AGjDPG 3: Focus Session: Big Data (Contributed Talks)

Time: Monday 17:15–18:45 Location: HE 101

AGjDPG 3.1 Mon 17:15 HE 101

Quantifying the behavior of stock correlations under Market stress — Tobias Preis[1,2,3], •Dror Kenett[4], H. Eugene Stanley[1], Dirk Helbing[2,5], and Eshel Ben-Jacob[4] — [1]Center for Polymer Studies, Department of Physics, 590 Commonwealth Avenue, Boston, MA 02215, USA, — [2]Chair of Sociology, in particular of Modeling and Simulation, ETH Zurich, Clausiusstr. 50, 8092 Zurich, Switzerland — [3]Artemis Capital Asset Management GmbH, Gartenstr. 14, 65558 Holzheim, Germany — [4]School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv, Israel — [5]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Understanding of correlations in complex systems is crucial in the face of crises, such as the ongoing financial crisis. However, in complex systems, such as financial systems, correlations are not constant but instead vary in time. Here we address the question of quantifying state-dependent correlations in stock markets, since reliable estimates of correlations are most needed to protect a port- folio. We find the striking result that the average correlation among stocks belonging to a given index scales linearly with market stress reflected by the normalized value of the index return. Consequently, the diversification effect of the portfolio value melts away in times of market losses, just when it would most urgently be needed. Our empirical findings could be used to anticipate diversification break- downs leading to protected individual portfolios and could contribute to increased stability of financial markets in general.

AGjDPG 3.2 Mon 17:30 HE 101

Quantifying Trading Behavior in Financial Markets Using Google Trends — Tobias Preis[1,2,3], •Helen Susannah Moat[1,4], H. Eugene Stanley[2], and Dirk Helbing[1,5] — [1]Chair of Sociology, in particular of Modeling and Simulation, ETH Zurich, Clausiusstr. 50, 8092 Zurich, Switzerland — [2]Center for Polymer Studies, Department of Physics, 590 Commonwealth Avenue, Boston, Massachusetts 02215, USA — [3]Artemis Capital Asset Management GmbH, Gartenstr. 14, 65558 Holzheim, Germany — [4]Department of Mathematics, UCL, Gower Street, London, WC1E 6BT, UK — [5]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA

Crises in financial markets affect humans worldwide. Detailed market data on trading decisions reflect some of the complex human behavior that has led to these crises. We show that massive new data sources resulting from human interaction with the Internet offer a different perspective on the behavior of market participants in periods before large market movements. Analyzing Google query volumes for search terms related to finance, we find that we can identify early warning signs of stock market moves. We further demonstrate that a trading strategy based on our analysis can outperform the market. Our results illustrate the unprecedented potential that combining extensive behavioral data sets offers for a better understanding of collective human behavior.

See also: http://www.tobiaspreis.de; http://www.suzymoat.co.uk; http://polymer.bu.edu/hes; http://www.soms.ethz.ch/people/dhelbing

AGjDPG 3.3 Mon 17:45 HE 101

Complex Economic Behavior Captured by Big Data — •Tobias Preis — Chair of Sociology, in particular of Modeling and Simulation, ETH Zurich, Clausiusstr. 50, 8092 Zurich, Switzerland — Center for Polymer Studies, Department of Physics, 590 Commonwealth Avenue, Boston, Massachusetts 02215, USA — Artemis Capital Asset Management GmbH, Gartenstr. 14, 65558 Holzheim, Germany

Financial market fluctuations are characterized by abrupt switches between upward trends and downward trends, which can last for hundreds of days or just a few minutes. Here, we address two questions. Firstly, can these ubiquitous switching processes be quantified? Secondly, what insight do new big social data sources offer into this collective behavior?

www.tobiaspreis.de

AGjDPG 3.4 Mon 18:00 HE 101

Who writes Wikipedia, a data-driven modeling of Wikipedia editorial activity — •Taha Yasseri, Róbert Sumi, and János Kertész — Institute of Physics, Budapest University of Technology and Economics, Budapest, Hungary

Recently developed Internet-based technologies facilitate collective cooperation of individuals more than ever. Among all possible examples, Wikipedia, the free encyclopaedia written by unknown volunteers from all around the world, is the one, in which self-organised value formation occurs with no external supervision or guidance. Moreover, similar to other online societies, the complete set of information about the activity of individuals is digitalised and collected in an easily accessible way, so that statistically analysing this big data corpus is achievable now.

In this contribution, we report on our recent investigations [1] on the accumulated data of the activity of Wikipedia editors in 34 languages (more than 1 billion records from more than 25 million editors), (i) to observe and evaluate the universalities and differences among various societies of editors and, (ii) to have an estimation of the geographical distribution of editors in the globe, based on the data-driven model for editor activities. We believe such studies could shed light to unknown aspects of the Wikipedia development process, its biases and limitations.

[1] T. Yasseri, R. Sumi and J. Kertész, Circadian patterns of Wikipedia editorial activity: A demographic analysis, to appear in *PLoS ONE, pre-print: arXiv/1109.1746.*

AGjDPG 3.5 Mon 18:15 HE 101

Analysis and modeling of human behavior observed in cyber space communication data — •kenta yamada[1], yukie sano[2], hideki takayasu[3], and misako takayasu[4] — [1]Waseda University, Japan — [2]Nihon University, Japan — [3]Sony CSL, Japan — [4]Tokyo Institute of Technology, Japan

Analysis and modeling of human behavior become major targets of twenty-one century science. Especially, huge data of articles in cyber space such as blogs and twitters are attracting attention because the data directly reflect trends and topics in the society.

In this presentation, we report our analysis and modeling of the huge blog database which contains 300,000 bloggers including 70 million articles from 11/01/2006 to 8/31/2011. We observed some characteristic patterns in appearance frequency of some key words per day. We categorize them into three patterns: ordinary words, news words and trendy words. Ordinary words like adverbs are characterized by stationary processes, while the frequency of the news words and trendy words are characterized by non-stationary processes. A news word such as "tsunami" shows a sharp increase by sudden appearance of news and the number decays slowly following a power law. In the case of a trendy word, the number of entries per day increases exponentially.

In order to understand the origin of these characteristic motions, we introduce a state transition type agent-based model similar to the SIR (Susceptible-Infected-Recovered) model which is a basic epidemic model. We show that our simple agent-based model reasonably reproduces these three typical patterns.

AGjDPG 3.6 Mon 18:30 HE 101

Mapping dietary Patterns and their Transitions: Implications for the Environment — Prajal Pradhan, •Dominik E. Reusser, and Jürgen P. Kropp — Potsdam Institute for Climate Impact Research

This study analyzes global, long term data on food consumption per country to identify typical patterns of diets. From these patterns, we derive typical food transition pathways on a global scale. Subsequently we assess the environmental consequences from green-house-gas (GHG) emission and anthropogenic inputs.

We used Self Organizing Map to identify the patterns and the transition pathways based on supply of 12 food groups from FAOSTAT dataset for a period 1961-2007. Data on energy output/input ratio for crop production and agricultural emission were used to estimate fossil energy used and GHG emission associated with the patterns.

We identified nine typical dietary patterns consisting high, moderate and low food intake with varied compositions along with a typical food transition pathway with one bifurcation. As expected, the high dietary patterns require higher fossil energy and lead to higher GHG emission. However related non-CO2 GHG emission intensities are relatively low.

Changes in dietary patterns are a part of the global change processes. Identification of past transitions is way to anticipate possible future transitions, which may supports policy processes.