

Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Kristýna Brunová
Advisor:	PhDr. Ladislav Křišťoufek, PhD.
Title of the thesis:	Are the more popular stocks also the risky ones? Google and Wikipedia searches in portfolio optimization

OVERALL ASSESSMENT (provided in English, Czech, or Slovak):

The thesis at hand investigates the connection between Google/Wikipedia frequency of searches for content associated with DJIA constituting companies and their riskiness, as assessed from the performance of an optimal portfolio. My comments are the following:

Chapter 1 provides an introduction into the investigated topic, while Chapter 2 provides a literature survey. I appreciate that the survey includes papers on the use of Google searches in other than portfolio optimization topics, as Google searches are a relatively new topic.

Chapter 3 describes the data used, the method used for data collection. I appreciate the detailed description of GSV and the meticulously described procedure of data selection. I have no comments here, except maybe a table with overall data statistics would be convenient.

Chapter 4 introduces the methodology used. I have several comments here:

- A minor typo - in the second paragraph of p. 20, the expression „negatively affects“ should probably be „positively affects“
- In performance evaluation, the portfolio is rebalanced every morning and the overall portfolio return for the given day is taken as a weighted average of separate stocks' daily returns on that given day. However, if the portfolio is rebalanced every morning, as search data were not yet available when the market closed on the previous day, the calculation of the portfolio performance is inaccurate as it overlooks the stocks' overnight return. Taking, e.g. the GE stock, it closes at 27.12 on April 28, opens at 26.95 and closes at 27.09 on April 29, hence in some cases the overnight return can be a significant component of the daily return.

Chapter 5 comments on the results. I have several comments here:

- All discussions of results mention that transactions costs were not taken into account. As the introduction of transaction costs into portfolio profitability should not be complicated, I would like to see this too as the work is presented for the benefit of real investors.
- Secondly, as the web-search frequency is assumed to be related to riskiness of each stock, I would like to see a comparison of web-search-based portfolio optimization as opposed to optimization based on volatility predictions (with the simplest prediction being average past volatility). This would show how much new data web-searches bring as opposed to traditional methods of stock riskiness evaluation.
- In section 5.2 I would like to see a more thorough conclusion on why more precise searches (with „stock“ included) lead to worse performing portfolio – more precise searches should lead to more precise data on the future riskiness.

Overall, this thesis is very nicely structured and written, the hypothesis is clear. Throughout the whole text the author proves a thorough understanding of the topic. Despite the text being rather short, this is a good thesis and thus deserves a good grade.

Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Kristýna Brunová
Advisor:	PhDr. Ladislav Křišťoufek, PhD.
Title of the thesis:	Are the more popular stocks also the risky ones? Google and Wikipedia searches in portfolio optimization

SUMMARY OF POINTS AWARDED (for details, see below):

CATEGORY	POINTS
<i>Literature</i> (max. 20 points)	20
<i>Methods</i> (max. 30 points)	20
<i>Contribution</i> (max. 30 points)	30
<i>Manuscript Form</i> (max. 20 points)	20
TOTAL POINTS (max. 100 points)	90
GRADE (1 – 2 – 3 – 4)	1

NAME OF THE REFEREE: Daniel Benčík

DATE OF EVALUATION: 24.05.2015



Referee Signature

EXPLANATION OF CATEGORIES AND SCALE:

LITERATURE REVIEW: *The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
20	10	0

METHODS: *The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
30	15	0

CONTRIBUTION: *The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
30	15	0

MANUSCRIPT FORM: *The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
20	10	0

Overall grading:

TOTAL POINTS	GRADE		
81 – 100	1	= excellent	= výborně
61 – 80	2	= good	= velmi dobře
41 – 60	3	= satisfactory	= dobře
0 – 40	4	= fail	= nedoporučuji k obhajobě