5. DISTRIBUTION DYNAMICS APPROACH

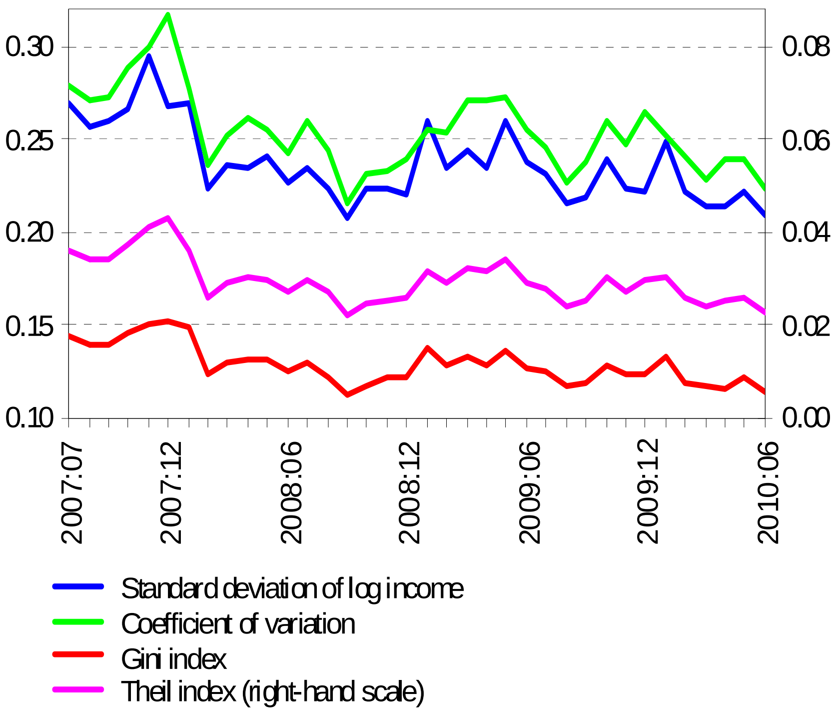
TRACING CHANGES IN INEQUALITY
(“σ-convergence”)

$I(y_t)$ = an index of inequality (e.g., the s.d. of logs, Gini index, Theil index, etc.) in an economic/social indicator y at time t.

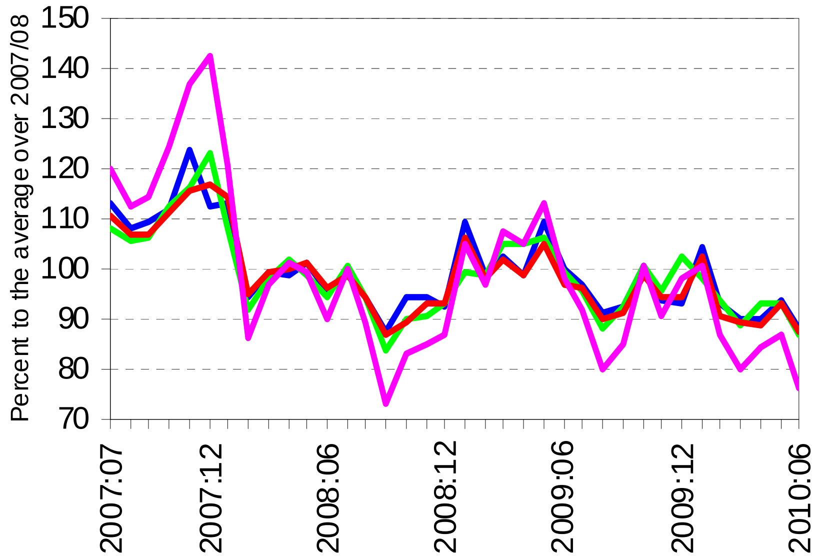
The question is: $I(y_{t+\tau}) < I(y_t)$?

Regional inequalities in real personal income per capita in Russia

Original values



Normalized to the time averages



EXAMINING THE SHAPE OF DISTRIBUTION AND ITS CHANGES OVER TIME

- Estimation of kernel density:

$$\hat{f}_t(y_t) = \frac{1}{Nh}\sum_i K\left(\frac{y_t - y_{it}}{h}\right),$$

h = smoothing bandwidth, $K(x)$ = kernel

(Epanechnikov kernel: $K(x) = 0.75(1 - x^2)$ if $x \in [0, 1]$, otherwise $K(x) = 0$;

Gaussian kernel: $\exp(-x^2/2)/\sqrt{2\pi}$; etc.)

- Testing identity of distributions:

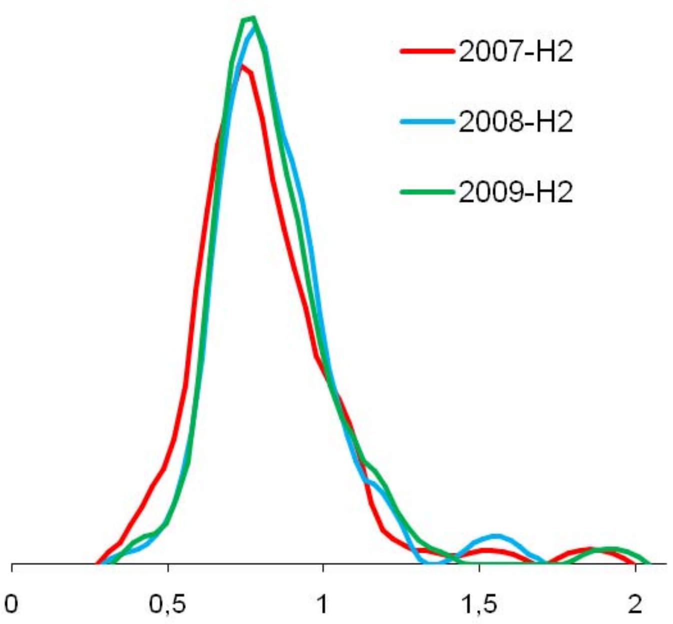
the two-sample Kolmogorov-Smirnov test; H_0: $f(y_t)$ and $f(y_{t+\tau})$ are identical

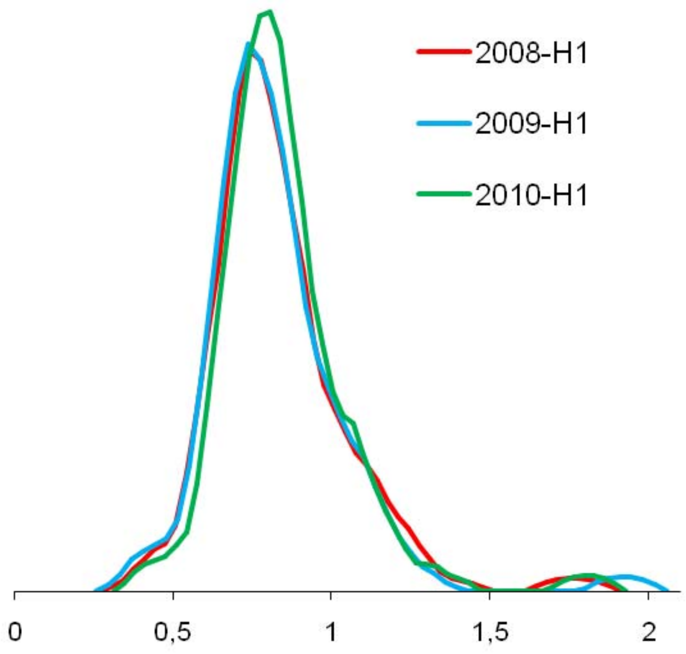
- Testing distribution for multimodality: see

Bianchi M. Testing for convergence: Evidence from non-parametric mulimodality tests. *Journal of Applied Econometrics*, 1997, V. 12, No. 4;

Herzfeld T. Inter-regional output distribution: a comparison of Russian and Chinese experience. *Post-Communist Economies*, 2008, V. 20, No. 4.

Kernel density estimates of real income per capita across Russian regions (incomes are averaged over a half-year)

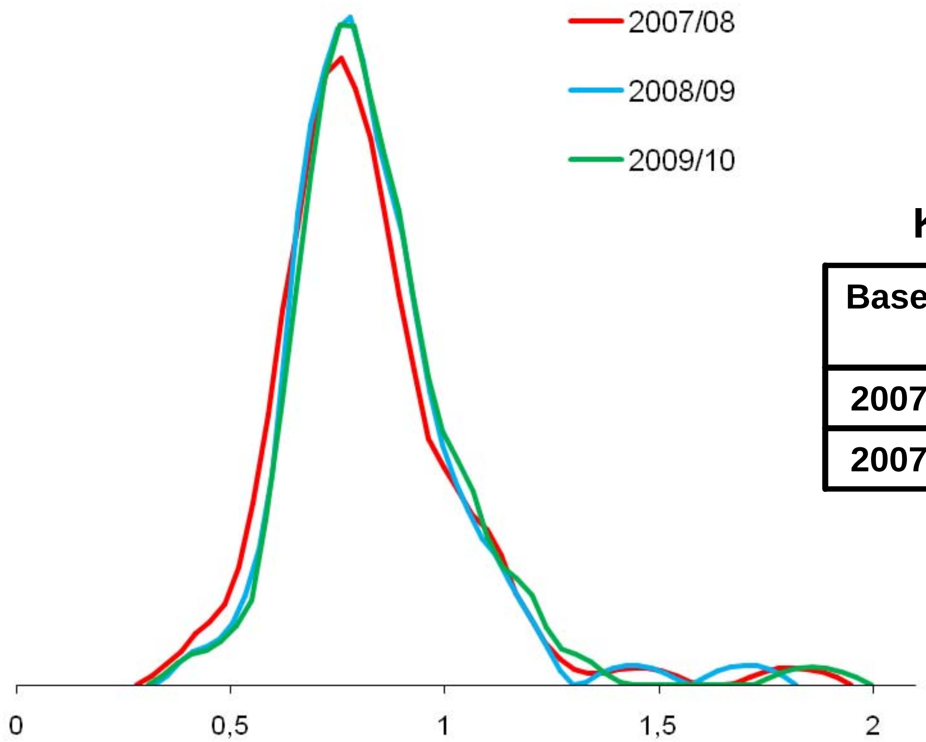




Kolmogorov-Smirnov test *p*-values

Base distribution in:	Compared with distribution in:	*p*-value	Compared with distribution in:	*p*-value
2007, 2nd half	2008, 2nd half	0.078	2009, 2nd half	0.322
2008, 1st half	2009, 1st half	0.977	2010, 1st half	0.684

Kernel density estimates of real income per capita across Russian regions (incomes are averaged over a 12-month periods)



Kolmogorov-Smirnov test *p*-values

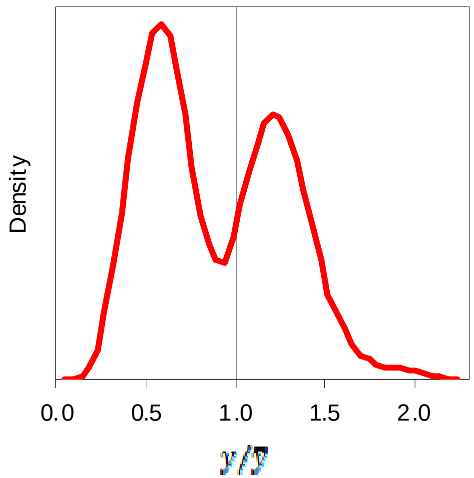
Base distribution in:	Compared with distribution in:	*p*-value
2007:07–2008:06	2008:07–2009:06	0.684
2007:07–2008:06	2009:07–2010:06	0.684

POLARIZATION

Polarization means that a society is divided into a few groups, with substantial intra-group homogeneity and inter-group heterogeneity.

Such a state is considered as a source of tensions and social conflicts.

(Sometimes, only the case of two groups is called polarization, while cases of more groups are referred to as stratification.)



Convergence clubs
(polarization)

Testing a distribution for multimodality:

H_0: $f(y)$ has m modes
against
H_1: $f(y)$ has more than m modes

However, this detects polarization, but does not provide its quantitative characterization.

POLARIZATION INDICES

- The Esteban-Ray index: $P^{(\mathrm{ER})} = \dfrac{1}{\bar{y}} \sum_{i=1}^{N} \sum_{j=1}^{N} n_i^{1+\alpha} n_j \left| y_i - y_j \right|$,

where n_i, n_j are population shares of locations i and j; $0 < \alpha < 1.6$ is the degree of "polarization sensitivity."

The number of groups is not involved.

- The Wolfson index: $P^{(\mathrm{W})} = (0.5 - L(0.5) - 0.5G(y)) \cdot \bar{y} / m(y)$,

where $L(0.5)$ is the value of the ordinate of the Lorentz curve at the median income, $m(y)$; $G(.)$ is the Gini index.

Two groups are implied. (The index measures the distance between the given distribution and the "perfectly bimodal" one.)

- The Zhang-Kanbur index:

$$P^{(\mathrm{ZK})} = Th_{\mathrm{between}}(y)/Th_{\mathrm{within}}(y_{(1)}, \ldots, y_{(m)}) \ ,$$

where (arbitrary) *m* is the number of groups. Locations are grouped by researcher.

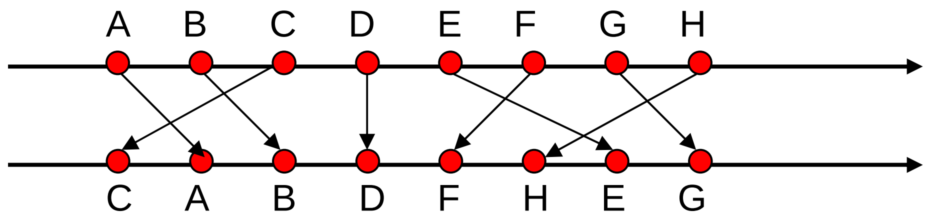
For comparison of the polarization indices, see:

Estban J., Ray D. A comparison of polarization measures. *UFAE and IAE Working Paper* No. 700.07, 2007.

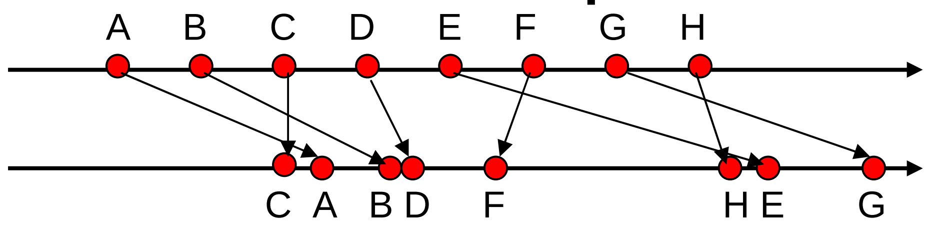
INTRADISTRIBUTION MOBILITY

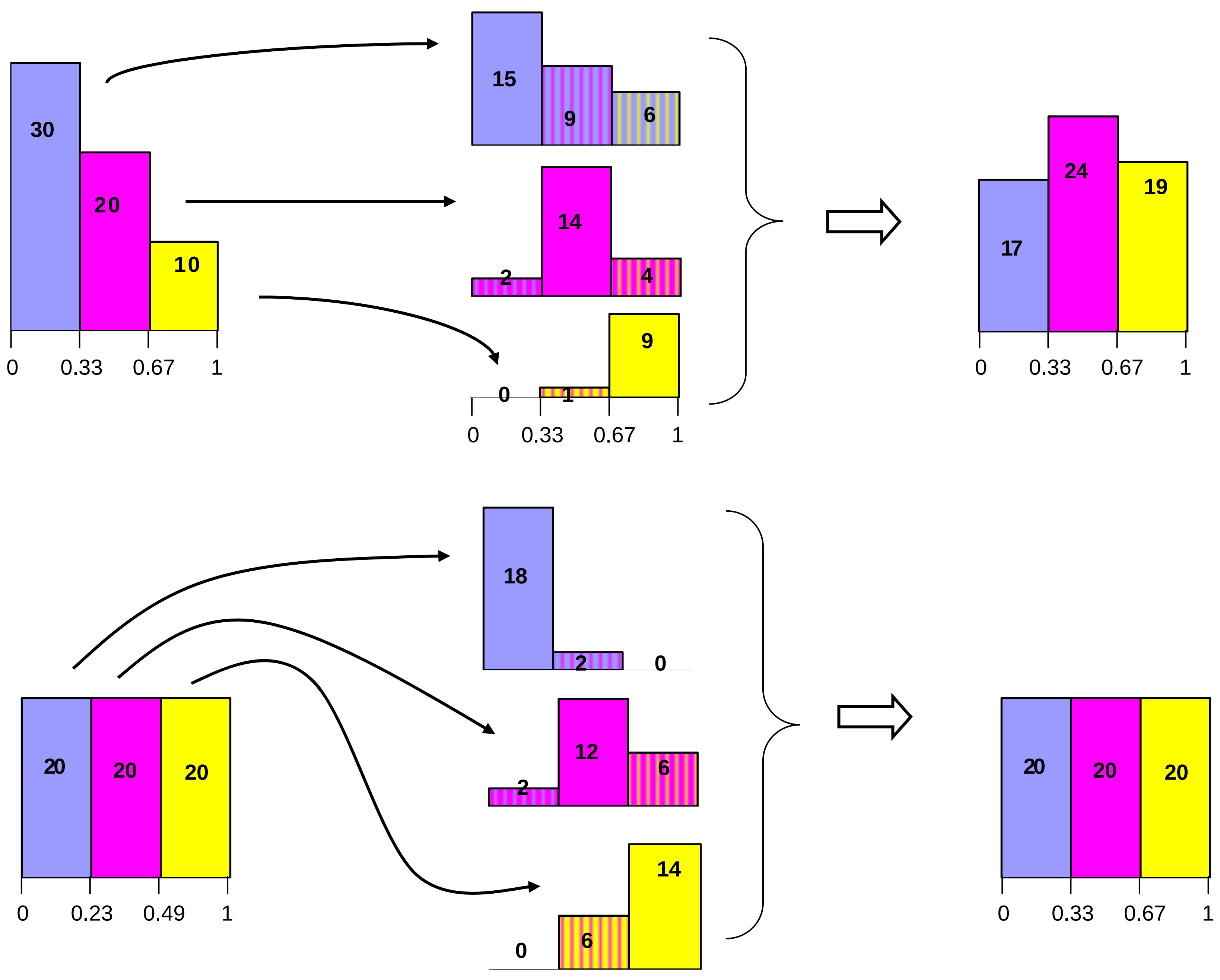
- Generally, **social mobility** is the movement of individuals or groups (classes, ethnic groups, regions, entire nations) through a system of social hierarchy or stratification.
- It may refer to education, health status, literacy, etc.
- The case when **changes in income** are dealt with may be considered as a special type of social mobility, namely, **economic mobility**.
- In terms of distributions, mobility is the movement of different parts of the distribution to other positions.

- **Rank mobility**: change in the positions of locations in the income space relative to one another, or changes in the order of locations.



- **Quantity mobility**: the movement of locations in the income space irrespective of their relative positions.





30
20
10
0
0.33
0.67
1
15
9
6
14
2
4
9
0
1
0
0.33
0.67
1
17
24
19
0
0.33
0.67
1
18
2
0
12
2
6
14
0
6
20
20
20
0
0.23
0.49
1
20
20
20
0
0.33
0.67
1

TRANSITION PROBABILITY MATRIX

The income space is divided into K income classes.

$\mathbf{P} = (p_{ij})$, where i, j are income classes.

p_{ij} is the fraction of locations that transferred from class i to class j over time span $[t, t + \tau]$.

Or, it is an estimate of probability of being in income class j at time $t + \tau$ on condition that in has been in class i at time t.

A diagonal element p_{ii} shows the fraction of immobile locations, i.e. those remained in the same income class i.

The difference between rank and quantity mobility lies in the way of constructing income classes.

Analyzing rank mobility, these classes represent income quantiles. Thus, the number of locations in each class is always constant and equal to K/M. However, the boundaries of classes at times t and $t + \tau$ in income terms will be different in the general case.

Considering quantity mobility, the entire possible range of incomes, e.g. [min(min(y_{rt}), min($y_{r,t+\tau}$)), max(max(y_{rt}), max($y_{r,t+\tau}$))] or $[0, \infty)$, is divided into K equal or unequal intervals representing the income classes. In this case, the boundaries of income classes are time-invariant, but the numbers of locations in classes vary over time. In particular, some income classes may turn out to be empty in the beginning or end of the time span.

EXAMPLE

	[0, 0.33]	(0.33, 0.67]	(0.67, 1]
[0, 0.33]	**0.5**	**0.3**	**0.2**
(0.33, 0.67]	**0.1**	**0.7**	**0.2**
(0.67, 1]	**0**	**0.1**	**0.9**

0.5	⇨	**0.283**
0.333		**0.4**
0.167		**0.317**

	1st tertile	2nd tertile	3rd tertile
1st tertile	**0.9**	**0.1**	**0**
2nd tertile	**0.1**	**0.6**	**0.3**
3rd tertile	**0**	**0.3**	**0.7**

0.333	⇨	**0.333**
0.333		**0.333**
0.333		**0.333**

GINI INDEX OF (RANK) MOBILITY

Yitzhaki, S., and Q. Wodon (2004). Mobility, inequality, and horizontal equity. In Amiel, Y., and J. A. Bishop (eds.), *Research on Economic Inequality*, Vol. 12. Oxford: Elsevier, pp. 179-199.

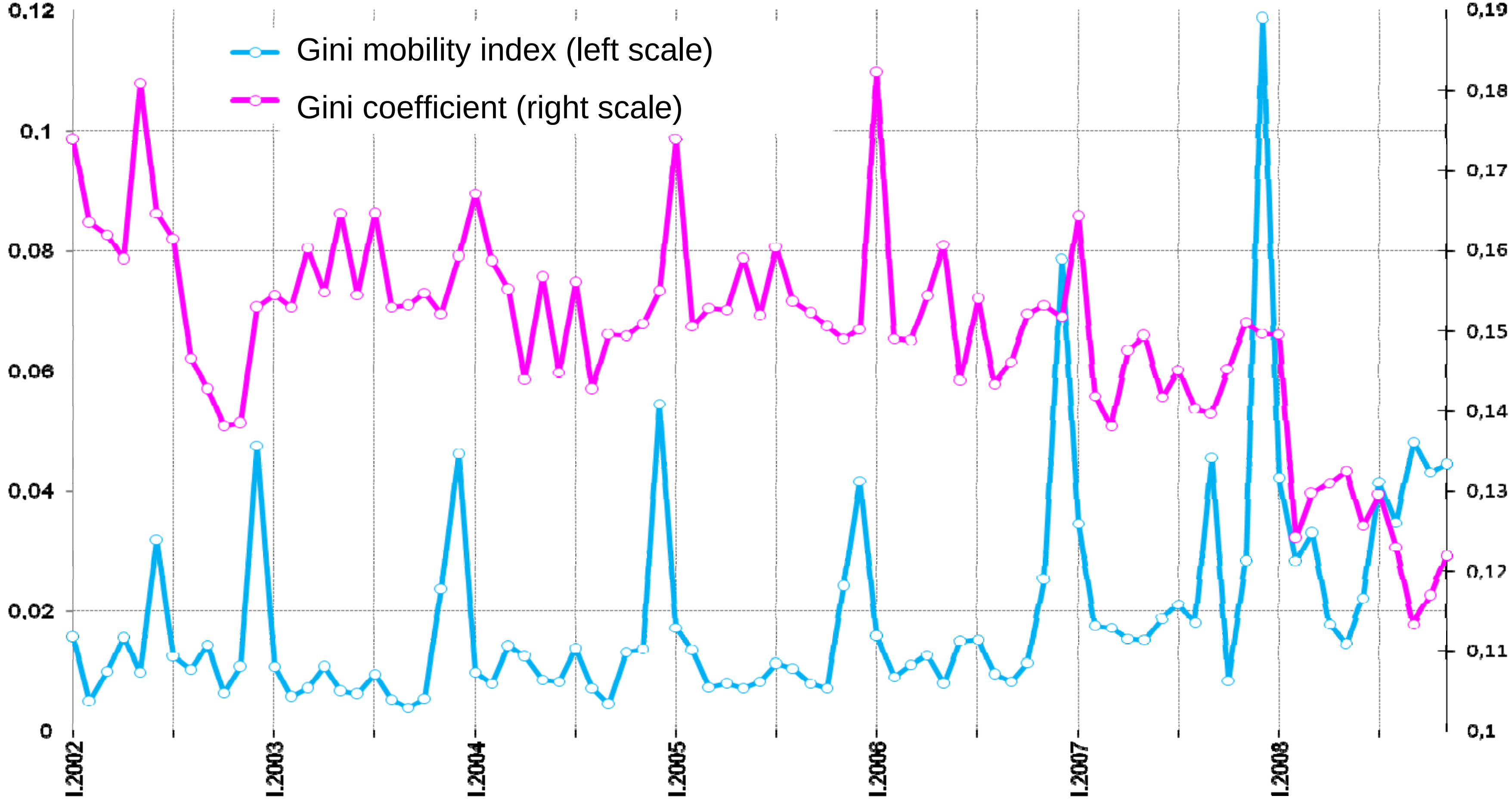
- Gini correlation coefficients:

$\Gamma_{t,t+\tau} = \mathrm{cov}(y_t, R(y_{t+\tau}))/\mathrm{cov}(y_t, R(y_t))$,
$\Gamma_{t+\tau,t} = \mathrm{cov}(y_{t+\tau}, R(y_t))/\mathrm{cov}(y_{t+\tau}, R(y_{t+\tau}))$

- Gini mobility indices: $M_{t,t+\tau} = (1 - \Gamma_{t,t+\tau})/2$, $M_{t+\tau,t} = (1 - \Gamma_{t+\tau,t})/2$

- Gini symmetric index of mobility: $S_t = \dfrac{G_t M_{t+\tau,t} + G_{t+\tau} M_{t,t+\tau}}{G_t + G_{t+\tau}}$

RANK REAL INCOME MOBILITY OF RUSSIAN REGIONS



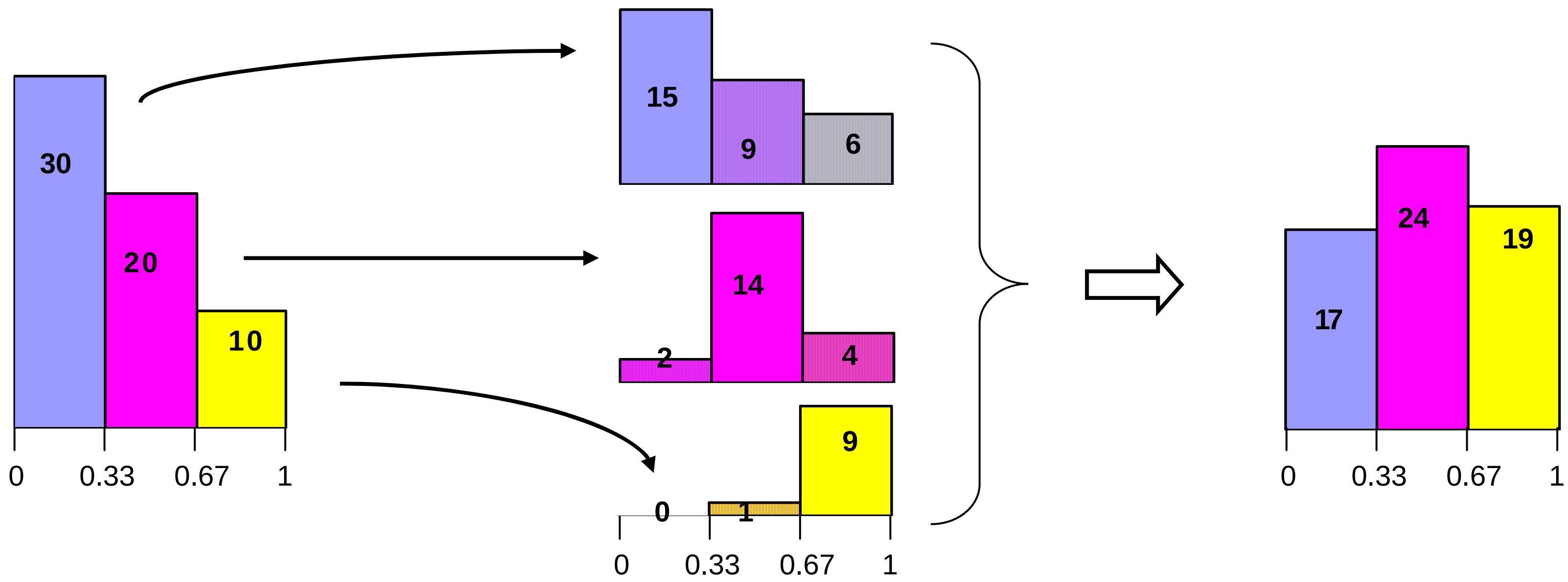
ANALYZING QUANTITY MOBILITY

- The law of motion: $f_{t+\tau}(y) = \Lambda \bullet f_t(y)$, where Λ is an operator mapping distribution at t into that at $t+\tau$
- The operator is assumed to be time-invariant; thus we have a discrete-time Markov process; and $f_{t+n\tau}(y) = \Lambda^n \bullet f_t(y)$
- Taking $n \to \infty$ yields the ergodic distribution, $f_\infty(y) = \Lambda_\infty \bullet f_t(y)$, such that $f_\infty(y) = \Lambda_\infty \bullet f_\infty(y)$, where Λ_∞ is the limit of Λ^n with $n \to \infty$.

The ergodic distribution is the long-run limit of the distribution. Depending on unimodality or multimodality of the ergodic distribution, it can be judged whether the existence of convergence clubs is to be expected in the long run.

- What are operator Λ and operation $\bullet$?

DISCRETE VERSION. RECALL:



0.5
0.333
0.167

	[0, 0.33]	(0.33, 0.67]	(0.67, 1]
[0, 0.33]	**0.5**	**0.3**	**0.2**
(0.33, 0.67]	**0.1**	**0.7**	**0.2**
(0.67, 1]	**0**	**0.1**	**0.9**

0.283
0.4
0.317

DISCRETIZATION

The distribution $f_t(y)$ is represented by $1{\times}K$ or $K{\times}1$ stochastic vector $\mathbf{f}_t = (f_{it})$; its i-th element is the probability of income to fall into class i (at time t).

Then Λ is a $K{\times}K$ Markov chain transition matrix $\mathbf{P}$ (if $\mathbf{f}_t$ is a raw vector) or transposed matrix $\mathbf{P}'$ (if $\mathbf{f}_t$ is a column vector); $\bullet$ is the matrix multiplication:

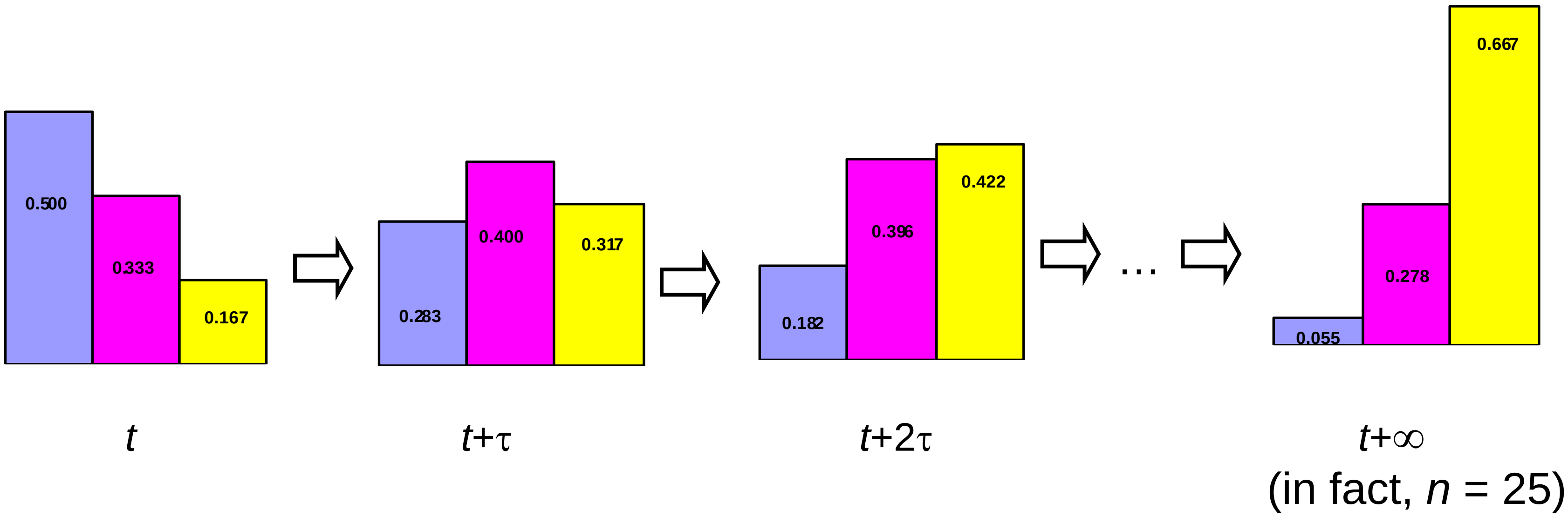
$f_{t+\tau}(y) = \Lambda \bullet f_t(y) \;\sim\; \mathbf{f}_{t+\tau} = \mathbf{f}_t\mathbf{P}$ or $\mathbf{f}_{t+\tau} = \mathbf{P}'\mathbf{f}_t$.

$$\begin{pmatrix} 0.5 & 0.1 & 0 \\ 0.3 & 0.7 & 0.1 \\ 0.2 & 0.2 & 0.9 \end{pmatrix} \begin{pmatrix} 0.5 \\ 0.333 \\ 0.167 \end{pmatrix} = \begin{pmatrix} 0.283 \\ 0.4 \\ 0.317 \end{pmatrix}$$

In Λ_∞, all columns (rows) are equal to one another and to ergodic vector.

This property can be used to approximate Λ_∞ with desirable precision.

ERGODIC DISTRIBUTION



$$\Lambda_{\infty} \approx \Lambda^{25} = \begin{pmatrix} 0.055 & 0.055 & 0.055 \\ 0.278 & 0.278 & 0.278 \\ 0.667 & 0.667 & 0.667 \end{pmatrix}$$

STOCHASTIC KERNEL

- Infinitesimal income classes: $y^{(i)} + dy$
- This gives continuous (hence infinite) number of rows and columns in Λ
- Then Λ is a **stochastic kernel**, or a transition probability function which is a generalization of transition probability matrix
- Λ is a probability density of incomes at $t + \tau$ conditional on incomes at t: $\Lambda = f(y_{t+\tau} | y_t)$
- $\bullet$ is integration: $f_{t+\tau}(y) = \Lambda \bullet f_t(y)$ ~

$$\sim \quad f_{t+\tau}(y_{t+\tau}) = \int_{-\infty}^{\infty} f(y_{t+\tau} | y_t) f_t(y_t) dy_t$$

ESTIMATING

- The stochastic kernel is estimated in a manner like the univariate distributions are:

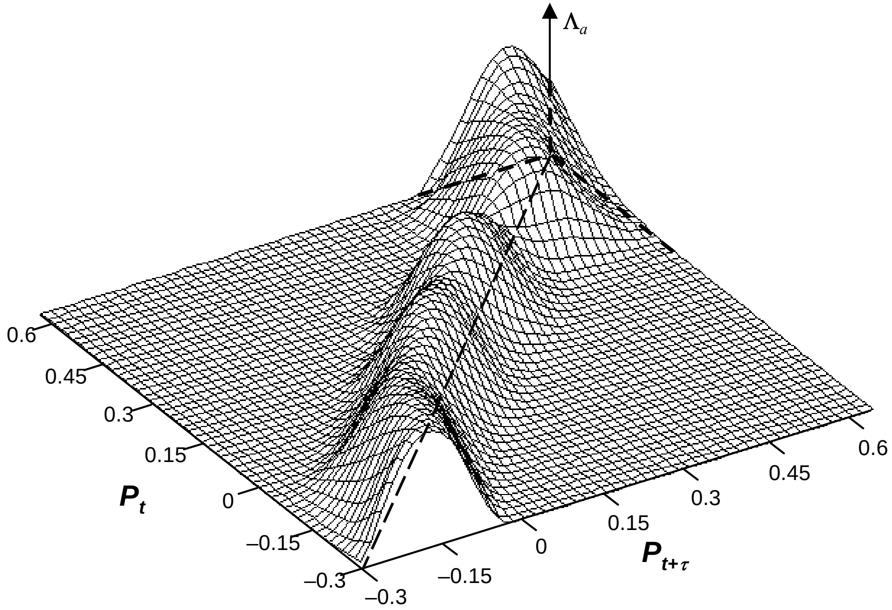
$$\hat{f}(y_{t+\tau}|y_t) = \frac{\frac{1}{Nh^2}\sum_r K\left(\frac{y_{t+\tau} - y_{r,t+\tau}}{h}\right)K\left(\frac{y_t - y_{rt}}{h}\right)}{\hat{f}_t(y_t)}$$

- Estimation of Λ^n**:**

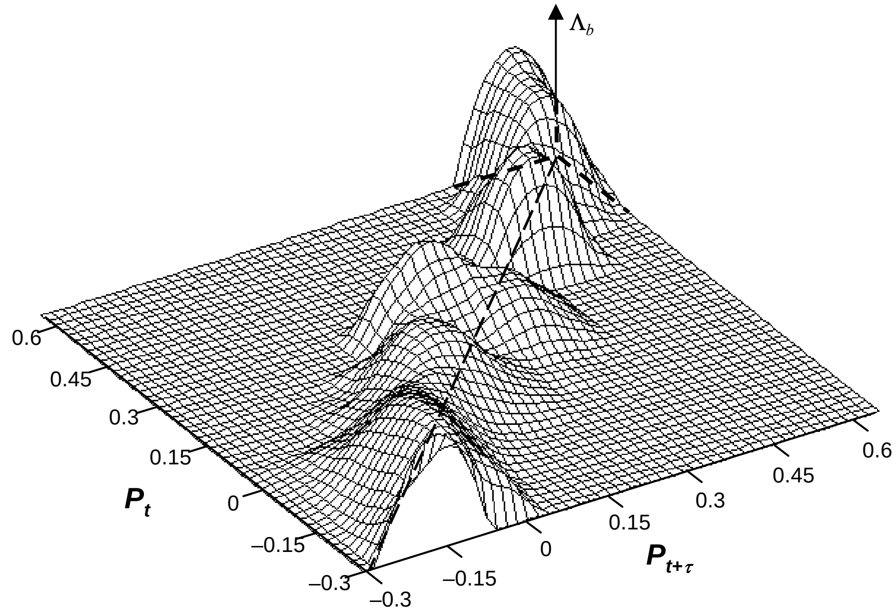
$$\Lambda^n = f(y_{t+n\tau}|y_t) = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f(y_{t+n\tau}|y_{t+(n-1)\tau})\ldots f(y_{t+2\tau}|y_{t+\tau})f(y_{t+\tau}|y_t)dy_{t+(n-1)\tau}\ldots dy_{t+\tau}dy_t$$

- Since Λ^∞ degenerates into $f(y_{t+\infty}|y_t^{(i)}) = f(y_{t+\infty}|y_t^{(j)})$ for each pair $y_t^{(i)}$ and $y_t^{(j)}$, the fulfilment of this condition accurate to 10^{-m} can be used as a criterion of convergence of Λ^n to Λ^∞

STOCHASTIC KERNELS: THE COST OF THE STAPLES BASKET ACROSS REGIONS OF RUSSIA, 2001–2010

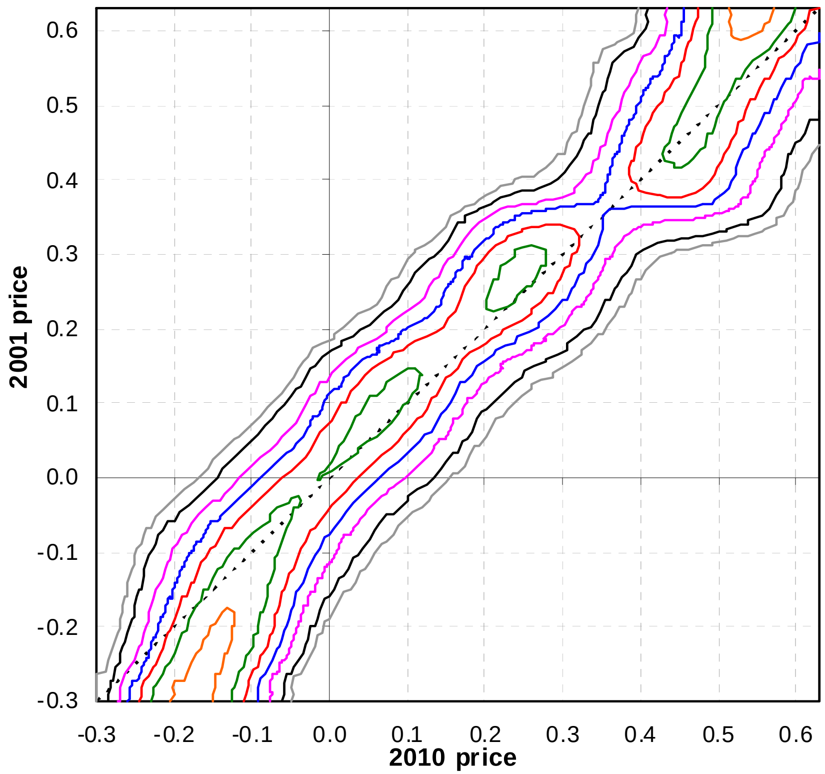


Estimated with the use of nine yearly transitions

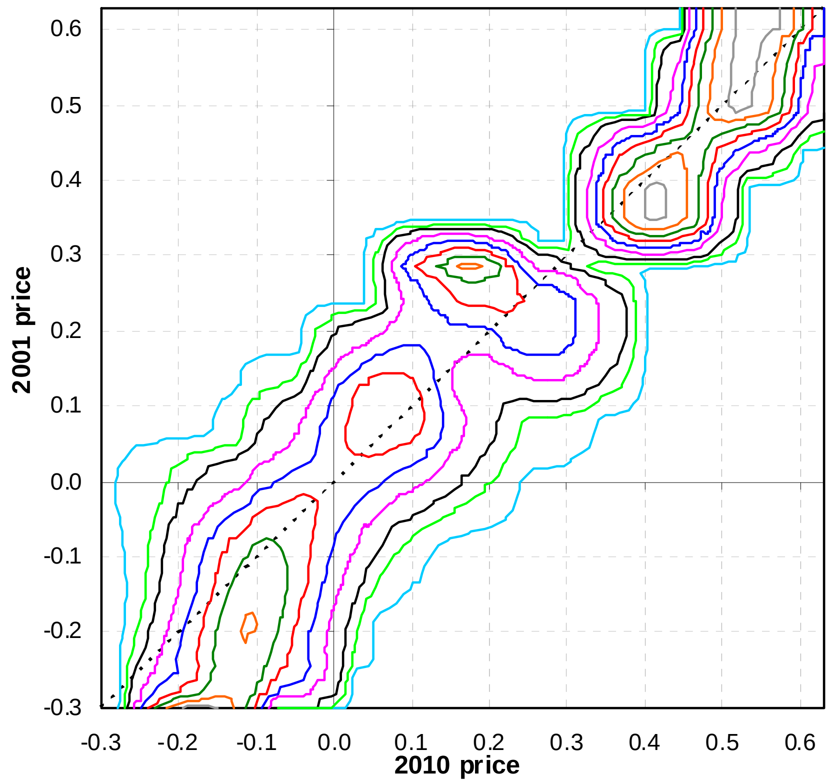


Estimated using the nine-year transition from 2001 to 2010

STOCHASTIC KERNELS, contour plots

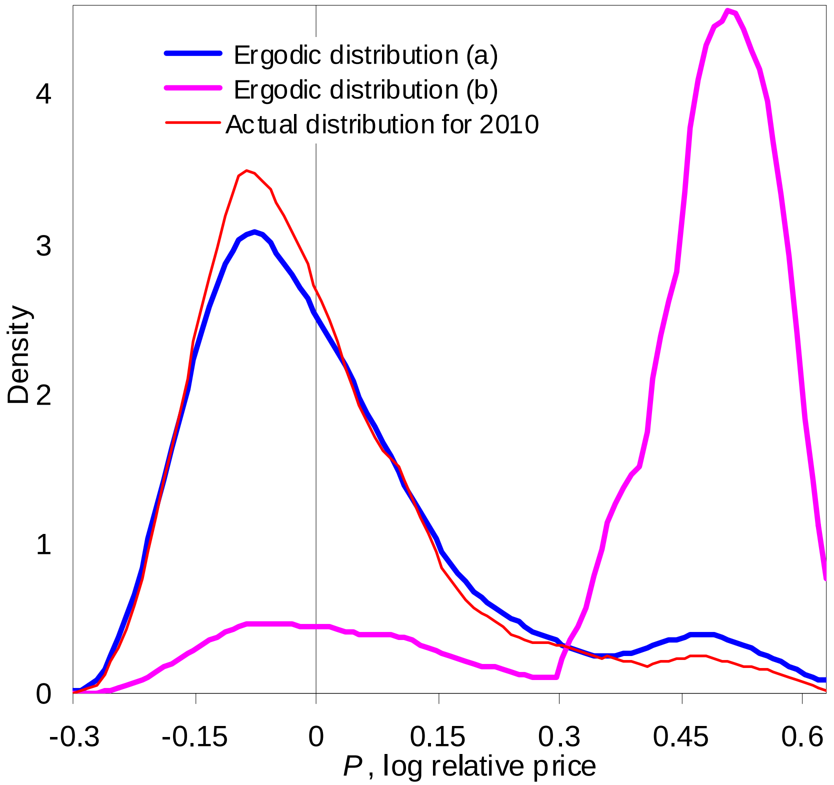


Estimated with the use of nine yearly transitions

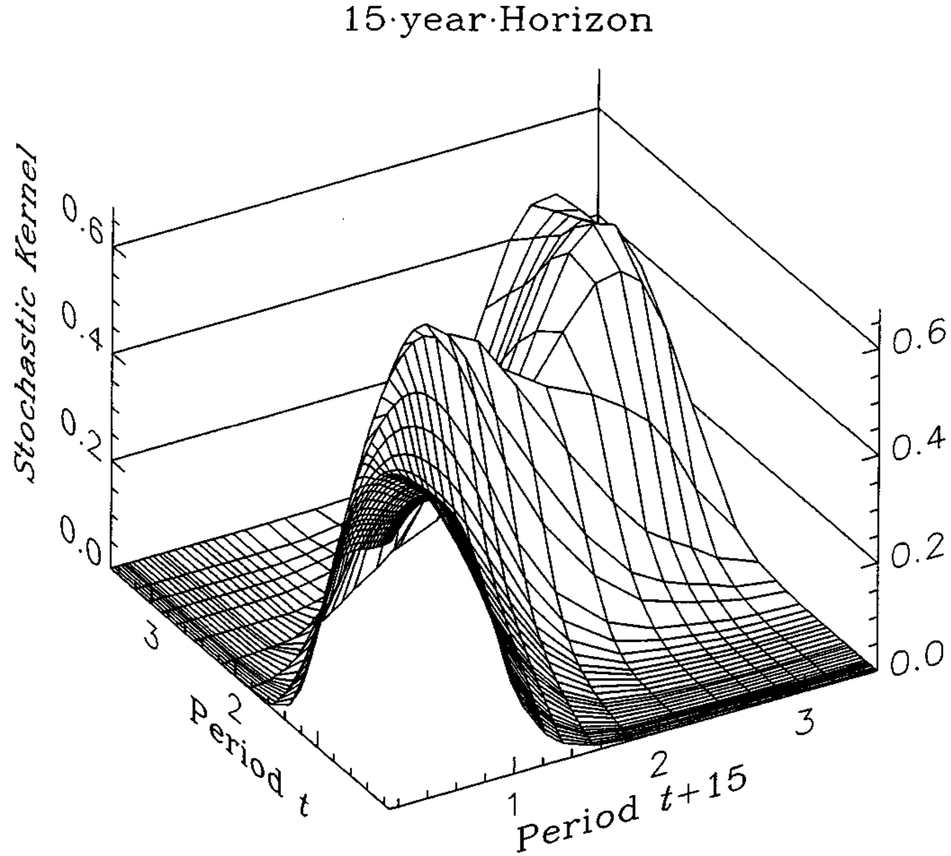


Estimated using the nine-year transition from 2001 to 2010

ERGODIC DISTRIBUTIONS



RELATIVE INCOME DYNAMICS ACROSS 105 COUNTRIES



Danhy Quah. Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs. *Journal of Economic Growth*, 1997, Vol. 2, No. 1