ECOFORUM

[Volume 9, Issue 1(21), 2020]

IS BIG DATA A BIG DEAL FOR CAPITAL MARKET COMPANIES IN THEIR TRANSFORMATION PROCESS?

Davit GAMKRELIDZE
Ilia State University, 0162, Georgia
davit.gamkrelidze@iliauni.edu.ge

Abstract
The following article discusses the importance of using big data, especially in the operation of capital market companies, both in terms of benefits and potential risks. Given the growing dynamic business environment, capital market companies have to transform their operations in order to accommodate the raising demands. Fast business decision making is of particular importance in this process. Structured use of data plays a major role in decision-making, especially as the amount of large digital data in the modern world grows at an unprecedented rate.
Author of the article focuses on the statistical and econometric techniques required for the analysis of big data. The article also highlights some use cases and the growing interest of capital market companies in introducing big data analytical technologies and the relevant challenges and benefits. In addition, using so-called "Simpson's Reversal Paradox" author explains that using big data and digging deep into details might be counterproductive and lead to loss of global picture and wrong decision-making.

Key words: *Analytical techniques; Big Data; Capital Markets; Simpson's Reversal Paradox; Electronic Trading*

JEL Classification: *C55, C70, C80, G10, G11, G15, G20*

I. INTRODUCTION

From the 20th century onward, along with the growth of globalization, the in-depth study of the evolution, role and influence of international financial markets became the subject of consideration of many scholars and research institutions (Alfaro, Chanda et al, 2000; Pagano, 1993; Prati and Schinasi, 1997; Dicle and Levendis, 2013); Nowadays, achieving significant economic development and progress in the modern world has become practically impossible without efficiently functioning, structured and properly regulated financial markets. For example, by using data on 49 countries from 1976 to 1993, the World Bank Policy research department found that stock market liquidity - as measured by stock trading relative to the size of the market and economy - is positively and significantly correlated with current and future rates of economic growth, capital accumulation, and productivity growth within the country (Levine and Zervos, 1996).

"One of the most significant economic developments of the last quarter century was the convergence and integration of financial markets and previously separate financial market segments. In the context of globalization, there is growing interaction between them between the insurance market, banking and the stock market" (Morosan, 2015, p.1). In general, financial markets allow the parties to trade with various financial instruments and assets, to manage and insure financial risks and to conduct an analysis and investment activity. A modern company can reach investors and raise money in a different ways. As a result, flow of funds within an economy, and the factors that having effect on this flow, might have an important role in the economic environment in which a company operates. Although we speak mainly about corporate firms, many different entities can raise money in the capital markets. These might include sovereign governments, local authorities, international financial institutions, supranational bodies (Choudhry, Joannas, et al, 2010).

Given the growing dynamic business environment in the modern digital world, proper analysis of data is increasingly important in the operations of companies presented on the financial markets. Therefore, in recent decades, the principles of financial markets in combination with the raising role of big data have been actively researched by world's leading universities and other solid international organizations. There are academic publications and analytical/policy reports on this topic out there, which gives indication on the increased scientific significance of the topic (Guerard and Rachev, 2013; Jothimani, Shankar et al, 2014).

Big data basically is the information asset characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value (Mauro, Greco et al, 2016 p.1). Over the last decades, the concept of big data has been applied to many different industries. For example, big data technologies have enabled more creative and innovative development for smart cities (Rotuna, Cirnu et al, 2017). Also, "the shifts to practice-based evidence for medicine, and data-driven rather than strictly hypothesis-driven biomedical research, represent the big changes driven by big data" (Sanchez and Verspoor, 2014, p. 18).

Further evidence of the above-mentioned is a summary of the number of papers published in 2010-2015 with "big data" as the subject and listed in the CNKI and Web of Science databases, illustrated in Figure 1:

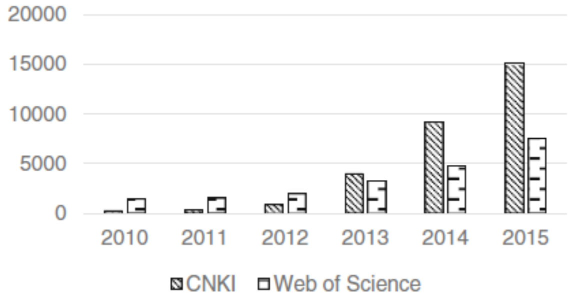


Figure 1 - Number of papers published in 2010 - 2015 with big data as the subject (Ye and Li, 2017).

Going forward, if we narrow focus and concentrate on only four subject areas such as business and management, computer science, decision science and social science, 227 articles have been published from 1996 to 2015. Figure 2 illustrates an abrupt increase in number of journal articles in the big data and big data analytics research area from 2013 onwards until 2015 and figure 3 demonstrates that the vast majority of the publications are research papers (Sivarajah, Kamal et al, 2017):

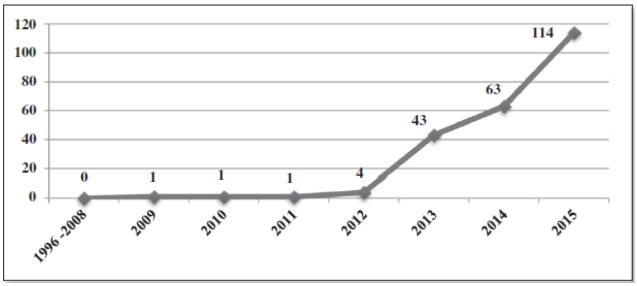


Figure 2 – Total number of papers published (from 1996 to 2015)

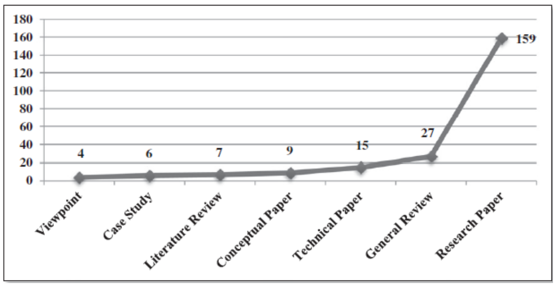


Figure 3 – Classification of publication types (from 1996 to 2015)

Source: Sivarajah, Kamal et al, 2017

Keeping focus on our research topic, from the different types of financial markets, it is especially interesting to explore the capital market[6]-or long-term securities market, because capital markets have been relatively slow to adopt big data strategies. Considering the sector of financial services, big data has gained far more traction within retail banking because of the increasing desire of these financial institutions to profile their customers in a similar manner to pioneers of big data adapters such as Google or Amazon. There is less incentive to apply big data in this manner on the institutional side of the capital markets (O'Sheaa and Shah, 2014).

Capital Market companies, like investment banks, hedge funds and others have always been quite data-intensive. However, since the emergence of issues related to big data and relevant business intelligence, the way those companies use data changes quite apparently. They have never been data-poor but the necessity to explore market conditions tremendously quickly is changing the essence of financial trading companies and transforming them from data consumers to data-driven firms, depending on the skills and expertise of data scientists and engineers, in addition to financial analysts and experts (Bank for International Settlements, 2016).

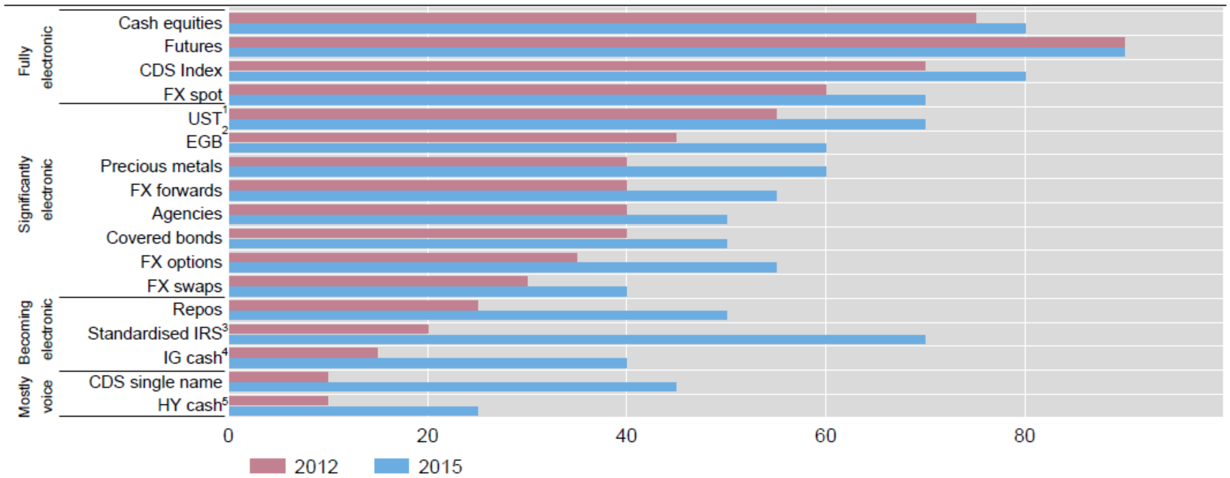
The sections below illustrate how electronic trading has grown in fixed income markets, and therefore how it generates a big set of data, that require specific techniques to analyze. Article gives insight into the benefits and challenges of big data for capital market companies, including the risk of losing global picture, explained through the Simpson Paradox.

[6] Capital market is a financial marketplace for its participants to deal with selling and buying financial instruments; Commonly capital markets comprise the bond markets and stock (also called equity) markets. Capital market participants basically include: Commercial Banks, Investment banks, Credit Unions, Mutual Funds, Pension Funds, Insurance Companies, Leasing Companies, Stock Exchanges and etc

II. Evolution of Electronic Trading and Data Analysis Techniques

It is worth noting how electronic trading[7] evolved in fixed income markets. Traditionally, fixed income trading has been organized around large banks or securities houses, called dealers and their clients. However, fixed income markets experienced a major shift starting in the late 1990s as electronic communication networks (ECNs[8]) started to evolve. (Bank for International Settlements, 2016, pp.4-5).

Electrification in some fixed income segments is now nearly as developed as in cash equities and foreign exchange, as illustrated in Figure 4:



[1] US Treasuries. [2] European government bonds. [3] Standardised interest rate swaps. [4] Investment grade cash bonds. [5] High-yield cash bonds.

Sources: Greenwich Associates (2014); McKinsey & Company and Greenwich Associates (2013).

Figure 4 - State of electrification in various asset classes (Bank for International Settlements, 2016, pp.4-5).

Electronic trading is just one example of increased digital universe, as a result of which data is growing at a tremendous rate. These changes led to significant increase of the data flow, which on the other hand activates doubts on the quality of data. Basic qualitative characteristics of the data are: Completeness, Consistency, Accuracy, Validity and Timeliness (Dejaeger, K., Hamers, B. et al, 2010). Logically the higher the quality of the data is the higher is the chance of organizational success, because of the reliance on fact-based decisions, instead of human intuition and habit. The increase in data is a minor problem, in itself, while the increase in percentage of unstructured data, in the overall data volume, is what is concerning all, including big exchanges (Hammer, Kostroch et al, 2017).

Big data require specific technology and analytical methods for its transformation into value. Therefore, big data becomes a big deal for the companies. In order to make this data understandable and presentable, companies have to go through essential steps, like e.g.:

Data Acquisition and Warehousing - The integral complexity of BD and exponentially growing demands develop unprecedented problems in BD engineering such as data acquisition and storage (Wang &Wiebe, 2014);

Formatting and cleaning – identifying one-off transactions in order to get sophisticated picture and reasonable trends; Formatting and cleaning is essential for proper understanding of data; Advocates of BD and BDA perceive that in identifying a better way to mine and clean the BD can result in big impact and value (Chen, Chiang et al., 2012);

Data aggregation and integration - this process challenge relates to aggregating and integrating clean data mined from large unstructured data (Sivarajah, Kamal et al, 2017);

Data visualization and interpretation - this step makes data understandable for users that is the data analysis and modelling results are presented to the decision makers to interpret the findings for extracting sense and knowledge (Simonet, Fedak, & Ripeanu, 2015).

Out of those stages mentioned above, last one might be perhaps the most challenging one from the analytical point of view. Depending on the specific needs and circumstances, different techniques for data

[7] The term "electronic trading" refers to the transfer of ownership of a financial instrument whereby the matching of the two counterparties in the negotiation or execution phase of the trade occurs through an electronic system.

[8] An ECN is a system that electronically matches buy and sell orders for securities. ECNs operate as virtually centralized marketplaces, aggregating offers to trade and matching them against incoming trade requests

analysis might be employed, many of which fall under methodology categories of econometric/time series, statistics and related fields:

Table 1

Descriptive Predictive Prescriptive Analytics	⟺	What has happened What could happen What should we do
Mathematical and Statistical models	⟺	Probabilistic models based on Math and Statistics are usually used to define relationships between variables. In some cases that relationship can be expressed exactly, while in other cases a random or another probabilistic component must be factored into the model. Regression and statistical models estimating the probability of how variables will affect a specified outcome.
Econometric Models	⟺	Statistical models used in econometrics, specifying the statistical relationship that is believed to hold between the various economic quantities pertaining to a particular economic phenomenon.
Markov Model	⟺	A stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event.
Game Theory Models	⟺	Mathematical models of strategic interaction between rational decision-makers
Parametric and Non-Parametric Recurrent Data Analysis	⟺	Used in various applied fields including economics and business. Capturing the trend, estimating the rate and predict the total number of recurrences, As well as displaying non-parametric graphical estimates of the mean cumulative number or cost of recurrence per unit versus age
Logical Dynamics and Dynamical Systems	⟺	Describing the behavior of the complex dynamical systems, usually by employing differential equations or difference equations.
Principal Components Analysis	⟺	Statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables (entities each of which takes on various numerical values) into a set of values of linearly uncorrelated variables called principal components

(Huisman 2017)

(Guidolin, 2012)

Given the complexity and scale of the information available to be scrutinized, it's quite natural that firms are looking to computers to do more of the work. Machine learning[9] is hardly a new concept in computer science, but with the growth of big data technologies, it is now being applied in real life situations. For example, a predictive machine learning technique is used to examine the financial news articles using textual representations (Kavitha, Vadhana et al, 2015).

III. Simpson's Reversal Paradox and Risk of Misleading Data for Capital Market Companies

The data analysis techniques presented in previous section provide an abstract models that organize data elements and explore how they relate to each other as well as to real world entities properties. Simply saying, they help us to go deep inside the data and investigate trends.

However, one key moment here is the fact that, when you dig deeper into the data, new risk arises – and the risk is losing control over broader picture. This phenomenon is known as "Simpsons Reversal Paradox", which was first described by Edward H. Simpson in a technical paper in 1951 (Pearl, 1999).

Let's simplify our discussion explain how Simpsons Paradox might affect capital market companies. Imagine an institutional investor wants to make a profitable investment in one of the two corporations: A or B. Assume Investor measures financial performance of those companies by quarterly Return On Equity (ROE), calculated as net income (NI) divided by shareholders' equity (SE). Below is presented the relevant data and ROE of each of those two companies, for each quarter of the year:

Table 2 (data is in $thousands)

	Quarter 1			Quarter 2			Quarter 3			Quarter 4		
	SE	NI	ROE	SE	NI	ROE	SE	NI	ROE	SE	NI	ROE
Company A	10,000	7,000	70%	9,000	7,000	78%	4,000	1,000	25%	5,000	2,000	40%
Company B	4,000	3,000	75%	5,000	4,000	80%	10,000	3,000	30%	9,000	4,000	44%

[9] Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.

It seems company B is more profitable in each quarter of the year, isn't it? Obviously, according to the ROEs calculated, Company B shows better performance than company A - and this is true for each quarter. But still, do you see anything paradoxical here? If you don't, then let's make our observation simpler. Let's present same data in terms of two halves of the year, where:

Half 1 = Quarter 1 + Quarter 2

Half 2 = Quarter 3 + Quarter 4

Once we sum up the data of individual quarters, new ROE ratios are calculated in a same manner: net income divided by shareholders' equity per same half. Now, we get the following picture:

Table 3 (data is in $ thousands)

	First Half of Year (Quarter 1&2)			Second Half of Year (Quarter 3&4)		
	SE	NI	ROE	SE	NI	ROE
Company A	19,000	14,000	74%	9,000	3,000	33%
Company B	9,000	7,000	78%	19,000	7,000	37%

The table still shows that company B is still more profitable than company A.

And still nothing paradoxical here?

Then let's make our observation more concise and general – let's have a look on the whole year statistics rather than halves or quarters. So if you sum up the figures of all quarters/halves and calculate data for the total year, we will end up with the following table:

Table 4 (data is in $ thousands)

	Year All Quarters		
	SE	NI	ROE
Company A	28,000	17,000	61%
Company B	28,000	14,000	50%

Here is the key - unexpectedly the picture has changed dramatically. Now yearly data shows that company A is actually better/more accurate than company B, even though table 4 consists of exactly same data which were used in previous tables.

Above example shows that data might be misleading and it applies many different companies on capital market. For example, "in investment management, it might seem obvious, at least at first glance, that positive returns in every sector guarantee that active total returns will be positive. However this is not always true. Beating every sector can result in total returns below the benchmark. Data compelling for one hypothesis can be reversed in every subpopulation thus supporting the contrary position like it is shown below" (Bassett, 1996):

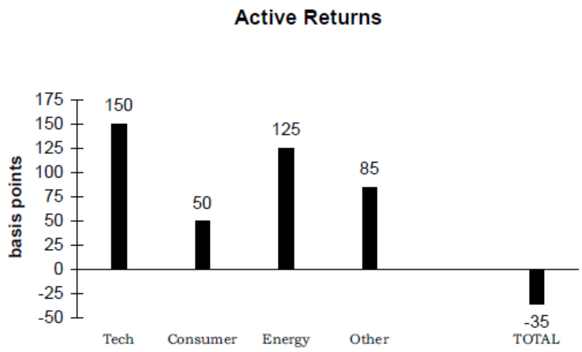


Figure 5: Comparison of active returns of the investment portfolio with S&P500 by sectors

Simpson's paradox can be found in economics, which is described in Appendix A of this article.

Above examples reveal that data might be misleading. This is what we mean under the risk of losing global picture. In other words, it is good to apply special techniques and go deeper into the data, but there is a risk of losing bird's eye view.

As big data are increasingly embedded into social and spatial decisions, processes, and institutions, the links between signifier and signified might become ever more obfuscated. That is, as we begin to place more and more trust in big data and the software, algorithms, and machines that are used produce and analyze such data, we may tend to lose sight of the very things that such data represent (Graham and Shelton, 2013).

IV. Other Opportunities, Challenges and Possible Solutions for Big Data in Capital Market

In capital market companies big data strategies are usually focused on three main directions - focused on revenue generation, meeting compliance requirements, and operational efficiency and cost reduction. For example, Sentiment analysis a good illustration of a revenue generating focus of big data. Specifically, big data strategy can be used to gather and process information to create a clear picture of market tendencies and sentiments in order to use relevant trading strategies, as well as to value individual securities. More recently however, firms have focused on increasing their capabilities to perform news analytics, thus coming up with the use of big data strategies. (O'Sheaa and Shah, 2014, pp.27-28).

In order to avoid the loss of global picture mentioned in the previous section and at the same time maintain control over details, companies use sophisticated data visualization tools and softwares. There are number of data visualization tools out there, lik e.g.: Qlikview, Spotfire, IBM, SAS, Microsoft Power Business Intelligence, Tableau, Excel, D3 and so on (Kandel, Paepcke, et al, 2012).

Those basically include dashboards providing a comprehensive picture, namely 360-degree view for business users with their most important metrics in one place, with updating feature in real time and available on many different devices. Users can explore the data behind their dashboard using intuitive tools that make finding answers easy and everything happens with single click. Nowadays creating a dashboard is simple, thanks to hundreds of connections to popular business applications. (Ferrari and Russo 2016)

Many financial sector companies have very specific software requirements and, therefore, decide to build big data solutions from scratch. For this task, data engineers' skills are clearly vital. For those smaller enterprises that cannot afford hiring skilled data engineers, the solution is often found in fintech development outsourcing, which might stimulate demand for financial expertise within software companies (Lee and Shin, 2017).

Big data technology can also be used to reduce the cost of data storage. Capital markets firms often face difficulties with data aggregation for reporting to satisfy both comprehensive reporting needs. The process of collection, transformation, and aggregation of data can be a challenge because of resource constraints within the company. "The appeal of big data in this context is that firms are able to tier storage - older or less well-used data sets can be stored in cheaper data containers-and they can establish a long-term data retention facility that will not be challenged by increasing volumes of input over time" (O'Sheaa and Shah, 2014, p.29).

However, it's important to note that spending significant amount of money on big data analytics doesn't guarantee that companies are getting the results that they want - or even that they know what they are doing. Apart from increased variety, volume and velocity of the data there are number of pros and cons identified by different researches so far. According to NewVantage Partners' 2018 big data Maturity Survey Enterprises report multiple advantages as well as disadvantages of big data, including the following (https://www.datamation.com/big-data/big-data-pros-and-cons.html):

Table 6 Advantages and Disadvantages of big data

Advantages of Big Data	Disadvantages of Big Data
Better decision-making	Need for talent
Increased productivity	Data quality
Reduce costs	Need for cultural change
Improved customer service:	Compliance
Fraud detection	Cybersecurity risks
Increased revenue	Rapid change
Increased agility	Hardware needs
Greater innovation	Costs
Faster speed to market	Difficulty integrating legacy systems

List of the 2018 executive survey participants included 77.2% Financial Service companies, like e.g. Goldman Sachs, JP Morgan, Citi Group, Wells Fargo, Bloomberg and etc. One of the key finding of the survey is that 97.2% of organizations are investing in big data and AI initiatives. However, investment levels continue to be relatively lower and modest. It can be expected that this investment trend is only like to continue in order to compete with data-driven competitors. (NewVantage Partners LLC, 2018, p.8).

V. CONCLUSION

Business decisions may usually have a strong financial effect, especially in financial trading, where even single decision, regardless of its size, might result in huge loss or gains. Over the past decades, advances in computer technologies in combination with the evolution of market structure resulted in emergence of electronic securities trading, leading to dramatic change of global securities markets. Financial securities transactions that used to be conducted by person and over the counter using phone are now basically executed by automated trading platforms. As a result the world is facing a vast increase in the exchange volumes, significant fragmentation of the markets and even bigger increase in the number of orders.

Financial trading markets have never been lacking large data sets. The emergence of big data and relevant advanced business intelligence will make them even more heavily dependent on data analysis and should change the way in which these markets' participants operate.

Big data has become a big deal and capital market companies have been transforming to handle it. It becomes apparent that business intelligence applications are increasingly backing trading process and dealing with large amounts of data. This start to happen for large trading organizations and investment banks, but also noticeably for relatively smaller business entities and even individual traders. However it is revealed big data in capital markets is still far from being considered mature, because capital markets have been relatively slow to adopt relevant strategies and latest developments, namely data science and high-frequency trading are still relatively new in the financial trading. Companies who want to integrate to global markets, among other things, will have to think on transforming its operations and automating big data analysis process.

Implementation of the technology dealing with variety, volume and velocity of the data comes with its own cost. In spite of this, apparently capital market companies are increasingly investing in these technologies with two main purposes: to stay competitive and to ensure investors earn optimal return on investments. Companies are trying to catch up and explore how to use these technologies not just in the regulatory space, but also to create value through improved modelling and analysis for trading purposes. Companies' performance measurement techniques and risk management will be increasingly reliant on scrutinizing historical data and combining it with real - time data to apply portfolio weights, model liquidity, and performance attribution over a determined period.

This is a growth area with huge potential, as the amount of digital information (particularly unstructured) is growing at staggering rates - more data was created during the last two decades than in the rest of human history. Investing in big data technologies will help companies to stay up-to-date, reduce costs, dig into trends, go deep into figures, investigate changes, but they should not forget Simpson's Reversal Paradox and the risk of losing global picture – data might be misleading.

VI. APPENDIX A

Simpson's paradox can be found in economics, while analyzing GDP and per capita GDP growths. "Aggregation is a common practice in economics and business, and it often causes counterintuitive "anomalies," which are simply manifestations of Simpson's paradox or the Yule–Simpson's effect. Aggregated data below shows a decline in the per capita GDP of the world from the mid-1990s into the turn of the twenty-first century. Since the per capita GDP of the world decreased by nearly 1% over the period, one may be bewildered by the apparent decreased standard of living in the world and wonders why all the technological advances in the late 1990s, which presumably lifted productivity, regressed the society as a whole. When the same data are partitioned into developing and developed countries, however, the per capita GDP actually showed an increase for both the developed and developing countries over the period. While the per capita GDP of the developed countries increased 1.72% from 1996 to 2001, the per capita GDP of the developing countries increased 4.32%. How could both the developed countries and developing countries increase the per capita GDPs while the world's per capita GDP decreased (Ma, 2015, pp.5-6)"?

Table 5 Comparing per capita GDPs of the developed and developing countries in 1996 and 2001

	1996		2001		Change	
	Per capita GDP (USD)	Population (millions)	Per capita GDP (USD)	Population (millions)	Per capita GDP (%)	Population (%)
World	5,357	5,780	5,305	6,180	−0.97	6.92
Developed countries	21,823	1,185	22,199	1,218	1.72	2.78
Developing countries	1,110	4,595	1,158	4,962	4.32	7.99

Source: World Bank (2014). The developed countries are labeled as high-income countrIes, while developing countries include low- and middle incomes countries in World Bank. The per capita GDPs are in 2013 US dollar (USD). Populations are rounded in million

"In the example shown in Table 5, the populations of the developed and developing countries changed unevenly from 1996 to 2001. Whereas the population of the developed countries increased by 2.78%, the population of the developing countries increased, at a much higher rate, 7.99%. This uneven change of population size in the compared entities is a main cause for the reversal - both the developed and developing countries had an increase in the per capita GDP, yet the world as a whole actually had a decrease in per capita GDP from 1996 to 2001" (Ma, 2015, p.5).

VII. ACKNOWLEDGMENT

This work was supported by Shota Rustaveli National Science Foundation (SRNSF) [Grant number: PHDF 18-492; Project Title: "International Financial Markets and Capital Market Development in Georgia"]

VIII. REFERENCES

1. Alfaro, L., Chanda, A., Kalemli-Ozcan, S., Sayek, S. (2004) FDI and Economic Growth: The Role of Local Financial Markets". Journal of International Economics, volume 64, pp. 89-112.
2. Pagano, M. (1993) Financial Markets and the Macroeconomy. Bocconi University, Milan, Italy and CEPR, London, UK. European Economic Review 37 (1993) 613-622. North-Holland.
3. Prati, A., Schinasi, G. (1997) European Monetary Union and International Capital Markets: Structural Implications and Risks, International Monetary Fund Research Department.
4. Dicle, M. F., and Levendis, J. 2013 The impact of technological improvements on developing financial markets: The case of the Johannesburg Stock Exchange. Review of Development Finance 3 (2013) 204–213. Available online at www.sciencedirect.com. Last accessed on 30-Sep-2019.
5. Morosan, G. (2015) Integration of Romanian Insurances Market in EU, ECOFORUM, Volume 4, Special Issue 1.
6. Levinne, R., Zervos S. (1996), Stock Markets, Banks and Economic Growth. The World Bank, Policy Research Department, Finance and Private Sector Development Division.
7. Choudhry, M., Joannas, D., Landuyt, G., Pereira, R. and Pienaar, R. (2010) Capital Market Instruments: Analysis and Valuation, 3rd edn., Macmillan Publishers Limited.
8. Guerard, J. B., Rachev, S. T., Shao, B. P. (2013) Efficient global portfolios: Big data and investment universes, Journal of Research and Development, Volume 57, no. 5, Paper 11.
9. Jothimani, D., Shankar, R., Yadav S. (2014) A Big Data Analytical Framework For Portfolio Optimization, City University of Hong Kong, International Conference on Frontiers of Finance.
10. Mauro A., Greco, M., Grimaldi, M. (2016) A Formal Definition of Big Data Based on its Essential Features, Library Review, Vol. 65 Iss: 3, pp.122 – 135
11. Rotună, C., Cîrnu E.C., Smada, D., Gheorghiță, A. (2017) Smart city applications built on big data technologies and secure IOT, ECOFORUM, Volume 6, Issue 3(13).
12. Martin-Sanchez, F., Verspoor K. (2014) Big Data in Medicine Is Driving Big Changes, Health and Biomedical Informatics Centre, The University of Melbourne, Parkville VIC 3010 Australia.
13. Ye, M., Li, G. (2017) Internet big data and capital markets: a literature review, Financial Innovation. 3. 6. 10.1186/s40854-017-0056-y.
14. Sivarajah, U., Kamal, M.M., Irani, Z., Weerakkody, V. (2017) Critical analysis of Big Data challenges and analytical methods, Journal of Business Research, Volume 70.
15. O'Shea, V., Shah, R. (2014), Big Data in Capital Markets: At the Start of the Journey. Prepared for Thompson Reuters.
16. Bank of International Settlements (2016) Electronic trading in fixed income market, available on the BIS website (www.bis.org) ISBN 978-92-9197-420-7 (online). Last accessed on 30-Sep-2019.
17. Hammer, C.L., Kostroch D. C., Quirós G., STA Internal Group. (2017), Big Data: Potential, Challenges, and Statistical Implications, IMF Discussion Note.
18. Dejaeger K., Hamers, B., Poelmans, J., Baesens, B. (2010), A Novel Approach to the Evaluation and Improvement of Data Quality in the Financial Sector, Department of Decision Sciences and Information Management, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium
19. Wang, Y., Wiebe, V. J. (2014) Big Data Analytics on the characteristic equilibrium of collective opinions in social networks, International Journal of Cognitive Informatics and Natural Intelligence (IJCINI), 8(3), 29–44.

20. Chen, H., Chiang, R. H., Storey, V. C. (2012). Business intelligence and analytics: From Big Data to big impact. MIS Quarterly, 36(4), 1165–1188.
21. Simonet, A., Fedak, G., Ripeanu, M. (2015). Active Data: A programming model to manage data life cycle across heterogeneous systems and infrastructures, Future Generation Computer Systems, 53, 25–42.
22. Huisman, D. O., (2017), To What Extent do Predictive, Descriptive and Prescriptive Supply Chain Analytics Affect Organizational Performance? University of Twente.
23. Guidolin, M. (2011), Markov Switching Models in Empirical Finance, Bocconi University, Milan, Italy.
24. Kavitha, S., Vadhana R., P., Nivi, A. (2015) Big Data Analytics in Financial Market. International Journal of Research in Engineering and Technology, Volume 04, Issue.
25. Pearl, J. (2011), Simpson's paradox: An anatomy, https://escholarship.org/uc/item/3s62r0d6.
26. Bassett, W. (1997) Simpson’s Paradox and Investment Management.
27. Ma, Z. (2015) Simpson’s paradox in GDP and per capita GDP growths, Journal of Empirical Economics, Article No.: 0921.
28. Graham, M., Shelton, T. (2013), Geography and the future of big data, big data and the future of geography, Dialogues in Human Geography, Dialogues in Human Geography, Volume: 3, Issue: 3, page(s): 255-261.
29. Kandel, S., Paepcke, A., Hellerstein, J.M., Heer, J. (2012) Enterprise Data Analysis and Visualization: An Interview Study, http://vis.stanford.edu/files/2012-EnterpriseAnalysisInterviews-VAST.pdf
30. Ferrari, A., Russo, M (2016), Introducing Microsoft Power BI, Microsoft Press.
31. Lee, I., Shin, Y.J. (2017) Fintech: Ecosystem, business models, investment decisions, and challenges, Business Horizons, 1424; No. of Pages 12.
32. https://www.datamation.com/big-data/big-data-pros-and-cons.html. Last accessed on 30-Sep-2019.
33. NewVantage Partners (2018) Big Data Executive Survey 2018.