic convergence is usually investigated by testing for the stationarity of the income differential series using unit root tests. Further, the use of cointegration tests helps detect a common stochastic trend which is interpreted as evidence of convergence.

In contrast to the aforementioned efforts, this paper employs a completely different nonparametric methodology. In a series of seminal papers, Quah (1993a,b, 1996a,b,c, 1997) criticized the standard econometric approach to income convergence, arguing that its focus on the first (β-convergence) and second (σ-convergence) moments of the income distribution describes the dynamics of a representative economy but fails to characterize the evolution of the entire income distribution over time. Instead, Quah used stochastic kernels to study changes in the external shape of the entire distribution as well as intradistributional mobility, which allowed him to detect convergence clubs indicating the existence of multiple steady states. Following Quah (1996b, 1997), we use kernel density estimates to examine the shape of the regional income distribution and employ transition probability functions to investigate distributional dynamics and intradistributional mobility in four large emerging economies. Furthermore, we employ regression analysis to identify the factors driving these dynamic patterns.

3. DATA

Data on regional output per capita were obtained from the CEIC database as well as from the national statistical offices and cover all 31 Chinese provinces and 27 Brazilian states, 32 of the 36 Indian states and union territories, and 79 of the 85 Russian regions.[1] The sample period was determined largely by the availability of regional-level data. In particular, regional output data for Russia are reported only since 1994, while Brazil and India rebase their regional output levels periodically but do not always publish linked series for all prior years making it difficult to compile data over longer periods of time. To enable cross-country comparisons, we use the 1993–1994 Indian GDP series comprising the period 1994–2014 instead of the latest but shorter 2011–2012 series.[2] The 1993 and 2004 series are linked by assuming that the percentage revision in the overlapping year 2004 reflected a drift in the annual estimates, which was then distributed back to the year 1993. For Brazil, we favor the 2002 series covering the period 1995–2014 over the shorter 2010 series.[3]

An implicit region-specific GDP deflator derived from per-capita GDP in current and constant prices was used to convert all data into constant 2004 prices expressed in the national currency units. The median regional per-capita GDP in a given country and year was set equal to 100 and served as numeraire. The median was chosen over the mean because the former is a more meaningful measure in the case of the skewed income distributions exhibited by the countries in the sample. Moreover, the median facilitates cross-country comparisons because it is located in the same spot across transition matrices and density distributions. The highest regional per-capita output in a given year could also serve as numerarie but it conceals the richness of intradistributional dynamics by allowing the mass of regions to move either toward or away from the top-performer benchmark. In contrast, the median splits the sample in the middle and reveals transitions for regions above and below the benchmark. Robustness checks in Section 7 show that ultimately the choice of numeraire does not influence the results.

The descriptive statistics of the sample along with two common inequality measures are reported in Table 1. On average, Russian regions are by far the

TABLE 1. Descriptive statistics of regional GDP per capita

	Brazil	China	India	Russia
Sample size	27	31	32	79
Sample period	1995–2014	1993–2015	1994–2014	1994–2015
GDP per capita				
Mean	8509	5585	3200	13,571
Median	6799	4203	2755	11,330
Range	30,300	13,425	7869	53,190
Coeff. of variation	0.73	0.64	0.55	0.58
Max (% of median)	452	405	317	602
Min (% of median)	17	48	32	28
Inequality measures				
Gini coefficient	0.35	0.30	0.27	0.31
Theil index	0.21	0.16	0.13	0.19

Note: Mean, median, and range are expressed in international dollars and are averages of the corresponding annual levels across all sample years. The per-capita GDP of each region was converted using the national purchasing power parity conversion factor reported in the World Bank's World Development Indicators database.

wealthiest with average output per capita almost three times the level of Chinese provinces. By comparison, Indian states exhibit the lowest average per-capita GDP and regional inequality in the sample. Brazilian states are almost three times wealthier than their Indian counterparts but they are also the regions with the most uneven distribution.

The regression model contains a number of variables that cover various aspects of regional economies. An effort was made to include the same variables in each regression and to measure these in a uniform way across the four countries in the sample, although this was not always possible due to limited data availability at the regional level. Despite the fact that the regression model in Section 8 focuses on distribution dynamics and is different from a standard growth model specification, we choose to include many of the variables common in the existing growth literature. Physical capital enters the equation in the form of gross fixed investment as a percentage of GDP. Only in the case of Brazil, the absence of investment data at the regional level compels us to follow previous studies [Ferreira (2000)] in using industrial electricity consumption (in megawatt hours per capita) as a proxy.

Education is expressed as average years of schooling which were calculated as follows:

$$e_{it} = \sum_j s_j N_{jit}, \tag{1}$$

where N_{jit} is the share of the population in region i and year t with j being the highest level of education attained. The weights s_j represent the length of the respective schooling cycle in years and were taken from Badunenko and Tochkov

(2010). Due to differences in census data across the countries in the sample, the population used to calculate average years of schooling consisted of individuals aged 15–64 in China, 15–80 in India, over 25 years of age in Brazil, and employed individuals in Russia.

The role of the government is assessed via government expenditure expressed as a percentage of GDP. In addition, for China, India, and Russia we include a variable on the importance of state-owned enterprises (SOE).[4] The international aspect was taken into account by variables for trade (the sum of exports and imports) and FDI, both expressed as a percentage of GDP.[5] Infrastructure, telecommunications, and innovation entered the regression model expressed as length of roads and railways (in km per 100 square kilometers), number of mobile phone subscribers (per 1000 people), and research and development (R&D) spending as a percentage of GDP, respectively.[6] We also considered urbanization (urban as a percentage of total population) and health (life expectancy at birth) as potential determinants, as well as the impact of industrial pollution in the case of China (industrial waste water discharge in tonnes per 100,000 yuan of GDP) and Russia (toxic industrial waste in tonnes per 100,000 rubles of GDP).

Institutional factors were assessed via the rule of law and property rights. The rule of law was proxied by the inverse of the number of crimes per 1000 people in Russia, India, and Brazil. For China, we use instead the number of registered lawyers per 100,000 people. Property rights were measured as the number of patent applications per 100,000 people, with the exception of Brazil for which no regional-level data is reported and Russia where the share of privatized enterprises per 100 SOE was used as proxy.[7]

The regression variables described above represent a unique regional-level data set that was collected from a large number of specialized databases, statistical publications, and websites of national and regional statistical authorities, ministries, and agencies. The descriptive statistics are presented in Table 2.

4. DISTRIBUTIONAL ANALYSIS

The first step of the analysis involves estimating a probability density function of regional output per capita within a country using a kernel function.[8] Let $X_1, \ldots, X_n$ be a sample of n independent and identically distributed observations on a random variable X. The density value $f(x)$ at a given point x is estimated by the following kernel density estimator:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right), \tag{2}$$

where h denotes the bandwidth of the interval around x and K is the kernel function.[9] The kernel estimator assigns a weight to each observation in the interval around x with the weight being inversely proportional to the distance between the observation and the x. The density estimate consists of the vertical sum of frequencies at each observation. The resulting smooth curve allows us to visualize

TABLE 2. Descriptive statistics of the regression variables

	Brazil	China	India	Russia
Capital	50.41	38.29	25.25	20.35
	(39.67)	(10.32)	(12.87)	(7.88)
Education	5.60	7.34	5.00	11.86
	(1.10)	(1.15)	(1.32)	(0.38)
Government spending	14.96	14.71	33.34	22.35
	(4.89)	(10.18)	(49.89)	(11.17)
Infrastructure	32.52	1268.92	629.65	114.19
	(24.07)	(1615.77)	(859.4)	(84.98)
Telecom	–	122.17	–	478.69
		(150.87)		(517.09)
FDI	35.76	2.78	1.56	14.47
	(21.02)	(3.04)	(2.40)	(59.72)
Trade	13.64	28.19	–	39.85
	(12.70)	(33.92)		(41.98)
SOE	–	54.54	17.30	6.20
		(18.90)	(30.70)	(2.83)
R&D	–	0.87	0.09	0.82
		(1.05)	(0.15)	(0.94)
Urbanization	71.44	41.39	30.42	69.01
	(10.08)	(17.40)	(10.19)	(12.80)
Pollution	–	231.02	–	4.90
		(137.09)		(6.24)
Health	69.62	71.24	–	64.98
	(2.75)	(3.42)		(2.69)
Rule of law	0.04	8.62	0.54	0.05
	(0.08)	(9.89)	(0.99)	(0.16)
Property rights	–	60.37	0.33	1.92
		(45.83)	(1.00)	(3.47)

Note: Means and standard deviations (in parenthesis) of the regression variables defined in Section 3.

the shape of the distribution of regional output per capita and detect the presence of “convergence clubs” represented by modes.

The kernel density distributions of regional output per capita are shown in Figure 1. The two-dimensional graphs represent snapshots of the smoothed distribution in the early (1995), midpoint (2005), and final years (2013–2015) of the sample period. In 1995, the majority of Chinese provinces are concentrated around the median, while only a few wealthy regions exhibit output per capita that is more than four times the median level. Over the 2000s the probability mass around the median clearly shifts to the right, indicating convergence with the wealthier regions. By 2015 the polarization of the bimodal configuration has lessened as the pronounced mode on the far right of the distribution continues shifting toward the median and the probability mass around the median keeps declining.

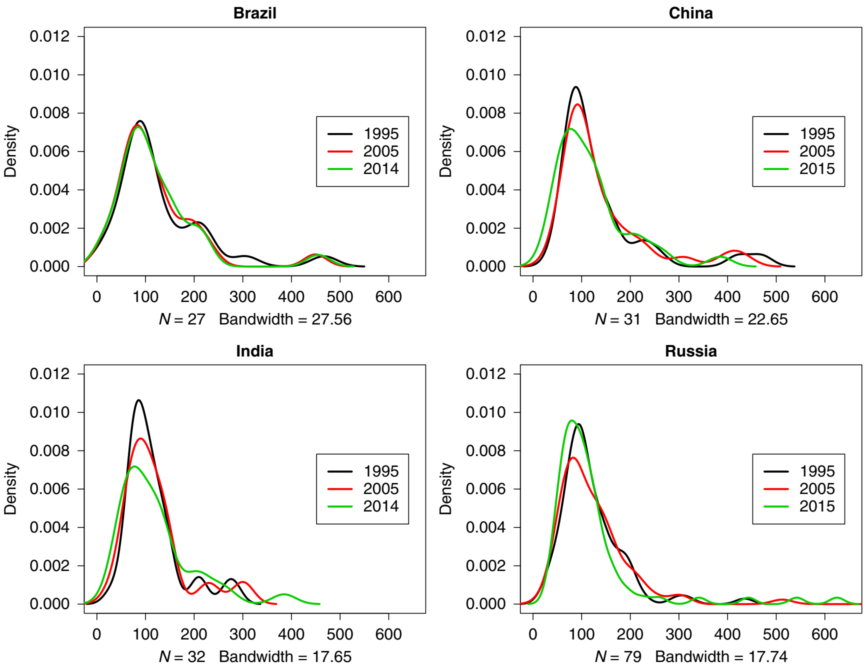

FIGURE 1. Distribution dynamics of regional income.

In India, the initial distribution with two distinct clubs is similar to that of China, except that the richest regions exhibit an output per capita of only around 2.5 times the median. Although the decrease in the probability mass around the median over the sample period is more dramatic than in China, Indian regions experience a convergence toward the rich mode at the same time as wealthy regions manage to increase their output per capita to levels exceeding 300% of the median. By 2014 a group of Indian states managed to achieve per-capita output levels contributing to a growing mode at between two and three times the median level. However, the few richest regions diverge by shifting even further away toward the top of the distribution.

In the early years after the break up of the Soviet Union, only very few wealthy regions in Russia record outputs per capita exceeding three times the median. The 2000s saw a substantial decline in the probability mass around the median, which was channeled mostly into a dramatic extension of the tail of the distribution on the right. This implies that a few regions have managed to break up from the mass and achieve output per capita levels of more than six times the median (and in the case of the oil-producing region of Tyumen almost nine times the median). The financial crisis triggered by falling oil prices and Western economic sanctions imposed in the aftermath of the Ukranian crisis seem to have had a devastating effect on regional inequality. The 2015 distribution reflects a strong polarization between the growing number of regions that found themselves again around or even below the median and a minority of regions that still maintain their lead with per-capita output levels of between 300% and 600% of the median. Regions with twice the median level of per-capita GDP in 2005 appear to be the major casualties of the polarization.

Brazil's distribution of regional output per capita looks different than in the other countries of the sample because it exhibits a relatively large mode at around twice the median level, which in turn reduces the probability mass around the median. Another distinction is the fact that the shape of the distribution barely changes over the sample period. The modes at the top and the bottom of the distribution remain virtually identical, while the regions with two or three times the median output level seem to have converged to the median between the late 1990s and early 2000s. These patterns suggest that the rapid economic growth in Brazil over the 2000s contributed neither toward regional convergence or catching up nor caused any significant changes in regional disparities.

5. TRANSITION PROBABILITIES

The distributions in Figure 1 illustrate the changes in the shape of the distribution of regional output and expose convergence or divergence tendencies between the probability mass centered around the median and the modes at higher levels of output per capita. To gain a better understanding of the dynamics within the distribution, we employ transition probability matrices.[10] Let Q_t denote the distribution of regional output per capita across a given country at time t. The distribution at time $t+1$ is then described by

$$Q_{t+1} = M \times Q_t, \tag{3}$$

where M is a finite discrete first-order Markov transition matrix that contains a complete description of the distributional dynamics as it maps Q_t into Q_{t+1}. The transition matrix is given by

$$M = \begin{pmatrix} p_{ij} & \cdots & p_{iN} \\ \vdots & \ddots & \vdots \\ p_{Nj} & \ldots & p_{NN} \end{pmatrix}, \tag{4}$$

where p_{ij} with $i, j = 1, \ldots, N$ is the probability of a transition from an initial state i at time t to a state j at time $t+1$. The main diagonal of the matrix consists of the probabilities that an observation remains in the same state in t and $t+1$.

Assuming that the transition probabilities from t to $t+1$ are time-invariant and independent of any previous transitions, the evolution of intradistributional mobility can be studied by iterating equation (2) k times. As $k \to \infty$, the iteration yields

$$\lim{}_{k\to\infty} M_j^k = \delta_j > 0, \ \sum \delta_j = 1. \tag{5}$$

The limiting probability distribution, δ_j, is the unconditional or ergodic distribution.[11] In other words, equation (4) describes the convergence to a steady-state distribution independent of the initial distribution. Accordingly, the ergodic distribution allows us to analyze the long-run tendencies of regional output per capita assuming that the observed dynamics continue to hold.

In our analysis, we employ Markov transition matrices, which report the probability of regions moving from one state associated with a certain level of output per capita to another over a period of 5 years.[12] For this purpose, we discretize the state space into six intervals chosen in such a way that each interval contains an approximately equal number of transitions. The corresponding transition matrices are presented in Table 3. The first column denotes the initial state, while the first row indicates the state to which a region has transitioned after 5 years. The advantage of using the median as a reference point is that it is located between states 3 and 4 across all countries.

The wealthiest regions in China with output per capita that is between two and five times the median are very likely to remain in the same state throughout the sample period. In contrast, poor regions with output per capita of less than 80% of the median have a 25% chance to transition to a higher level over 5-year periods. The regions slightly below or above the median exhibit higher mobility which in both cases is directed more strongly toward the median than away from it. The estimated ergodic distribution, shown at the bottom of the matrix, suggests that if these convergence patterns would remain constant in the long run, around half of all regions would end up at or around the mode, while another third would be positioned at higher levels of the distribution.

India presents a profoundly different picture. The poorest regions have a minimal chance of transitioning to higher levels of output per capita, whereas the likelihood of rich regions to remain at the top of the scale is estimated at 93%. Despite the strong persistence at both ends of the distribution, high levels of mobility are present in the middle. In contrast to China, the regions at or below the median have a stronger tendency to converge toward lower-income levels, while for those at between 1.2 and 1.5 times the median level the picture is mixed. As a result, the ergodic distribution features a large mode at the lowest income level where 30% of regions converge in the long run and a smaller but substantial mode at the top of the distribution.

In Russia, the persistence at both ends of the distribution is similarly high as in India but the mobility across states is more intense than in most other countries. This is illustrated by the fact that regions transition across two and in some cases even three states over a period of only 5 years. For instance, a region that initially had an output per capita of around 70–80% of the median was able to double its income relative to the median within 5 years, albeit with a probability of only 5%. In general, the regions in the middle of the distribution do not offer a clear pattern of convergence, although the overall probability of moving toward the bottom of the distribution seems stronger. In the long run, these tendencies produce a bimodal distribution similar to the one in India, with a large mode at the lower

Table 3. Markov transition matrices and ergodic distributions (5-year transitions)

	Brazil, 1995–2014						
State	[15; 62)	[62; 80)	[80; 98)	[98; 132)	[132; 196)	[196; 467]	n
[15; 62)	0.99	0.01	0	0	0	0	67
[62; 80)	0.09	0.78	0.13	0	0	0	68
[80; 98)	0	0.18	0.67	0.15	0	0	67
[98; 132)	0	0	0.12	0.78	0.10	0	68
[132; 196)	0	0	0	0.06	0.90	0.04	67
[196; 467]	0	0	0	0	0.15	0.85	68
Ergodic	0.67	0.07	0.05	0.07	0.11	0.03	405
	China, 1993–2015						
State	[43; 77)	[77; 87)	[87; 100)	[100; 127)	[127; 195)	[196; 488]	n
[43; 77)	0.75	0.25	0	0	0	0	93
[77; 87)	0.14	0.56	0.30	0	0	0	93
[87; 100)	0	0.16	0.69	0.16	0	0	93
[100; 127)	0	0	0.19	0.67	0.14	0	93
[127; 195)	0	0	0	0.15	0.76	0.09	93
[195; 488]	0	0	0	0	0.11	0.89	93
Ergodic	0.07	0.13	0.25	0.21	0.19	0.15	558
	India, 1994–2014						
State	[27; 70)	[70; 84)	[84; 100)	[100; 120)	[120; 148)	[148; 393]	n
[27; 70)	0.91	0.04	0.03	0.01	0	0	90
[70; 84)	0.36	0.46	0.15	0.02	0	0	91
[84; 100)	0.01	0.33	0.44	0.18	0.01	0.03	90
[100; 120)	0	0	0.24	0.48	0.21	0.07	91
[120; 148)	0	0	0.01	0.15	0.80	0.04	90
[148; 393]	0	0	0	0	0.07	0.93	92
Ergodic	0.31	0.07	0.08	0.09	0.19	0.25	544
	Russia, 1994–2015						
State	[19; 69)	[69; 84)	[84; 100)	[100; 123)	[123; 164)	[164; 892]	n
[19; 69)	0.89	0.10	0.01	0	0	0	223
[69; 84)	0.19	0.56	0.19	0.01	0.05	0	224
[84; 100)	0.02	0.32	0.49	0.15	0.02	0	223
[100; 123)	0	0.03	0.20	0.49	0.28	0	224
[123; 164)	0	0	0.02	0.27	0.58	0.12	224
[164; 892]	0	0	0	0	0.18	0.82	225
Ergodic	0.32	0.17	0.12	0.12	0.16	0.11	1343

Note: The numbers in the tables represent the probabilities of transitioning from an initial state (the column on the left) to a final state (the row at the top). Each state is defined by a lower and upper bounds expressed in real per-capita regional income as a percentage of the median.

levels of output per capita and a much smaller but distinct mode near the top of the scale.

In Brazil, the poorest regions record the highest persistence among all four countries with a negligible chance of achieving higher-income levels relative to the median. Regions in all remaining states have a probability of between 67% and 90% of remaining in the same income range, while their likelihood of transitioning to lower levels of output per capita is higher than joining the top performers. The stronger downward tendencies lead to a steady-state distribution that is similar to Russia's ergodic distribution. Close to 70% of the regions end up at the lowest income levels, while only 3% would exhibit output per-capita levels exceeding twice the median value.

6. MOBILITY METRICS

The extent and speed of mobility across states in the transition matrix can be measured formally using three metrics. The average length of stay in the initial state i [Prais (1955)] is given by

$$L_i = \frac{1}{1 - p_{ij}}, \tag{6}$$

where p_{ij} denotes the probability of remaining in the initial state $i = j = 1, \ldots, N$. This metric, which measures the level of persistence expressed in years, can be calculated for all states by using the probabilities along the main diagonal of the transition matrix.

In contrast, the mobility index [Shorrocks (1978)] assesses the extent of mobility across states and can be calculated as follows:

$$S(M) = \frac{n - tr(M)}{n - 1}, \tag{7}$$

where n is the rank of the transition matrix M and $tr(M)$ is its trace. A complete lack of mobility results in a value of zero, while the instant transition away from the initial state is assigned a value of $\frac{n}{n-1}$.

To measure the speed of convergence toward the ergodic distribution, we use an asymptotic half-life measure [Shorrocks (1978)] given by

$$h = \frac{-\log 2}{\log \mid \lambda_2 \mid}, \tag{8}$$

where λ_2 is the second largest eigenvalue of the transition matrix. This metric can be interpreted as the number of years it takes the distribution to move halfway toward its long-run steady state.

The formal measures of persistence, mobility, and convergence speed are presented in Table 4. India, Russia, and Brazil exhibit the highest persistence at the bottom of the distribution with poor regions remaining at lower levels of output per capita for twice (Russia) and even 15 times (Brazil) as long as those at the top of the income scale. In general, duration is lowest, and thus mobility highest,

TABLE 4. Mobility and convergence speed metrics (5-year transitions)

	Brazil	China	India	Russia
Duration in				
State 1	100.0	4	11.1	9.1
State 2	4.6	2.3	1.9	2.3
State 3	3.0	3.2	1.8	1.9
State 4	4.6	3.0	1.9	1.9
State 5	10.0	4.2	5.0	2.4
State 6	6.7	9.1	14.3	5.6
Mobility	0.21	0.34	0.40	0.43
Half life	52.2	11.6	15.2	9.6

Note: Duration in a given state is measured in years. The mobility index can take values between 0 (no mobility) and 1.2 (instant transition away from the initial state). Half life represents the years it takes the distribution to move halfway toward its long-run steady state.

in the middle of the distribution. In comparison, wealthy regions in China endure the longest in their respective states, while those at the lower-income levels have the highest mobility.

Despite the lack of movement at both ends of their distributions, India and Russia have the highest overall mobility, mainly due to the significant shifts in the middle of the distributions. Brazil features the lowest mobility, which is largely driven by complete stagnation at the lowest levels of per-capita output. Brazil's regions, on the other hand, would need more than 50 years to move halfway toward their ergodic distribution because of low overall mobility and the fact that more than 70% of regions would have to wind up at the lowest income levels in the long run.

7. ROBUSTNESS TESTS

The transition probability matrix approach has two major drawbacks that might distort the distributional dynamics.[13] First, it uses continuous data on regional output per capita to estimate a discrete model. Second, the discretization of the state space into states i and j, with $i, j = 1, \ldots, N$ is somewhat arbitrary. To avoid these potential issues and test for the robustness of the results, we focus on transition probabilities in a continuous state space and, following Quah (1997), estimate a stochastic kernel that maps the distribution Q_t into $Q_{t+\tau}$ as follows:

$$Q_{t+\tau}(x_{t+\tau}) = \int g(x_{t+\tau} \mid x_t) Q_t(x_t) dx, \quad \textbf{(9)}$$

where the conditional density function $g(x_{t+\tau}|x_t)$ describes the probability of the transition to a certain state in $t + \tau$ given the initial state in t. In line with Hyndman et al. (1996), the conditional density is estimated using a kernel estimator given by

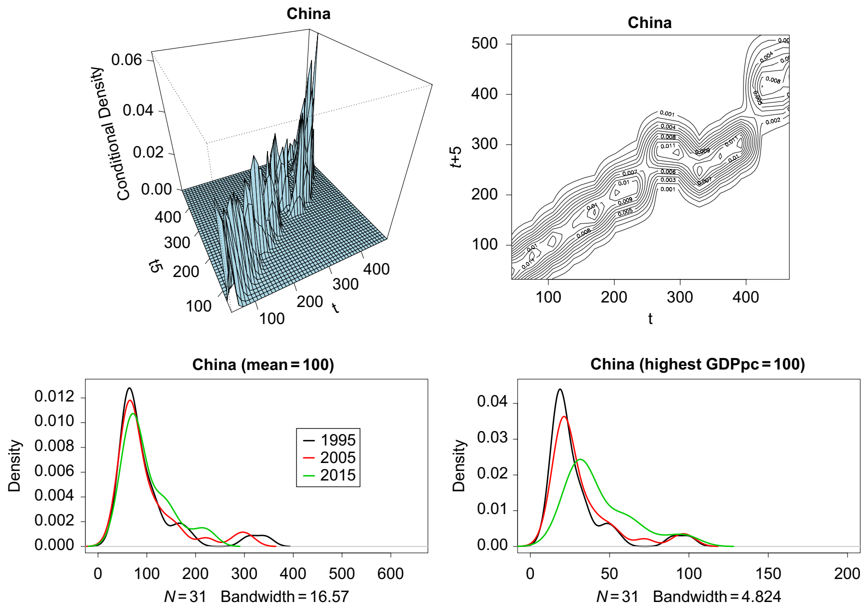

FIGURE 2. Robustness tests Panel A: Stochastic kernels/Panel B: Alternative benchmarks.

$$\hat{g}(x_{t+\tau} \mid x_t) = \frac{\hat{z}(x_{t+\tau}, x_t)}{\hat{f}(x_t)}, \tag{10}$$

where $f(x_t)$ is the marginal density from equation (1) and $\hat{z}(x_{t+\tau}, x_t)$ is the joint density given by

$$\hat{z}(x_{t+\tau}, x_t) = \frac{1}{nhb} \sum_{i=1}^{n} K\left(\frac{x_{t+\tau} - X_{i,t+\tau}}{b}\right)\left(\frac{x_t - X_{it}}{h}\right), \tag{11}$$

with h and b denoting the bandwidth of the interval around x_t and $x_{t+\tau}$, respectively. The visual representation of the stochastic kernel produces three-dimensional graphs with corresponding two-dimensional contour plots.

The resulting stochastic kernels for China are shown in Panel A of Figure 2.[14] The contour plots provide a two-dimensional view of the distributions, where contours represent points of equal frequency. The vertical dimension measures the conditional probability of a region with output per capita of x percent of the median in $t+5$ years, given an income level of y percent of the median in year t shown on the horizontal axis. As with the Markov transition matrix, peaks along the main diagonal indicate high persistence and lack of intradistributional mobility, while peaks off the diagonal denote high probability of convergence or divergence.

The graphs largely confirm the results from Section 5, but they also provide more detailed insights into the movements at higher-income levels that were obscured by the discretization of the state space. Although the transition matrix

TABLE 5. Robustness test: Markov transition matrices (alternative benchmarks)

State	[31; 56)	[56; 64)	[64; 73)	[73; 91)	[91; 144)	[144; 344]	n
	China (mean = 100)						
[31; 56)	0.72	0.27	0.01	0	0	0	93
[56; 64)	0.13	0.51	0.37	0	0	0	93
[64; 73)	0	0.08	0.63	0.29	0	0	93
[73; 91)	0	0	0.22	0.59	0.19	0	93
[91; 144)	0	0	0	0.14	0.78	0.08	93
[144; 344]	0	0	0	0	0.06	0.94	93
Ergodic	0.02	0.03	0.15	0.19	0.26	0.35	558
	China (highest GDPpc = 100)						
State	[10; 17)	[17; 20)	[20; 24)	[24; 30)	[30; 47)	[47; 100]	n
[10; 17)	0.58	0.34	0.08	0	0	0	93
[17; 20)	0.05	0.33	0.40	0.22	0	0	93
[20; 24)	0	0.03	0.34	0.46	0.16	0	93
[24; 30)	0	0	0.05	0.38	0.56	0.01	93
[30; 47)	0	0	0	0.03	0.61	0.35	93
[47; 100]	0	0	0	0	0.04	0.96	93
Ergodic	0.00	0.00	0.00	0.01	0.10	0.89	558

for China in Table 3 shows very low levels of mobility in the top income range, the graphs in Figure 2 reveal convergence toward the median at output levels between three and four times the median and at the top income levels of the distribution, where the peaks are located underneath the main diagonal. In contrast, at the output level of three times the median there is a distinct peak above the diagonal implying a movement away from the median and toward higher-income levels.

The choice of the median as the benchmark for comparative analysis might also affect the results. Accordingly, the mean and the highest per-capita output for a given year were used as alternatives to test the robustness of the findings. Panel B in Figure 2 displays the corresponding kernel densities. Clearly, the patterns and dynamics are very similar regardless of the benchmark. In all three graphs (median, mean, and top level) the probability mass around the median decreases consistently over time, implying convergence toward higher levels of per-capita output. Furthermore, the opposite convergence from the top toward the median or mean is reflected in the shift of the rich mode to the left. When the wealthiest region is chosen as a benchmark that has to remain static, this intradistributional convergence tendency is revealed by a larger shift of the major mode to the right.

Table 5 presents the transition matrices for the two alternative benchmarks. The differences in the transition probabilities between the mean and the median benchmark are relatively small. The ergodic distributions are both unimodal, except that the mode in the one relative to the mean is at the highest range of per-capita output rather than in the middle. This is caused by the fact that the small differences in transition probabilities have a compounding effect in the long run. These patterns intensify under the top-performer benchmark with increases in mobility across states that compensate for the inability of the rich mode to converge toward the median, as mentioned above.

8. DETERMINANTS OF DISTRIBUTION DYNAMICS

The distribution dynamics analysis focuses on the probability of moving across the state space of the transition matrix but does not explain the determinants of the transition process. Therefore, we relate the location in a given state of the Markov transition matrix to potential explanatory variables by employing a multinomial logit regression model. In particular, the dependent variable in the regression is categorical, with six possible outcomes defined as follows depending on where the output per worker of a given region i is located within the state space in year t and year $t+\tau$:

$$y_{i,t+\tau} = \begin{cases} 1 & \textit{if below the mean in } t \textit{ and remains in the same state in } t+\tau \\ 2 & \textit{if below the mean in } t \textit{ but moves to a lower state in } t+\tau \\ 3 & \textit{if below the mean in } t \textit{ but moves to a higher state in } t+\tau \\ 4 & \textit{if above the mean in } t \textit{ and remains in the same state in } t+\tau \\ 5 & \textit{if above the mean in } t \textit{ but moves to a lower state in } t+\tau \\ 6 & \textit{if above the mean in } t \textit{ but moves to a higher state in } t+\tau \end{cases}. \quad \textbf{(12)}$$

The vector of independent variables contains a number of region-specific regressors described in Section 3. Their coefficients are estimated via maximum likelihood but are difficult to interpret in the current context as they represent logs of odds relative to a base category. Instead, we derive the marginal effect of a given regressor x on the probability of each category of the dependent variable as follows:

$$\frac{\partial P_m}{\partial x} = P_m \left(\beta_m - \sum_{q=1}^{6} P_q \beta_q \right) = P_m(\beta_m - \bar{\beta}), \quad \textbf{(13)}$$

where P_m is the probability of $y_{i,t+\tau} = m$, with $m = 1, \ldots, 6$ representing the categories of the dependent variable, while β_m is the estimated coefficient of regressor x for outcome m from the multinomial logit regression.[15] The marginal effects are computed at the sample averages of the regressors.

The results from the multinomial regressions for Brazil, China, India, and Russia are shown in Tables 6–9, respectively. The analysis for each country focuses mostly on those regions that move either toward or away from the mean since they reveal the patterns of distribution dynamics. Accordingly, a variable contributes to regional convergence if the coefficients for below-average regions that move to lower levels and above-average regions that move to higher levels

Table 6. Marginal effects in the multinomial regression for Brazil

	Below/ lower	Below/ same	Below/ higher	Above/ lower	Above/ same	Above/ higher
Capital	0.269***	0.404**	−0.030	0.161*	−0.673***	−0.131
	(0.092)	(0.189)	(0.079)	(0.096)	(0.206)	(0.108)
Education	−15.53***	−66.69***	−6.85*	24.33***	63.17***	1.64
	(5.70)	(11.48)	(3.71)	(6.67)	(11.70)	(3.23)
Government spending	3.09**	11.33***	0.91	−1.27	−12.46***	−1.60*
	(1.20)	(2.27)	(0.68)	(1.23)	(2.43)	(3.23)
Infrastructure	0.364**	−0.400	−0.029	−0.433*	0.384	0.025
	(0.142)	(0.396)	(0.041)	(0.255)	(0.397)	(0.145)
FDI	−0.314*	0.035	−0.92	0.06	0.502**	−0.072
	(0.173)	(0.192)	(0.122)	(0.140)	(0.196)	(0.146)
Trade	0.705	−0.945	−0.257	−0.193	0.48	0.21
	(0.461)	(0.933)	(0.378)	(0.418)	(0.764)	(0.313)
Urbanization	−0.950**	−1.83**	−0.411	0.927**	1.87**	0.396
	(0.383)	(0.763)	(0.263)	(0.372)	(0.758)	(0.286)
Health	2.13	−25.29***	2.62*	5.67**	13.35*	1.52
	(2.28)	(7.24)	(1.57)	(2.65)	(7.84)	(1.91)
Rule of law	−16.79	4.38	12.86	39.64**	−32.97	−67.13
	(93.95)	(22.46)	(71.86)	(18.21)	(23.15)	(80.82)
Log likel.		−126.09		Observations		216
Chi-square		368.34***		McFadden Pseudo R^2		0.594

Note: Marginal effects are derivatives of the probabilities from the logit regression with respect to the regressors. Standard errors are in parenthesis. *** $p < .01$. ** $p < .05$. * $p < .10$.

of output per capita carry a negative sign. Similarly, an indication of convergence tendencies is a positive sign of the coefficients for below-average regions that move upward and for above-average regions that move downward.

Physical capital accumulation seems to be largely contributing to either persistence or divergence in regional income across the sample. The effect is most prominent in China where significance of the coefficients is achieved only for movements toward both ends of the distribution. Below-average regions in Brazil and Russia and above-average regions in India and Russia also exhibit strong persistence or polarization tendencies. Only in above-average regions in Brazil and Russia capital indicates some probability of convergence toward the mean but in both cases the magnitude of the coefficients is smaller than the divergence forces. These findings indicate that physical capital accumulation is more effective in promoting growth in wealthier regions.

Education has a more heterogeneous impact across the sample. It does not play a significant role in regional dynamics in Russia, which might be explained by the fact that Russia has the highest average levels of education in the sample and the

TABLE 7. Marginal effects in the multinomial regression for China

	Below/ lower	Below/ same	Below/ higher	Above/ lower	Above/ same	Above/ higher
Capital	0.242**	−0.375	−0.463	0.027	0.223	0.345*
	(0.121)	(0.487)	(0.319)	(0.277)	(0.333)	(0.204)
Education	−3.98**	−33.63***	−9.94***	8.54***	33.48***	5.53**
	(1.89)	(6.48)	(3.86)	(3.04)	(5.75)	(2.57)
Government spending	0.064	3.41***	1.14**	0.616	−3.89***	−1.33***
	(0.165)	(0.906)	(0.533)	(0.729)	(0.622)	(0.344)
Infrastructure	−0.002	−0.017***	−0.004	−0.002	0.02***	0.006**
	(0.001)	(0.006)	(0.004)	(0.002)	(0.005)	(0.002)
Telecom	0.005	0.026	0.026	−0.003	−0.033	−0.020
	(0.004)	(0.044)	(0.030)	(0.01)	(0.037)	(0.021)
FDI	−0.031	−5.07**	−2.72*	1.73*	5.27***	0.820
	(0.614)	(2.32)	(1.62)	(1.02)	(1.67)	(0.702)
Trade	−0.034	−0.730	−0.382	−0.466**	1.79***	−0.174
	(0.043)	(0.462)	(0.272)	(0.228)	(0.526)	(0.233)
SOE	0.077	0.606**	−0.148	0.137**	−0.661**	−0.012
	(0.055)	(0.288)	(0.196)	(0.070)	(0.293)	(0.109)
R&D	−3.43*	7.91	8.16**	−0.085	−6.86***	−3.43*
	(1.75)	(5.29)	(4.05)	(0.076)	(2.64)	(2.08)
Urbanization	0.116	−0.157	−0.718***	−0.05	0.637***	0.172*
	(0.088)	(0.269)	(0.176)	(0.142)	(0.177)	(0.097)
Pollution	−0.001	0.108***	0.049**	−0.035*	−0.083***	−0.037**
	(0.010)	(0.034)	(0.023)	(0.019)	(0.025)	(0.017)
Health	−0.631*	−3.66**	2.92*	−1.01*	1.88	0.495
	(0.331)	(1.82)	(1.52)	(0.598)	(1.30)	(0.774)
Rule of law	0.179	−3.13***	2.08**	−0.341	1.58*	−0.366**
	(0.404)	(1.03)	(1.02)	(0.232)	(0.822)	(0.167)
Property rights	0.021	−0.38***	0.035	0.106***	0.173***	0.043
	(0.038)	(0.126)	(0.103)	(0.040)	(0.054)	(0.035)
Log likel.		−188.15		Observations		341
Chi-square		587.95***		McFadden PseudoR^2		0.611

Note: Marginal effects are derivatives of the probabilities from the logit regression with respect to the regressors. Standard errors are in parenthesis. *** $p < .01$. ** $p < .05$. * $p < .10$.

regional variation comparatively small. In China and Brazil, schooling leads to strong convergence toward the mean from above but it also prevents the poorest regions from slipping further toward the bottom. The latter is also true for India, while there are some signs that wealthier regions diverge from the mean thanks to education.

Table 8. Marginal effects in the multinomial regression for India

	Below/ lower	Below/ same	Below/ higher	Above/ lower	Above/ same	Above/ higher
Capital	−0.88***	−1.85***	−0.67**	0.033	3.18***	0.199*
	(0.252)	(0.315)	(0.299)	(0.214)	(0.502)	(0.117)
Education	−10.44***	−13.52***	−2.62	1.79	23.65***	1.14*
	(2.42)	(2.94)	(2.85)	(1.95)	(4.18)	(0.687)
Government spending	0.06***	−0.041	0.088***	0.024	−0.125**	−0.006
	(0.019)	(0.042)	(0.026)	(0.045)	(0.052)	(0.005)
Infrastructure	−0.001	0.002	0.000	−0.004*	0.003	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.000)
FDI	0.015	0.079	0.047	0.019	−0.175	0.016
	(0.053)	(0.347)	(0.189)	(0.115)	(0.259)	(0.022)
SOE	0.000	−0.004	0.008	0.045**	−0.051**	0.001
	(0.010)	(0.012)	(0.007)	(0.021)	(0.021)	(0.005)
R&D	54.33***	−20.92	−28.79	−5.59	1.24	−0.276
	(17.32)	(39.49)	(40.50)	(29.89)	(43.02)	(3.55)
Urbanization	−0.591*	−0.879***	−0.824**	0.078	2.15***	0.068
	(0.319)	(0.255)	(0.327)	(0.270)	(0.482)	(0.060)
Rule of law	3.36	29.05***	27.92***	8.14	−67.52***	−0.946
	(9.05)	(8.50)	(8.77)	(8.85)	(16.08)	(1.36)
Property rights	−4.31	−1.27	−50.77***	1.47	55.75***	−0.887
	(5.07)	(1.78)	(12.99)	(13.34)	(16.62)	(3.24)
Log likel.		−188.89		Observations		210
Chi-square		276.89***		McFadden Pseudo R^2		0.423

Note: Marginal effects are derivatives of the probabilities from the logit regression with respect to the regressors. Standard errors are in parenthesis. *** $p < .01$. ** $p < .05$. * $p < .10$.

Government spending is found to have made modest contributions to convergence in most cases. In China, it promotes transition toward the mean from below and prevents the top performers from diverging. In India, the probability of convergence toward the mean from below is stronger than toward the bottom. In Brazil, government spending curbs divergence at the top but is unable to stop divergence at the bottom. The least positive effect is recorded for Russia where persistence below average and a decrease in the probability of moving toward the mean from above result in divergence.

Infrastructure, telecommunications, and trade are shown to promote regional divergence. FDI has mostly the opposite effect, helping regions converge toward the mean, except in India where it is not a significant factor. The presence of SOEs prevents rich regions from moving toward the top, generating convergence in the process. For below-average regions, SOEs do no matter, except in Russia where they are an impediment. R&D spending contributes to convergence in China and

TABLE 9. Marginal effects in the multinomial regression for Russia

	Below/ lower	Below/ same	Below/ higher	Above/ lower	Above/ same	Above/ higher
Capital	−0.251	1.22***	−0.325	0.636**	0.706*	0.453**
	(0.202)	(0.435)	(0.291)	(0.313)	(0.412)	(0.215)
Education	2.60	9.23	−3.47	7.37	−11.69	−4.04
	(2.84)	(7.65)	(5.22)	(5.88)	(7.44)	(3.89)
Government spending	−0.191	2.18***	0.376	−1.2***	−0.771*	−0.401
	(0.173)	(0.447)	(0.353)	(0.346)	(0.414)	(0.267)
Infrastructure	0.02	0.212***	0.11***	−0.11***	−0.22***	−0.013
	(0.016)	(0.036)	(0.026)	(0.033)	(0.038)	(0.017)
Telecom	−0.004	−0.02***	−0.02***	0.01**	0.03***	0.006*
	(0.003)	(0.006)	(0.005)	(0.005)	(0.006)	(0.003)
FDI	−0.091	−0.72***	0.243**	−0.34**	0.717***	0.193***
	(0.080)	(0.194)	(0.101)	(0.154)	(0.157)	(0.061)
Trade	−0.32***	−0.33***	0.001	0.038	0.49***	0.11**
	(0.067)	(0.079)	(0.057)	(0.080)	(0.079)	(0.045)
SOE	−0.652	−7.25***	−3.37***	1.85*	8.69***	0.739
	(0.515)	(1.52)	(1.11)	(1.08)	(1.44)	(0.917)
R&D	−2.17*	−2.7***	−3.12	5.62***	1.52	0.847
	(1.29)	(0.361)	(2.04)	(1.84)	(2.68)	(1.27)
Urbanization	−4.55***	−2.70***	−0.95***	0.307	3.14***	0.65***
	(0.175)	(0.361)	(0.237)	(0.263)	(0.341)	(0.180)
Pollution	−1.05***	−2.99***	−1.65***	0.884*	3.50***	1.3***
	(0.403)	(0.677)	(0.551)	(0.464)	(0.594)	(0.312)
Health	−4.49	−6.04***	−0.562	0.143	5.37***	1.54*
	(6.22)	(1.48)	(1.09)	(1.17)	(1.53)	(0.799)
Rule of law	−34.66	50.62***	−35.2***	16.65	−19.28	−92.59
	(66.05)	(16.27)	(13.37)	(14.37)	(17.83)	(87.83)
Property rights	−4.88	−2.99**	0.303	−0.158	2.50**	0.83**
	(0.554)	(1.36)	(0.763)	(0.948)	(1.06)	(0.395)
Log likel.		−934.16		Observations		790
Chi-square		755.35***		McFadden PseudoR^2		0.288

Note: Marginal effects are derivatives of the probabilities from the logit regression with respect to the regressors. Standard errors are in parenthesis. *** $p < .01$. ** $p < .05$. * $p < .10$.

Russia, whereas in India it drags poor regions down. Urbanization exacerbates regional divergence with the exception of Brazil where it encourages transitions toward the mean from both below and above. Health and the rule of law promote convergence except for Russia where these factors seem to favor wealthier regions and obstruct poor regions. In Russia and India, property rights lead to regional polarization, while in China they appear to trigger a convergence effect from above.

As can be seen from the above discussion, it is not a simple task to derive a clear set of policy recommendations from the regression analysis because various factors have diverging or opposite effects across the sample. We believe that this is evidence for the presence of heterogeneity across the four large economies. Although they share many features such as rapid economic growth and regional inequalities, the factors behind these dynamics reveal major institutional, structural, cultural, and social differences that are not easily summarized in the context of a comparative analysis.

We test the robustness of the regression results by employing a different definition of the dependent variable. While the six states in equation (12) remain the same, the classification of shifts up or down the distribution after the 5-year period now depends on the magnitude of the shift. Only transitions that amount to more than 5% of the median are considered. This procedure tests for the possibility that the dynamics in the original regressions neglect large shifts that might occur within a given state while focusing on minor shifts across states for regions that are closer to a state's bounds. The results (available upon request) remain largely robust in the sense that most factors in each country show very similar convergence or divergence patterns as in Tables 6–9.

9. CONCLUSIONS

The rapid growth achieved by large emerging economies has been spread unequally across their regions resulting in an increasing income gap between poor and wealthy areas. This paper examines the distribution of regional output per capita and the evolution of its shape in four major emerging economies between the mid-1990s and mid-2010s. Although Brazil, China, India, and Russia exhibit rapid growth and face similar problems regarding regional income disparities, the results reveal different patterns of shifts within the regional income distribution across the four countries. China and Brazil represent the two extremes. China's regions with below-median output per capita have one the highest probability of transitioning toward higher-income levels, while their Brazilian counterparts are almost completely trapped at the bottom of the distribution. Moreover, although in both countries there are strong convergence tendencies, these lead in China to a concentration of probability mass at the middle-income levels in the long run, while the majority of regions in Brazil end up in a steady state at the bottom of the distribution. Although India has the lowest and Russia one of the highest regional income inequality levels in the sample, they display similar divergence patterns exemplified by high persistence at both ends of the distribution. The high mobility in the middle is directed toward the top- and the bottom-income levels resulting in the long run in a bimodal distribution with a larger poor mode.

Our findings further indicate that government spending and the rule of law are the major driving forces behind regional convergence, except in Russia where they have the opposite effect. Innovation and property rights also promote convergence in China, but cause regional divergence in India. A larger share of SOE increases the likelihood that rich provinces will move down toward the median, while trade strengthens tendencies in the opposite direction. From a policy perspective, the most relevant variables are those that help regions with initial below-median output per capita improve over time without necessarily constraining the growth of rich regions. Our analysis identified government spending and the rule of law (in China and India), FDI (in Russia and Brazil), innovation (in China and Russia), and health (in China and Brazil) as potential factors in this category. These results reveal heterogeneity across the countries in the sample, which impedes the formulation of a clear set of universal policies that can be implemented to diminish regional inequality in large emerging economies.

NOTES

1. Three Indian union territories (Dadra and Nagar Haveli, Daman and Diu, and Lakshadweep) and the Chechen Republic in Russia were dropped due to lack of data. Three autonomous regions of Russia (Nenents, Yamalo-Nenents, and Khanty-Mansi) were treated as subdivisions of the regions they are part of. The Indian state of Telangana was counted as part of Andhra Pradesh, from which it was carved out in 2014.
2. Annual data in India are reported for fiscal rather than calendar years. The fiscal year begins on April 1 and ends on March 31. For simplicity, single years are used to denote fiscal years. For instance, fiscal year 2004–2005 will be referred to as 2004.
3. Regional data for the years before 2002 are retropolated and harmonized with the national data series by the Brazilian Institute of Geography and Statistics (IBGE).
4. In China, this variable takes the form of gross industrial output of SOE as a percentage of total industrial output, whereas in Russia and India it is measured as the share of SOE in the total number of registered enterprises.
5. Trade was excluded from the regression for India because of insufficient data over the sample period.
6. Due to data limitations, we excluded telecommunications from the model for Brazil and India, and innovation from the regression for Brazil.
7. The crime rate and patent applications have been used in previous studies of large emerging economies as proxies for institutional quality [e.g., da Mata et al. (2007), Fan (2011), Baranov et al. (2015), and Smith and Thomas (2017)].
8. The following description of kernel densities has been used by the author in a different context in previous publications [e.g., Nenovsky and Tochkov (2014), Tochkov (2015), and Wamboye and Tochkov (2015)].
9. We use data-driven bandwidth selection (likelihood cross-validation) and a Gaussian kernel.
10. The following description of transition matrices has been used by the author in a different context in previous publications [e.g., Nenovsky and Tochkov (2014), Tochkov (2015), and Wamboye and Tochkov (2015)].
11. The ergodic distribution is unique if there is only one eigenvalue of M with modulus one.
12. We opt for the 5-year periods because annual transitions are too short to result in any significant movements across the six states despite the rapid growth of many regions in the sample. In addition, 5-year transitions control for the effect of short-run fluctuations that are likely to occur.
13. The following description of the disadvantages and robustness tests has been used by the author in a different context in previous publications [e.g., Nenovsky and Tochkov (2014)].

14. To save space, we present the results of the robustness tests for China. The conclusions about the other countries in the sample are similar and the corresponding figures are available upon request.

15. It is worth mentioning that the dependent variable in the regression is not directly linked to the probabilities of the transition matrix. The six categories define states of persistence or transition without reference to the corresponding probabilities. At the same time, the marginal effects estimated from the logistic regression are interpreted as the change in the probability of being in a given persistence/transition state in response to a marginal change in a given independent variable. Accordingly, these estimates are conceptually related to the transition probabilities but the latter are not included directly into the regression model.

REFERENCES

Akhmedjonov, A., M. Lau and B. Balci (2013) New evidence of regional income divergence in post-reform Russia. *Applied Economics* 45(16–18), 2675–2682.

Azzoni, C. (2001) Economic growth and regional income inequality in Brazil. *The Annals of Regional Science* 35, 133–152.

Badunenko, O. and K. Tochkov (2010) Soaring dragons, roaring tigers, growling bears: Determinants of regional growth and convergence in China, India and Russia. *Economics of Transition* 18, 539–570.

Baranov, A., E. Malkov, L. Polishchuk, M. Rochlitz and G. Syunyaev (2015) How (not) to measure Russian regional institutions. *Russian Journal of Economics* 1(2), 154–181.

Barro, R. and X. Sala i Martin (1992) Convergence. *Journal of Political Economy* 100, 223–251.

Barua, A. and P. Chakraborty (2010) Does openness affect regional inequality? A case study for India. *Review of Development Economics* 14(3), 447–465.

Bernard, A. and S. Durlauf (1995) Convergence in international output. *Journal of Applied Econometrics* 10, 97–108.

Cherodian, R. and A. Thirlwall (2015) Regional disparities in per-capita income in India: Convergence or divergence? *Journal of Post-Keynesian Economics* 37(3), 384–407.

da Mata, D., U. Deichmann, J. Henderson, S. Lall and H. Wang (2007) Determinants of city growth in Brazil. *Journal of Urban Economics* 62(2), 252–272.

Das, S., G. Sinha and T. Mitra (2010) Regional convergence of growth, inequality, and poverty in India: An empirical study. *Economic Modelling* 27(5), 1054–1060.

Fan, P. (2011) Innovation capacity and economic development: China and India. *Economic Change and Restructuring* 44(1–2), 49–73.

Fedorov, L. (2002) Regional inequality and regional polarization in Russia, 1990-1999. *World Development* 30(3), 443–456.

Ferreira, A. (2000) Convergence in Brazil: Recent trends and long-run prospects. *Applied Economics* 32(4), 479–489.

Friedman, M. (1992) Do old fallacies ever die? *Journal of Economic Literature* 30, 2129–2132.

Hering, L. and S. Poncet (2010) Income per capita inequality in China: The role of economic geography and spatial interactions. *The World Economy* 33, 655–679.

Holmes, M., J. Otero and T. Panagiotidis (2014) A note on the extent of U.S. regional income convergence. *Macroeconomic Dynamics* 18(7), 1635–1655.

Hyndman, R., D. Bashtannyk and G. Grunwald (1996) Estimating and visualizing conditional densities. *Journal of Computational and Graphical Statistics* 5, 315–336.

Kholodilin, K., A. Oshchepkov and B. Siliverstovs (2012) The Russian regional convergence process: Where is it leading? *Eastern European Economics* 50(3), 5–26.

Lehmann, H. and M. Silvagni (2013) Is There Convergence of Russia's Regions? Exploring the Empirical Evidence: 1995-2010. OECD Economics Department, Working Paper: No. 1083.

Lima, L., H. Notini and F. Gomes (2010) Empirical evidence on convergence across Brazilian states. *Revista Brasileira de Economia* 64(2), 135–160.

Lin, P., C. Lin and I. Ho (2013) Regional convergence or divergence in China? Evidence from unit roots tests with breaks. *Annals of Regional Science* 50(1), 223–243.

Mallick, J. (2014) Regional convergence of economic growth during post-reform period in India. *Singapore Economic Review* 59(2), 1–18.

Mello, M. (2010) Stochastic convergence across Brazilian states. *Brazilian Review of Economerics* 30(1), 23–52.

Mello, M. (2011) Stochastic convergence across U.S. States. *Macroeconomic Dynamics* 15(2), 160–183.

Milanovic, B. (2005) Half a world: Regional inequality in five great federations. *Journal of the Asia Pacific Economy* 10(4), 408–445.

Nenovsky, N. and K. Tochkov (2014) Transition, integration, and catching up: Income convergence between central and eastern Europe and the EU. *Mondes en Développement* 42(3/167), 73–92.

Neto, R. and C. Azzoni (2011) Non-spatial government policies and regional income inequality in Brazil. *Regional Studies* 45(4), 453–461.

Prais, S. J. (1955) Measuring social mobility. *Journal of the Royal Statistical Society. Series A, Part I* 118, 56–66.

Quah, D. (1993a) Galton's fallacy and the convergence hypothesis. *Scandinavian Journal of Economics* 95, 427–443.

Quah, D. (1993b) Empirical cross-section dynamics in economic growth. *European Economic Review* 37, 426–434.

Quah, D. (1996a) Convergence empirics across economies with (some) capital mobility. *Journal of Economic Growth* 1, 95–124.

Quah, D. (1996b) Empirics for economic growth and convergence. *European Economic Review* 40, 1353–1375.

Quah, D. (1996c) Twin peaks: Growth and convergence in models of distribution dynamics. *Economic Journal* 106, 1045–1055.

Quah, D. (1997) Empirics for growth and distribution: Stratification, polarization and convergence clubs. *Journal of Economic Growth* 2, 27–59.

Shorrocks, A. F. (1978) The measurement of mobility. *Econometrica* 46, 1013–1024.

Smith, N. and E. Thomas (2017) Regional conditions and innovation in Russia: The impact of foreign direct investment and absorptive capacity. *Regional Studies* 51(9), 1412–1428.

Tian, X., X. Zhang, Y. Zhou and X. Yu (2016) Regional income inequality in China revisitied: A perspective from club convergence. *Economic Modelling* 56, 50–58.

Tochkov, K. (2015) Regional integration and economic convergence in Northeast Asia: Evidence from the Greater Tumen Initiative. *Journal of Northeast Asia Development* 17, 85–108.

Wamboye, E. and K. Tochkov (2015) External debt, labor productivity growth, and convergence: Evidence from Sub-Saharan Africa. *The World Economy* 38(5), 856–877.