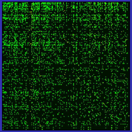
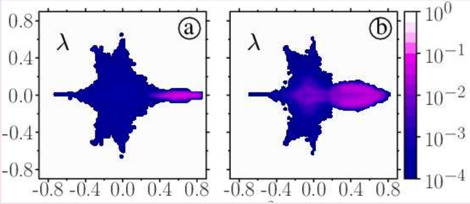
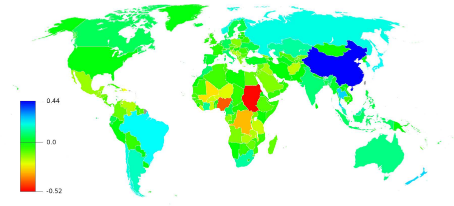
Anderson transition for Google matrix eigenstates



Dima Shepelyansky (CNRS, Toulouse)
www.quantware.ups-tlse.fr/dima
with O.V.Zhirov (BINP Novosibirsk)
and L.Ermann (CNEA TANDAR), K.Frahm (LPT), V.Kandiah (LPT), H.Escaith (WTO Geneve)



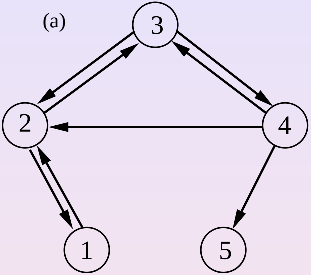


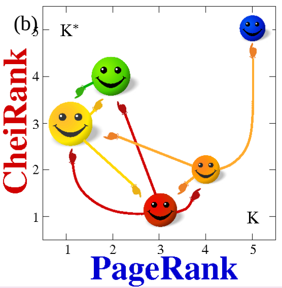
* Markov (1906) → Brin and Page (1998)
* Anderson transition on directed networks
* Applications: multiproduct world trade network (UN COMTRADE + OECD-WTO)

Support: EC FET Open project NADINE
Refs. at www.quantware.ups-tlse.fr/FETNADINE/ + arXiv:1409.0428

Google matrix construction rules

Markov chains (1906) and Directed networks





For a directed network with N nodes the adjacency matrix $\mathbf{A}$ is defined as $A_{ij} = 1$ if there is a link from node j to node i and $A_{ij} = 0$ otherwise. The weighted adjacency matrix is

$$S_{ij} = A_{ij} / \sum_k A_{kj}$$

In addition the elements of columns with only zeros elements are replaced by $1/N$.

Google matrix construction rules

Google Matrix and Computation of PageRank

$\mathbf{P} = \mathbf{SP} \Rightarrow \mathbf{P}$= stationary vector of **S**; can be computed by iteration of **S**.

To remove convergence problems:

- Replace columns of 0 (dangling nodes) by $\frac{1}{N}$:

$$S = \begin{pmatrix} 0 & 1/2 & 1/3 & 0 & 1/5 \\ 1 & 0 & 1/3 & 1/3 & 1/5 \\ 0 & 1/2 & 0 & 1/3 & 1/5 \\ 0 & 0 & 1/3 & 0 & 1/5 \\ 0 & 0 & 0 & 1/3 & 1/5 \end{pmatrix} \qquad S^* = \begin{pmatrix} 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 1/3 & 0 & 1 & 0 \\ 0 & 1/3 & 1/2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- To remove degeneracies of $\lambda = 1$, replace **S** by **Google matrix**

  $\mathbf{G} = \alpha \mathbf{S} + (1-\alpha)\frac{\mathbf{E}}{N}$; $GP = \lambda P$ => Perron-Frobenius operator

- α models a random surfer with a random jump after approximately 6 clicks (usually $\alpha = 0.85$); PageRank vector => P at $\lambda = 1$ ($\sum_j P_j = 1$).

- CheiRank vector P^*: $G^* = \alpha \mathbf{S}^* + (1-\alpha)\frac{\mathbf{E}}{N}$, $G^* P^* = P^*$
  ($\mathbf{S}^*$ with inverted link directions)
  Fogaras (2003) ... Chepelianskii arXiv:1003.5455 (2010) ...

Real directed networks

Real networks are characterized by:

- **small world property**: average distance between 2 nodes $\sim \log N$
- **scale-free property**: distribution of the number of ingoing or outgoing links $\rho(k) \sim k^{-\nu}$

PageRank vector for large WWW:

- $P(K) \sim 1/K^{\beta}$, where K is the ordered rank index
- number of nodes N_n with PageRank P scales as $N_n \sim 1/P^{\nu}$ with numerical values $\nu = 1 + 1/\beta \approx 2.1$ and $\beta \approx 0.9$.
- PageRank $P(K)$ on average is proportional to the number of ingoing links
- CheiRank $P^{*}(K^{*}) \sim 1/K^{*\beta}$ on average is proportional to the number of outgoing links ($\nu \approx 2.7; \beta = 1/(\nu - 1) \approx 0.6$)
- WWW at present: $\sim 10^{11}$ web pages

Donato *et al.* EPJB **38**, 239 (2004)

Anderson transition on directed networks

Anderson (1958) metal-insulator transition for electron transport in disordered solids

$H = \epsilon_n \psi_n + V(\psi_{na+1} + \psi_{n-1}) = E\psi_n \; ; \quad -W/2 < \epsilon_n < W/2$

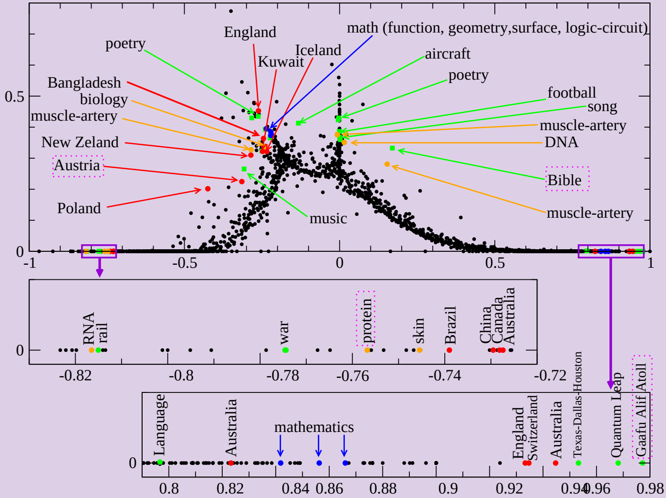
In dimensions $d = 1, 2$ all eigenstates are exponentially localized, insulating phase. At $d = 3$ for $W > 16.5V$ all eigenstates are exponentially localized, for $W < 16.5V$ there are metalic delocalized states, mobility edge, metalic phase

Random Matrix Theory - RMT (Wigner (1955)) for Hermitian and unitary matrices (quantum chaos, many-body quantum systems, quantum computers)

Google matrix, Markov chains, Perron-Frobenium operators:
=> complex spectrum of eigenvalues; new field of research
Can we have the Anderson transition for Google matrix? All the world would go blind if PageRank is delocalized What are good RMT models of Google matrix? Subspaces and core

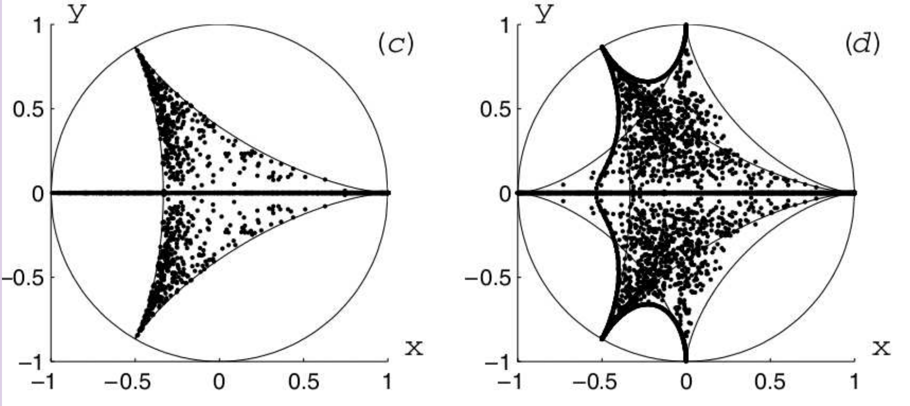
$$\mathbf{S} = \begin{pmatrix} \mathbf{S_{ss}} & \mathbf{S_{sc}} \\ 0 & \mathbf{S_{cc}} \end{pmatrix}$$

Wikipedia spectrum and eigenstates



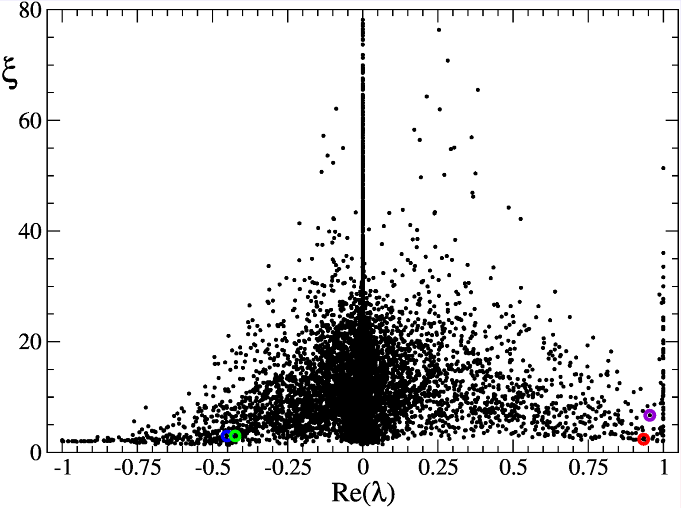
Spectrum S of EN Wikipedia, Aug 2009, $N = 3282257$. Eigenvalues-communities are labeled by most repeated words following word counting of first 1000 nodes. (Ermann, Frahm, DS 2013)

Spectrum of random orthostochastic matrices



Spectrum $N = 3$ (left), 4 (right) [K.Zyczkowski *et al.* J.Phys. A **36**, 3425 (2003)]

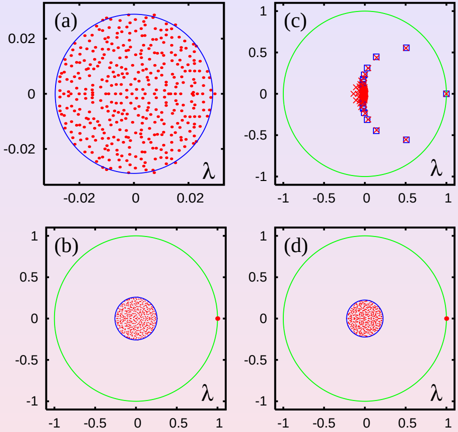
Localization features



Multiproduct world trade network $N = N_p \times N_c = 61 \times 227 = 13847$ (year 2008): small IPR values ξ. Small gaps in S of directed networks.

Random Matrix Models of directed networks

random matrix elements of G with sum equat unity in each column ($N = 400$)



(a) N positive random elements with unit sum in each column;
(c) triangular matrix with random elements;
(b),(d) $Q = 20$ nonzero elements in each column
- blue circle is theory with radius $\sim 1/\sqrt{N}, 1/\sqrt{Q}$

Anderson type models for directed networks

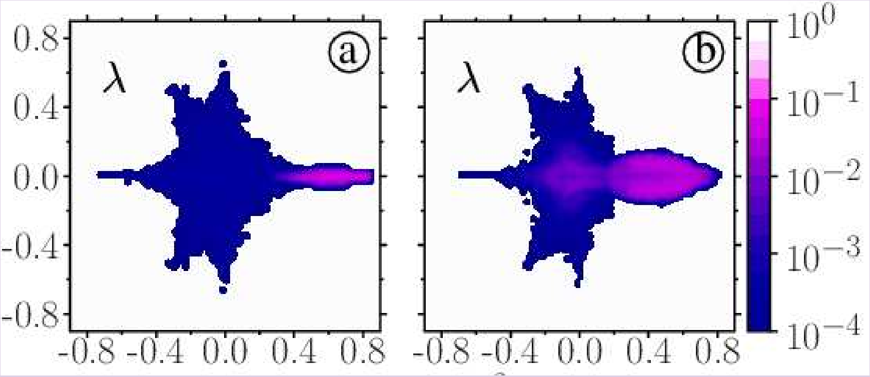
We use the usual Anderson model with diagonal disorder terms W_i and transitions V to nearby sites on a lattice in dimension d:

$$W_{\mathbf{i}}\psi_{\mathbf{i}} + V\psi_{\mathbf{i+1}} + V\psi_{\mathbf{i-1}} = \lambda\psi_{\mathbf{i}} \ , \tag{1}$$

where indexes in bold are vectors in d-dimensional space. On this we construct the matrices S and G for $d = 2, 3$. The matrix S is constructed as follows: each transition matrix element, corresponding to V terms, in the Anderson model in dimension d is replaced by a random number ε_i uniformly distributed in the interval $[0, \varepsilon_{max}/2d]$, the diagonal element $W_{\mathbf{i}}$ is replaced by unity minus the sum of all ε_i over $2d$ nearby sites $(1 - \sum_{i=1}^{2d} \varepsilon_i)$. The asymmetric matrix S constructed in this way belongs to the Google matrix class.

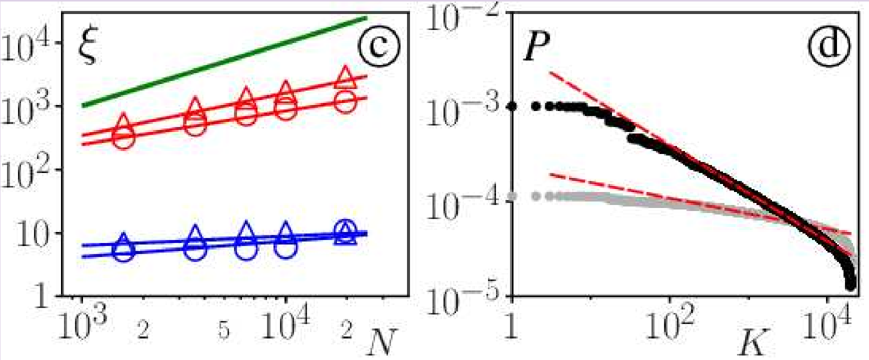
By replacing matrix elements in the model AD2 by blocks B of size 4×4 we obtain the **model AD2Z**. In a similar way we obtain the **model AD2ZS** with block shortcuts. In this case we restrict our studies only for dimension $d = 2$.

Anderson transition on directed networks



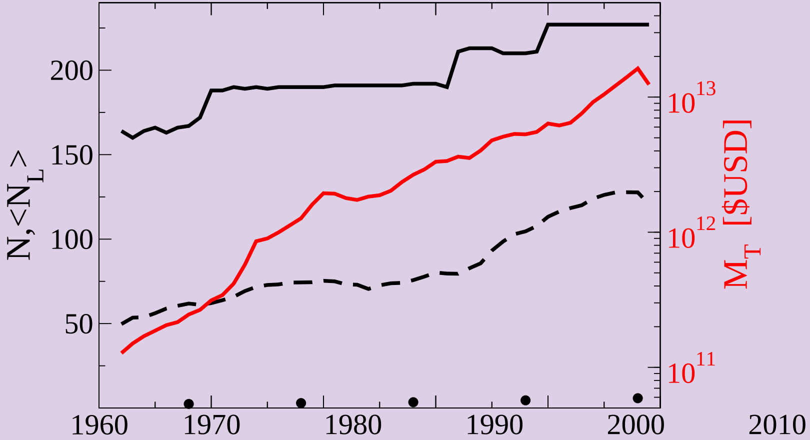
Panels show distribution of IPR values ξ (number of nodes contributing to an eigenstate) on the plane λ or eigenvalues of *G* matrix for two nodels of directed networks with disorder; color shows the ratio ξ/N, $\alpha = 0.85$.

Anderson transition on directed networks



Panel (c): dependence of ξ on *N* for AD2Z with triangles for states with λ located in the delocalized domain $\mathrm{Re}\lambda \in (0.3, 0.85)$ (red triangles, fit gives $\nu = 0.67$) and in the localized domain $\mathrm{Re}\lambda < -0.5$ (blue triangles, $\nu = 0.15$); for AD2ZS at $\delta = 0.25$ with circles for states with λ located in the delocalized domain $\mathrm{Re}\lambda \in (0.2, 0.85)$ (red circles, $\nu = 0.53$) and in the quasi-localized domain $\mathrm{Re}\lambda < -0.5$ (blue circles, $\nu = 0.25$); fits are shown by lines, green line shows $\xi = N$. Panel (d): dependence of PageRank probability *P* on PageRank index *K* for models AD2Z (gray symbols) and AD2ZS at $\delta = 0.25$ (black symbols); the fits for the range $K \in (100, 6000)$ are shown by dashed lines with $\beta = 0.16$ (AD2Z) and $\beta = 0.51$ (AD2ZS) for the parameters of panels (a,b).

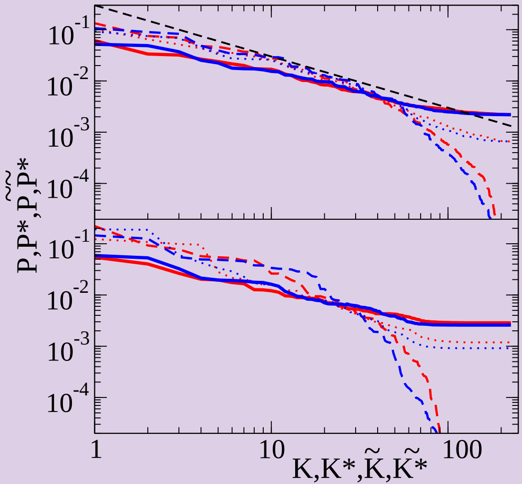
World trade network (WTN) of United Nations COMTRADE 1962-2010



Number of countries (black), links (dashed/points) and mass volume in USD (red)

Leonardo Ermann, DS arxiv:1103.5027 (2011)

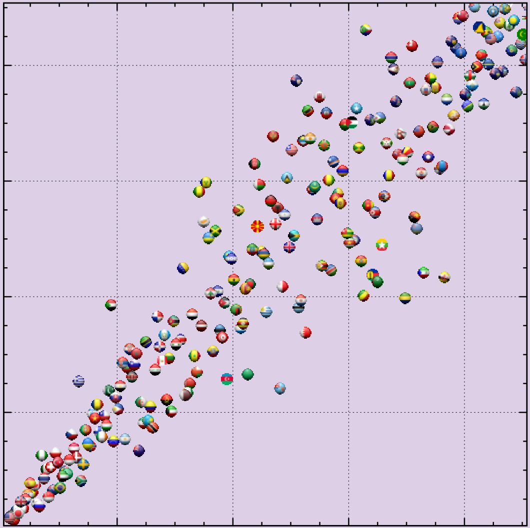
PageRank, CheiRank of World Trade



Year 2008: Probabilities of PageRank $P(K)$ (red), CheiRank $P^*(K^*)$ (blue) for all commodities (top) and crude petroleum (bottom), $\alpha = \mathbf{0.5}$; 0.85 (full/dotted); (dashed curves are for ImportRank, ExportRank); dashed line Zipf law $P \sim 1/K$; 227 countries

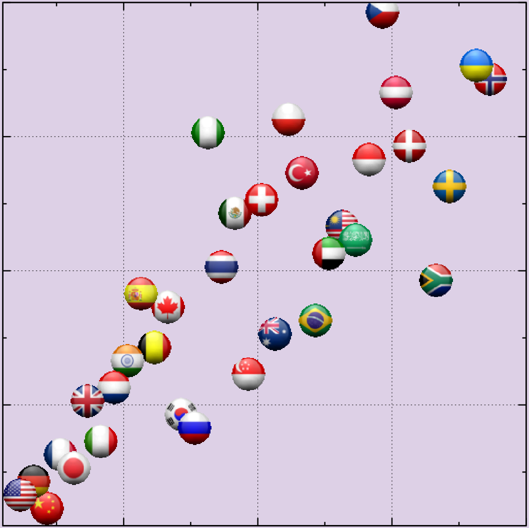
Ranking of World Trade

2008: All commodities



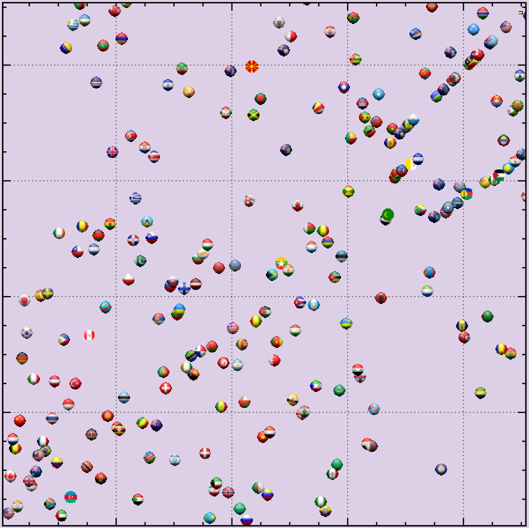
Ranking of World Trade

2008: All commodities



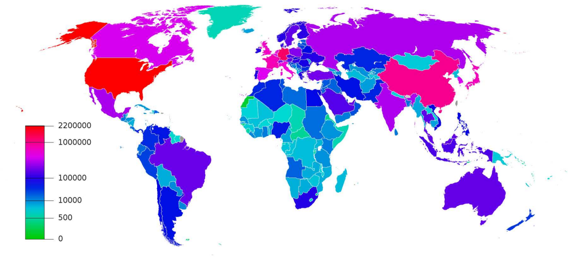
Ranking of World Trade

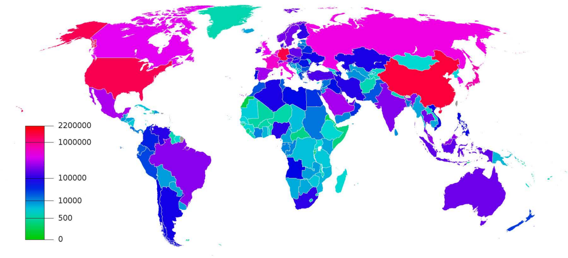
2008: Crude petroleum



Multiproduct WTN + WNEA of WTO

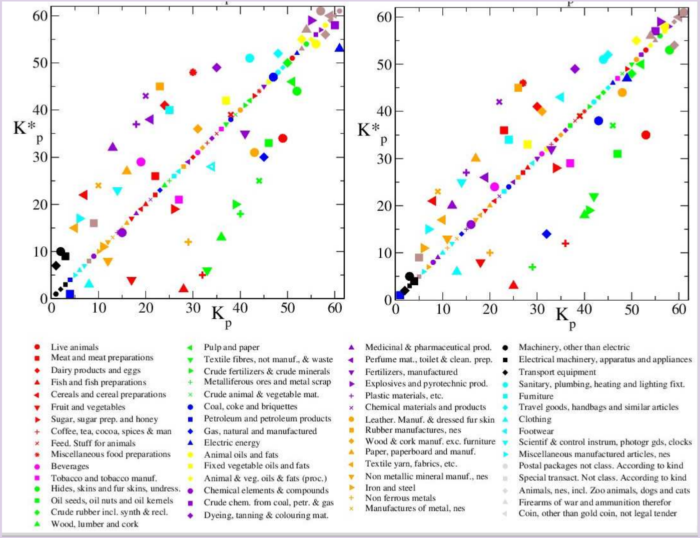
Example: year 2008, $N_c = 227$, $N_p = 61$, $N = 13847$, import (top) - export (bottom) in millions USD





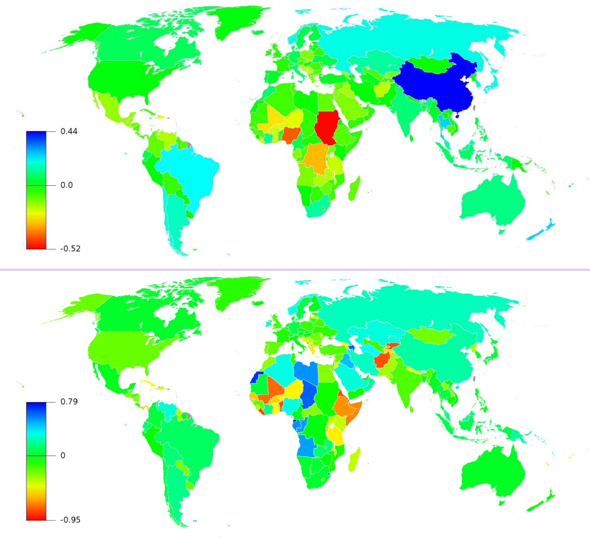
Multiproduct WTN: ranking of products

Democracy in countires, volume fraction in products => personalized vector in *G*. Left: 1993, right: 2008



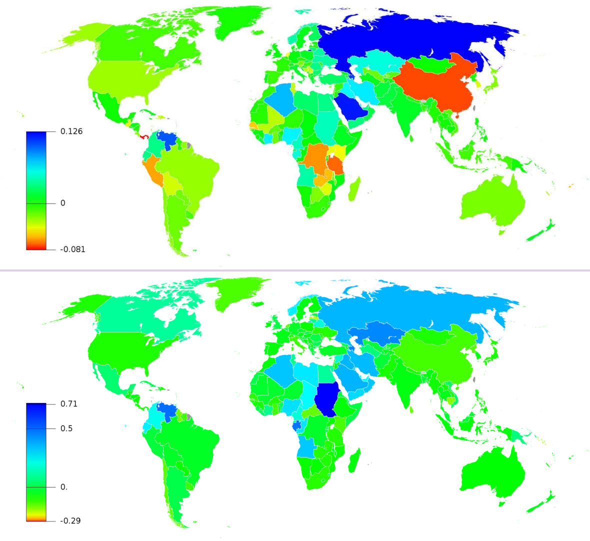
CheiRank-PageRank balance (2008)

$B_c = (P_c{}^* - P_c)/(P_c{}^* + P_c)$ (top - CheiRank-PageRank; bottom -Export-Import volume)



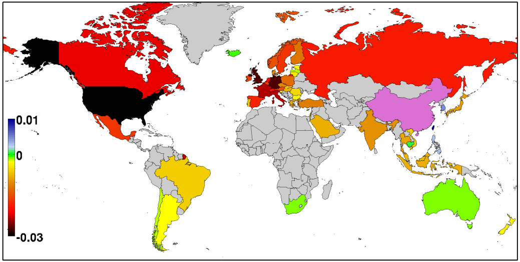
Sensitivity to price of petroleum (2008)

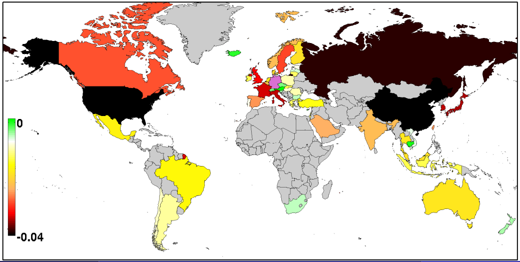
$dB_c/d\delta_p$ (top - CheiRank-PageRank; bottom -Export-Import volume)



WNEA of OECD-WTO (2008)

World network of economic activities: countries $N_c = 58$, activity sectors $N_s = 37$ (with V.Kandiah and H.Escaith (WTO Geneve)); $dB_c/d\sigma_{c'}$ sensitivity to labor cost of China (top), Germany (bottom)





Further applications of Markov chains and Google matrix ?

What would say Richard Prange? What will reply Shmuel Fishman ?

At celebration of 70th of Boris Chirikov, 17 July 1998, Toulouse