Report on **Bachelor / Master Thesis**

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

Student:	**Yu Zang**
Advisor:	Prof. Ing. Oldřich Dědek, CSc.
Title of the thesis:	Determinants of International Tourists Inflows: The case of China

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

This master thesis is aiming at estimating the factors determining the the incoming tourism to China based on statistics for 1998-2012. This is an interesting and important economic topic that was not covered extensively in the world research and it offeres also a wide field for exercising technical skills in econometrics. The author has decided to use gravity equation for that purpose – a model that received in the last 20 years a high attention because of its practical applicability.

Assessment of the literature used:

The choice of the literature deserves appreciation – the author found an extensive list of titles where all important papers were presented. However, some of the enlisted papers were not articles of high standard and a sort of critical examination would be essential. Expecting that seems to be a too demanding taskfor an unexperienced student researcher who very often trusted the authors who did not deserve it.

In describing the methodology I missed the critical literature on the gravity models, e.g. the papers by Baldwin, R. and Taglioni D. 2006. Gravity for Dummies and Dummies for Gravity Equations. Cambridge MA, NBER, Working Paper no. 12516 or Egger, P. 2002. An Econometric View on the Estimation of Gravity Models and the Calculation of Trade Potentials. The World Economy 25(2): 297–312.

Assessment of the manuscript form:

I found too many errors, omissions and sloppy approaches to statements. The latter can be illustrated on the negligence with the data/numbers/:

E.g. „The inbound tourists reached an amazing 129 million people and brought an income of 51.7 U.S dollars“ (p. 13). However, on p. 14 there is „The entry of foreign visitors throughout the year are 27,191,600 passengers“ (data for 2013 and 2012).

At the same time table 8.3 indicates mere 2719 visitors from abroad measured in 10000 (sic!), i.e. once again 27 million. In contrast to that table 3.1 (p. 14) enlists 20 touristically most important countries (the model works with 22) that comprise mere 2362 thousand tourists (i.e. mere 8.7% of the reported passangers). This is a highly mysterious statistics and a weird data selection (could it be that 91.3% of tourists were not a part of the estimation?).

Another case of inconsistency: „The panel data includes … 15 years (1998-2002)“ (p. 21).

Some sentences are not logically consistent. E.g. „In this study it is long panel. In a balanced panel, all entities have measurements in all time periods. In this study it will be balanced panel data“ (p.21 in the middle).

Report on **Bachelor / Master Thesis**

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

Student:	**Yu Zang**
Advisor:	Prof. Ing. Oldřich Dědek, CSc.
Title of the thesis:	Determinants of International Tourists Inflows: The case of China

There are too many times when the time dimension of the data is indicated as 1998-2002 when in reality the data are up to 2012.

There is a highly insufficient and non-standard style both in quoting references (e.g. missing year or quoting the given name) and, journal and in the list of references (e.g. missing the journal name, its number and pages, or city and institution publishing the source). Publications are sorted by given names, not by surnames.

The text would deserve editing by some native English speaker since there are too many mistakes in grammar.

Assessing the own contribution of the author:

Here I will distinguish between the ideas behind the specification of the model and the choice of data where both are supposed to be related to some theoretical framework. Econometric contribution will be assess later in the section on „Methods".

Unfortunately, the theories of economic gravity are delt with too quickly (pp. 36-40) and too late, when the data has been already chosen (pp. 21-35). Theories had apparently minimal impact on the specification of the model.

I have the following provisos to the construction of the model:

I cannot share the idea that the intensity of tourism is determined exclusively by the conditions in the visitors' countries (p. 5). E.g. the visa conditions are of high importance (see e.g. SWAGEL, P. 2009. The impact of post-9/11 visa policies on travel to the United States. *Journal of International Economics*. June 2009, Vol. 78, No. 1, pp. 86-99). Then also domestic political liberalisation, real exchange rate (inflation), status of hotels, transport infrastructure and services (e.g. proxied by home GDP) could be included, which the author did not consider...

Mixing the tourism for leisure with the tourism for business is mixing two subpopulations of data whose motive (and behavioural patters) can be expected to be very different. Explaining them in the same equation (or without subjecting them to a dummy) can be a source of serious bias and a loss in efficiency of estimation.

Another source of bias can be caused by choosing the tourist arrivals (total annual) to China as the dependent (explained) variable. Arrival of a "tourist" coming for 1-day supermarket shopping from Russia, Vietnam or Myanmar cannot be compared with a tourist coming for a month from the USA. Explaining the number of visitor-nights or the receipts from tourism could serve as a more meaningful variable.

The measure of distances in some cases can be misleading, e.g. 5795 km between Russia and China does not represent the real proximity at cross-border visits between Manchuria and the Amur area.

Report on **Bachelor / Master Thesis**

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

Student:	**Yu Zang**
Advisor:	Prof. Ing. Oldřich Dědek, CSc.
Title of the thesis:	Determinants of International Tourists Inflows: The case of China

The use of nominal exchange rates (relative to USD) could have some meaning in time series[1], even though it is the real exchange rates that would be more appropriate. However, that very nominal data used in cross-section estimates becomes completely meaningless. Another confusion comes when the PPP conversion factor is used in parallel with the previous. Why, instead, is the relationship between them (i.e. some sort of Exchange Rate Deviation Index) not used for assessing the evolution of purchasing parity?

Rodrigue's modification of the model by introducing TD (p. 37) in order "to suit the tourism" would deserve much more concrete explanation.

Replacing the GDP of partner countries with their GDP per capita is a serious interference with the theoretical definition of the gravity equation and its justification would deserve more explanation in this case than merely remarking pragmatically: "As some authors have already proved, that if GDP per capita used instead of GDP, all other variables will have significant effect on tourist inflows" (p. 39).[2]

However, the abandoned GDP comes back again when the WD variable is introduced. The economic meaning of the weighted distance (using shares on world GDP as a weight) should be explained in much more detail. This is a highly non-standard approach to the distance in the gravity models that distorts the whole model. It cannot be introduced by claiming that "several literatures suggested [these] approaches."[3] What about two countries with similar geographic distance, a similar GDP per capita but a highly different shares on the world's GDP? (as is the case of US and Canada)? Is then Canada so much less distant and can export more tourists to China due to that?

Assessing the methods applied:

Considering the just mentioned WD, it is no surprise that the reader gets confused by its intuitively unexpected results (e.g. receiving results with the negative and the positive signs of the distance, depending on the technique of estimation used).

This is not a unique case. As we compare alternative estimations we can see that estimated coefficients are highly non-robust, differing in significance (from significance at 1% level to complete insignificance), in signs and in the magnitude of elasticities. This relates to all explanatory variables.

The mix of data that was selected without considering some consistent theoretical reasoning can lead to spurious regressions, omitted variable distortions and

[1] I mean "some meaning" provided the same currency is still in use, which was not the case of the Eurozone countries (e.g. the data for Italy is 1736.207 for 1998 and 0.938627 for 1999) and if the relative inflation is not internalized by the exchange rate adjustments.
[2] Errors in grammar are left unchanged.
[3] Not all published literature is reliable and some can be outright flawed!

Report on **Bachelor / Master Thesis**

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

Student:	**Yu Zang**
Advisor:	Prof. Ing. Oldřich Dědek, CSc.
Title of the thesis:	Determinants of International Tourists Inflows: The case of China

problems with endogeneity and co-integration. It does not matter that the author applied the techniques of estimation and econometric controls quite skillfully (and mechanically). After checking the results I am not sure if the exercises produced in the thesis have proven anything credible.

I would recommend that the model be estimated as cross-section one-way panel where variables with exchange rate are replaced by their quotient (reflecting the purchasing power) and where the GDP and geographic distances are re-introduced. Plus some institutional variables added describing the progressing liberation and the cultural rapprochement of China to diferent partner countries.

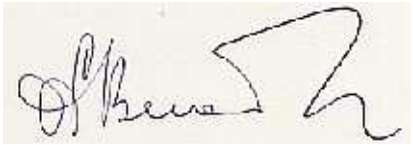
Final conclusion:

Considering the problems with model specification and estimation, I recommend the thesis to be defended, but just on a margin. My proposed grade is weak 3.

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

CATEGORY		**POINTS**
Literature	*(max. 20 points)*	17
Methods	*(max. 30 points)*	12
Contribution	*(max. 30 points)*	6
Manuscript Form	*(max. 20 points)*	6
TOTAL POINTS	*(max. 100 points)*	**41**
GRADE	**(1 – 2 – 3 – 4)**	**3**

NAME OF THE REFEREE: Doc. Ing. Vladimír Benáček, CSc.



DATE OF EVALUATION: 1 September 2014

_____

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.*

Strong	*Average*	*Weak*
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.*

Strong	*Average*	*Weak*
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.*

Strong	*Average*	*Weak*
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.*

Strong	*Average*	*Weak*
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