Factors affecting long-term economic growth-consistency and stability by soft regression estimation

Eli Shnaider*
Peres Academic Center,
10 Shimon Peres St.,
Rehovot, Israel
Email: eli-sh@012.net.il
*Corresponding author

Arthur Yosef
Tel Aviv-Yaffo Academic College,
2 Rabenu Yeruham St.,
Tel Aviv-Yaffo, Israel
Email: artyosef@gmail.com

Nava Haruvy
Netanya Academic College,
1 University St.,
Netanya, Israel
Email: navaharu@netvision.net.il

Abstract: This study demonstrates a challenge of building and validating a model of factors associated with long-term economic success of economies, as reflected by measures of their aggregate value of output/income per capita. We use two quantitative modelling techniques including multiple linear regression (MLR) and a newer technique of soft regression, a modelling tool based on fuzzy information processing technology. The objective of this study is to test and compare the two information-processing tools to find the more reliable and potentially helpful tool for policy decision making. The conclusions of this study are: 1) the soft regression tool generated more consistent and comprehensible results in comparison to the MLR method; 2) based on soft regression tool, the model displayed solid stability over extensive period under study; 3) based on MLR method, we could interpret the results as supporting the validity of the model; however, some of the results were contradictory, thus undermining the reliability of conclusions.

Keywords: multiple linear regression; MLR; soft regression; fuzzy information processing; economic growth; international competitiveness.

Factors affecting long-term economic growth-consistency

Biographical notes: Eli Shnaider is an Associate Professor of Economics in the Peres College in Israel. He received his PhD in Economics, and MSC in Computer and Information Sciences and in Economics from the Florida State University. He has over 30 years of research experience in the fields of quantitative economics, economic modelling and forecasting, and computational intelligence as focusing on applications in economics and management, international economics and international competition.

Arthur Yosef is a Senior Lecturer in Tel Aviv-Yaffo Academic College in Israel. He received his PhD in Mathematics Bar Ilan University in Israel, and has research experience on mathematical, statistical and information systems modelling and applications.

Nava Haruvy is a Full Professor of Economics in the Netanya Academic College in Israel. She received her PhD in Economics from the Hebrew University in Jerusalem, and has over 30 years of research experience in economics of the environment and natural resources, with a focus on water and wastewater economics. She has published over 40 papers in refereed journals, as well as many conference symposia proceedings and other papers. She has participated in international research projects on the development and environmental issues.

1 Introduction

The objective of this study is to examine the effectiveness of two information-processing tools (modelling tools) under conditions of severe uncertainty – the multiple linear regression (MLR) and the soft regression (SR) method. The idea is to utilise these tools to build a computerised model that will constitute effective policy supporting tool. The purpose of this paper is to compare and evaluate the two methods, while using the construction of economic growth model as a case study.

First stage in the modelling process is the theoretical definition of the model. We utilise several factors that are usually associated with the long-term economic success as reflected by various measures of standard of living. These factors include international competitiveness, human capital and degree of social progress. Absence of quantitative cross-national data for these variables induced us to use proxy variables. Building a quantitative model consisting of known concepts, but without directly measurable data is a challenging task and a good case study to demonstrate the degree of effectiveness and reliability of the two very different information-processing methods.

We decided to test the model over four decades by using cross-national data for the years 1965, 1978, 1992 and 2007 to confirm its long-term stability. Traditional modelling techniques such as MLR require precise and complete model specification, as well as additional restrictive and inflexible conditions, which is highly problematic under the conditions as described above. Hence, we also applied the SR method, based on fuzzy information processing, which does not require precise model specification, and can handle incomplete models without distorting the relative importance of participating variables. We compared the performance of the two information processing techniques regarding reliability and stability.
2 Literature

The literature associated with economic growth includes emphasis on investment in physical capital and technological progress (Thirlwall, 1987), including the ‘new growth theory’ (Koner, 1996). The background factors facilitating economic success include international competitiveness, human capital and the degree of social progress (Shnaider and Haruvy, 2008).

Greater degree of international competitiveness allows the country to capture a larger portion of the continuously expanding global markets and to facilitate long-term economic growth (Davidson and Kregel, 1997).

The concept of international competitiveness of a country (relative to other countries) is frequently used but rarely defined (Fagerberg, 1988). However, it refers to the ability of a country to realise economic policy goals, while there are links between the economic growth and balance of payment position of an open economy, which we refer to size of export. According to Fagerberg (1996), the words ‘technology’ and ‘competitiveness’ are two of the most popular buzzwords of our time, and policy-makers link the two. Lundvall (2010) presents a new perspective on the dynamics of the national and the global economy by basing international competitiveness of nations on innovation referred to level of technological development.

Investment in human capital constitutes one of the main factors promoting long-term economic growth (Seidi, 2001). Cohen and Soto (2007) presented data-set regarding years of school across countries for the years 1960–2000 (based on the OECD database on educational attainment and from surveys published by UNESCO). They found in standard cross-country growth regressions that their series showed significant coefficients for schooling describing impact of human capital on growth. Routledge and von Amsberg (2003) defined and characterised social capital in a simple growth model by incorporating the trading model into a growth model to connect among growth, labour mobility, and social capital.

According to Fagerberg (1996), the words ‘technology’ and ‘competitiveness’ are two of the most popular buzzwords of our time, and policy-makers link the two. He addressed the questions of what do we really mean when we talk about the international competitiveness of a country, and what does technology have to do with this from a long run perspective.

Hence, the most important factors affecting economic growth include international competitiveness, quality of human capital, and degree of social progress. The proxies we use include export per capita-proxy for the degree of international competitiveness, tertiary and secondary education-proxy for investment in human capital as well as the degree of social progress. Additional proxies include: high technology (exports of products associated with advanced technology)-represent both knowledge, as related to human capital, and international competitiveness, as reflected by the export sector of the economy, and natural population growth-representing the degree of social progress.

Regarding estimation methods, MLR is the very basic, widely used tool as found in Seber and Lee (2003), Brown (1990), Diaz et al. (1983) and Gilioni and Padberg (2002). Various soft computing/fuzzy information processing techniques have been introduced to handle severe cases of uncertainty, where standard statistical systems cannot function effectively due to challenges such as lack of knowledge or lack of data to define and
implement complete and precise model. The idea of fuzzy sets was introduced by Zadeh (1965), and since then has been widely applied in various branches of information processing, such as fuzzy linear programming (Saati et al., 2015), fuzzy genetic algorithms (Bera and Maiti, 2012), various fuzzy neural networks applications – such as fuzzy cognitive maps (Kosko, 1992), etc. In this article we apply soft computing tool – SR which is based on fuzzy information processing (Kandel et al., 2000; Maimon and Rokach, 2005). Comparison of MLR to SR appears in Yosef et al. (2015).

3 Model

In this section, we introduce a conceptual model, which is the precondition for constructing a quantitative model. We identify the following factors that influence economic success of various countries (reflected by standard of living) as components of our model. These factors are international competitiveness, human capital and social progress. As we stated above, there are no quantitative data for these factors, and it is necessary to use proxy variables instead. This constitutes an excellent example of practical case where policy makers are induced to make policy decisions under severe conditions of uncertainty: the relevant variables are not directly measurable, and the use of proxies is the only option for applying quantitative computing.

Therefore, it is critical to define and explain in details the components of the model in order to be able to evaluate to what extent the use of quantitative proxy variables reflects the originally intended variables of the model (will be explained in the next section):

1. *International competitiveness*: The more globally competitive is the country and the higher is the value of products and services sold in global markets (per capita), the greater is economic success and higher standard of living. The term ‘international competitiveness’ reflects:
   a. The ability of a given country to produce products and services in a competitive manner (high quality or low cost).
   b. The capability to supply globally large enough amount of products and services while generating substantial income per capita.

2. *Quality of human capital*: Human capital includes factors such as education, knowledge, skills, experience, and tradition. It is reflected by phenomena such as development of new technologies and products, research and development capabilities, high technology infrastructure, education and research facilities and infrastructure, and organisational and management skills.

   Human capital is an important factor in determining international competitiveness of the economy. Countries possessing substantial human capital capabilities are able to:
   a. Develop new products and thus enjoy at least temporary world monopoly (until imitations are developed).
   b. Develop new methods of production and organisation.
   c. Rapidly enter new high technology intensive markets initiated by their rivals (once it becomes apparent that such new markets hold substantial economic promise).
d Enable quick adaptation to technological developments by absorption of new technologies, techniques and methods (that have appeared successful in use by others).

e Enable improved efficiency and thus higher income generation in domestic non-tradable sector of the economy (including public sector).

3 Degree of social progress: We characterise socially advanced countries by:

a Degree of social sophistication and flexibility required for effective functioning of modern and internationally competitive economy.

b Social environment facilitating growth and retention of human capital.

c Higher degree of personal and economic freedom.

d Greater adaptability to ever-changing global economic-political environment.

e Greater flexibility to find a reasonable compromise between local traditions and the requirements necessary for functioning of successful modern economy.

There is a definite relation expected between the degree of social progress and the previously defined factor of human capital. In addition, we expect some relation between human capital (technology, knowhow) and international competitiveness. Hence, the factors included in this model are dependent of each other. This fact constitutes a severe limitation for information processing tools based upon assumption that all explanatory variables are independent (such as MLR).

4 Explanatory variables

We utilised the following variables as proxies for the three explanatory factors of our model: international competitiveness, human capital and social progress.

1 Export per capita-being a proxy for the degree of international competitiveness of a given economy in global markets (adjusted for the population size). This is not a very accurate proxy since in some countries re-export of imported components may constitute a substantial portion of their exports, while in other countries such re-exported component is small or insignificant. However, even in the case of the substantial re-export component, it still requires certain competitive capabilities that determine why one country is engaged in re-exporting of some items and not others.

2 Education-being a proxy for investment in human capital as well as a proxy of the degree of social progress:

a Tertiary education-percentage of the relevant population group that attends tertiary education institutions.

b Secondary education-percentage of the relevant population group that attends secondary education institutions.

3 High technology-actually refers to exports of products associated with advanced technologies. This variable represents a combination of the first two factors by having an important component representing both knowledge related to human capital and international competitiveness as reflected by the export sector of the economy.
Factors affecting long-term economic growth-consistency

4 Natural population growth—birth rate minus death rate. This is a proxy representing a degree of social progress. Rapid natural growth of population is in general associated with agrarian economies, where the agricultural sector is usually characterised by traditional (and technologically backward) methods of production, where large families are customary and where children since a very young age are utilised as labour force thus negatively affecting their investment in human capital. On the other hand, smaller families (that result in slower population growth) are usually associated with the aspiration to be part of the middle class (or above), to acquire education and skills needed for a successful career. Also, slower population growth requires less expenditures for infrastructure whose purpose is just to accommodate increase in numbers, and therefore more could be spent to improve quality (of human capital).

5 Method

In this study, we defined and tested a quantitative model of factors associated with economic growth, based on cross-national data obtained from the database, as well as previous files and publications of the World Bank. We tested the model over extended period of 40 years (1965, 1978, 1992 and 2007) to confirm the model’s validity and stability. We utilised two quantitative modelling tools—MLR and SR method.

When selecting data we avoided years with severe international economic crises (for example years 1973–1976, 1990–1991, and 2008–2013). We also chose similar intervals among the selected years. For 2014, the data were not available for large number of countries at the time we performed our study.

We utilised aggregate income per capita as a measurement of performance of various countries. However, there are several different methods to measure aggregate income/output on a cross-national basis. These methods are based on current or constant US dollars (USDs) relating to per capita gross national product (GNP), gross domestic products (GDP) or gross national income (GNI). All these measures are internationally accepted, and we could not figure out which one better represents ‘standard of living’ (another good example of uncertainty our tools had to overcome). In addition, not all the measurement methods appear in all the years selected in the same form. Hence, to maintain long-term consistency we utilised for each year two of these measures:


Regarding the explanatory variable of exports, there are also several measurement methods substantially differing from each other as far as numerical values. We could not
find any theoretical argument which measure is more appropriate for use in our study, so
to maintain long-term consistency we utilised for each year three of these measures:

- For 1965: Exports of goods and services, constant 2000 USD (World Bank, 2006
database); exports of goods and services, current USD (World Bank, 2006); and,
merchandise exports, current USD (World Bank, 2007).
- For 1978: Exports of goods and services, constant 1995 USD (World Bank, 2004);
exports of goods and services, current USD (World Bank, 2009); and, merchandise
exports, current USD (World Bank, 2004).
- For 1992: Exports of goods and services, constant 1995 USD (World Bank, 2009
database); exports of goods and services, current USD (World Bank, 2009); and,
merchandise exports, current USD (World Bank, 2004).
- For 2007: Exports of goods and services, constant 1995 USD (World Bank, 2014);
exports of goods and services, current USD (World Bank, 2014); and, merchandise
exports, current USD (World Bank, 2014).

6 Multiple linear regression

There are many correlation techniques (association rule analysis, neural networks,
decision tree, etc.), but the classic and most common of all is a MLR. MLR attempts to
model the relationship between one or more explanatory variables (independent
variables) and a response variable (dependent variable) by fitting a linear equation to
observed data. Every value of the explanatory variable is associated with a value of the
dependent variable. The linear regression model expressed mathematically as:

$$Y = a_0 + a_1X_1 + a_2X_2 + \ldots + a_nX_n + \epsilon$$

where $X_1, X_2, \ldots, X_n$ represent explanatory variables, $Y$ represent a response variable, $a_0, a_1, \ldots, a_n$ are regression coefficients, and $\epsilon$ is error term associated with observation.

MLR is a modelling tool, and as such, we distinguish between the important factors
correlated with the variable we model and the unimportant factors. Modelling by
definition is a process of simplification, where we attempt to simplify a complex reality
and try to understand it by focusing only on the most important factors, while leaving
unimportant factors out of the model, so that they will not obscure our ability to analyse
and understand the most important things. Therefore, by definition, modelling involves
a certain degree of imprecision, caused by the factors (supposedly unimportant) left out of
the model. Those factors that are left out of the model are in reality still interacting with
the dependent variable causing some variation in its behaviour that the explanatory
variables included in the model cannot explain. Those unexplained variations represent
the phenomenon of randomness. Randomness is supposed to cause minor deviations in
the behaviour of the depended variable versus its expected behaviour based on the
behaviour of explanatory variables. This of course expected to be the case if the factors
left out of the model are truly of minor importance and tend to cancel each other out over
a large enough number of measurements.

However, if for whatever reason one or more of the important factors influencing the
dependent variable is left out, and is causing deviation in the expected behaviour of the
Factors affecting long-term economic growth-consistency

dependent variable, this is already not a randomness error (normal and expected statistical imprecision) but a modelling error causing mistaken results of a large magnitude. It is termed ‘misspecification of model’ and leads to biased, distorted results. Regular statistical tests cannot detect misspecification of the model. In some cases, one can detect model misspecification because the coefficients of explanatory variables do not have logical signs (plus instead of minus or vice versa). However, in many cases misspecified models appear to be logical, signs of their coefficients appear to be correct and all the statistical tests look satisfactory.

In this case, we will discover the problem only when the model fails. There are two main factors leading to the model misspecification due to the incorrect set of explanatory variables:

1. We are not aware of some important factor(s) influencing the behaviour of dependent variable.

2. We are aware of some important factors, but have no idea (or no budget) how to measure it, and as long as we cannot measure a variable in a quantitative way, it cannot be included in a quantitative model.

In addition to the possibility of the model misspecification due to the incorrect set of the explanatory variables, there is an additional, even more difficult to handle source of the model misspecification: a requirement to utilise a correct functional form of the equation.

In general we apply linear function because of the convenience (assuming it is a close enough approximation of real behaviour), and not because we have definite theoretical proof that the function is linear. If we decide to use nonlinear specification, there is an infinite amount of possible versions, of which only one is correct. In economics (as well as other fields where laboratory tests are impossible), there is no way to find out theoretically which one of the infinite amount of possibilities is the correct specification.

An additional factor for the model misspecification arises from purely technical reasons, since it is assumed that explanatory variables are independent of each other; However, in reality very often explanatory variables are highly correlated mathematically among themselves (even if logically they are unquestionably separate factors). This often causes either one or both of the correlated explanatory variables to appear as statistically insignificant, and therefore redundant (even though based on common sense, they should definitely be a part of the model in order to have correct model specification).

The modelling case study utilised in this paper is a good example of practical difficulties to adhere to all the strict requirements of MLR. The modelling process raises many questions that are very difficult to answer positively. For example, do we know with certainty all the important factors that facilitate economic growth, are they measurable and appear in statistical publications and databases, etc.? Do we have any idea regarding the correct functional form of the equation?

To compare we utilise an additional quantitative tool SR.

7 Soft regression

SR is a modelling tool based on soft computing concepts such as fuzzy logic and fuzzy information processing (Zadeh, 1965). We will briefly describe several of the important
characteristics of the SR tool that are different from those of MLR. These characteristics are:

a. SR does not require precise model specification. This regression tool is based on fuzzy logic, which is designed in the first place to handle information under severe conditions of uncertainty and imprecision. The idea here is to give up on the possibility of building a precise model and satisfying ourselves with the opportunity to work with whatever data are available. We generate a partial/less-precise model that could still be very reliable in a general direction of its conclusions because it avoids the problem of misspecification bias. It allows using partial and unreliable available data to make reliable but not precise conclusions in comparison to misspecified MLR model based on these data, especially in the case of partial data (some potentially important variables excluded due to lack of data).

b. Adding or removing variables does not affect the relative importance of the variables. It does not change the significance and relative importance of the explanatory variables among themselves. This is in contrast to the behaviour of MLR, where addition or removal of an explanatory variable can change drastically the significance of other explanatory variables of the model. This characteristic of the SR adds an important feature of stability and confidence to the decision making.

c. We introduce the heuristically determined $\alpha$-cuts (for minimum and maximum values of membership function) and designed membership function to determine to what extent each value of the processed data is a member in predefined fuzzy set (Zadeh, 1965). This brings the membership function utilised in SR more in line with ‘human thinking’ and thus allowing the modeller to monitor the logic of the information processing throughout the analysis. This feature of the SR helps to handle the distortions due to outlying values in a user-based logical approach (in contrast to strictly mathematical method utilised in robust regression approach).

1. In the SR method we utilise the measure of similarity which indicates the degree to which explanatory variable behaves in a similar pattern (direct or inverse) in comparison to dependent variable. Therefore, the measure of similarity, $S_{Y,X_j}$, is an equivalent to the statistical measures of significance (t-tests or sig.). However, in addition to significant relation (similarity of $S_{Y,X_j} \geq 0.8$), there is an option of partial significance $0.7 < S_{Y,X_j} < 0.8$, so that as $S_{Y,X_j}$ is approaching closer to 0.7, it is closer to insignificance. In addition, of course, when the similarity measure is below 0.7, the explanatory variable is insignificant. The gradual transition from being fully significant to being fully insignificant adds additional stability to modelling process while utilising SR. The relative importance of partially significant variables approaches zero as the variable moves closer to the border of insignificance (0.7). The binary decision requirement of the MLR method often forces user into dilemma – what to do with respect to variable when sig. > 0.05, or other standard cut-off point of significance by a minor degree.
After computing similarity measures for all the explanatory variables, we calculate collective contribution of all the explanatory variables combined in explaining the behaviour of dependent variable. This measure is denoted $S_{i, j, k, \ldots, n}$. It reflects, to what degree all the explanatory variables combined – explain the behaviour of the dependent variable, and in this respect it is equivalent to R square adjusted.

The heuristic features of SR discussed above provide important tools for modelers to avoid distortions when the mathematical ‘number crunching’ results are in contradiction to common sense.

### 8 Data

We obtained the data for this study from World Bank database and previous World Bank hard-copy publications and files. In the case of variable ‘high technology’, we found variable by that name in the database of the World Bank, but this variable has data only for the later part of the period covered by our study. For the early part of the period, we used variable ‘machinery and transportation equipment exports’ which is essentially the same type of variable because during the period of 1960s–1980s the capabilities to produce machinery and transportation products represented good proxy of advanced technology capabilities of various countries at that period. Thus for our study covering period 1965–2007 we combined those two variables into a single variable named ‘high technology’ in order to cover the whole period.

For years 1992 and 2007, we excluded new countries of the former USSR and Yugoslavia. The reasons were:

1. Those countries anyways did not exist during early years covered by our study (1965 and 1978).
2. During the early years of their independence, these new countries experienced severe economic downturn, that lasted for many years and their performance was way below their potential.

The main reason for the severity of the downturn was the difficult transition from centrally managed and highly integrated economic entities to separate economies under transition into more market-oriented approach (at least to some degree).

### 9 Results

Table 1 presents the comparative results of MLR vs. SR for dependent variable GNI per capita. The table displays results of regression runs (both MLR and the SR) with GNI per capita being a dependent variable. The table displays results for all years covered by this study – 1965, 1978, 1992 and 2007. We present three different specifications for various measurements methods of exports per capita (as we discussed above).
Table 1  Dependent variable GNI per capita

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High-tech</td>
<td>0.002</td>
<td>0.000</td>
<td>0.595</td>
<td>0.006</td>
</tr>
<tr>
<td>Export 1</td>
<td>0.863</td>
<td>0.847</td>
<td>0.880</td>
<td>0.888</td>
</tr>
<tr>
<td>Secondary</td>
<td>0.010</td>
<td>0.014</td>
<td>0.240</td>
<td>0.006</td>
</tr>
<tr>
<td>School</td>
<td>0.848</td>
<td>0.848</td>
<td>0.839</td>
<td>0.847</td>
</tr>
<tr>
<td>Tertiary</td>
<td>0.050</td>
<td>0.050</td>
<td>0.581</td>
<td>0.722</td>
</tr>
<tr>
<td>School</td>
<td>0.867</td>
<td>0.867</td>
<td>0.827</td>
<td>0.827</td>
</tr>
<tr>
<td>Natural</td>
<td>0.009</td>
<td>0.000</td>
<td>0.801</td>
<td>0.046</td>
</tr>
<tr>
<td>Population</td>
<td>0.860</td>
<td>0.950</td>
<td>0.761</td>
<td>0.257</td>
</tr>
<tr>
<td>Growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{R}^2$/S\text{comb}</td>
<td>0.867</td>
<td>0.958</td>
<td>0.801</td>
<td>0.761</td>
</tr>
<tr>
<td>High-tech</td>
<td>0.011</td>
<td>0.014</td>
<td>0.050</td>
<td>0.006</td>
</tr>
<tr>
<td>Export 2</td>
<td>0.863</td>
<td>0.848</td>
<td>0.867</td>
<td>0.888</td>
</tr>
<tr>
<td>School</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Tertiary</td>
<td>0.595</td>
<td>0.210</td>
<td>0.002</td>
<td>0.722</td>
</tr>
<tr>
<td>School</td>
<td>0.880</td>
<td>0.839</td>
<td>0.771</td>
<td>0.827</td>
</tr>
<tr>
<td>Natural</td>
<td>0.046</td>
<td>0.057</td>
<td>0.878</td>
<td>0.248</td>
</tr>
<tr>
<td>Population</td>
<td>0.822</td>
<td>0.809</td>
<td>0.827</td>
<td>0.781</td>
</tr>
<tr>
<td>Growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{R}^2$/S\text{comb}</td>
<td>0.867</td>
<td>0.958</td>
<td>0.761</td>
<td>0.953</td>
</tr>
<tr>
<td>High-tech</td>
<td>0.002</td>
<td>0.010</td>
<td>0.040</td>
<td>0.001</td>
</tr>
<tr>
<td>Export 3</td>
<td>0.863</td>
<td>0.848</td>
<td>0.867</td>
<td>0.921</td>
</tr>
<tr>
<td>School</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.771</td>
</tr>
<tr>
<td>Tertiary</td>
<td>0.590</td>
<td>0.120</td>
<td>0.000</td>
<td>0.832</td>
</tr>
<tr>
<td>School</td>
<td>0.880</td>
<td>0.839</td>
<td>0.000</td>
<td>0.827</td>
</tr>
<tr>
<td>Natural</td>
<td>0.041</td>
<td>0.071</td>
<td>0.878</td>
<td>0.247</td>
</tr>
<tr>
<td>Population</td>
<td>0.822</td>
<td>0.809</td>
<td>0.827</td>
<td>0.781</td>
</tr>
<tr>
<td>Growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{R}^2$/S\text{comb}</td>
<td>0.867</td>
<td>0.966</td>
<td>0.844</td>
<td>0.844</td>
</tr>
</tbody>
</table>

Notes: Export 1: exports of goods and services, constant USD; Export 2: exports of goods and services, current USD; Export 3: merchandise exports, current USD; $\text{R}^2$: adjusted R squared.

Table 2 presents the comparative results of MLR vs. SR for dependent variable GDP per capita. At first glance at Table 2, the immediate impression is that in general the results are similar to the results of Table 1. This by itself is an important issue: it points towards the robustness of the study. Despite the fact that we utilise different measurements as dependent variables, both generate very similar results.

Both, Tables 1 and 2 utilise model specifications consisting of all five explanatory variables. In MLR regression runs, there were usually two (in some cases one or three) insignificant explanatory variables in all regression runs in both tables.
In the SR regression runs, all explanatory variables were significant (even though there were few cases where some variables came out only partially significant). The results appear in Tables 1 and 2, and they are very similar although based on two different methods of measuring levels of aggregate economic activities (GNI and GDP). This stability of the results increases confidence in the conclusions and adds to the robustness of the study.

Table 2  Dependent variable GDP per capita

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High-tech</td>
<td>0.006</td>
<td>0.856</td>
<td>0.003</td>
<td>0.867</td>
<td>0.048</td>
<td>0.863</td>
<td>0.012</td>
<td>0.858</td>
</tr>
<tr>
<td>Export 1</td>
<td>0.000</td>
<td>0.849</td>
<td>0.002</td>
<td>0.893</td>
<td>0.015</td>
<td>0.905</td>
<td>0.000</td>
<td>0.932</td>
</tr>
<tr>
<td>Secondary school</td>
<td>0.813</td>
<td>0.906</td>
<td>0.168</td>
<td>0.847</td>
<td>0.194</td>
<td>0.764</td>
<td>0.914</td>
<td>0.717</td>
</tr>
<tr>
<td>Tertiary school</td>
<td>0.017</td>
<td>0.872</td>
<td>0.001</td>
<td>0.866</td>
<td>0.043</td>
<td>0.835</td>
<td>0.001</td>
<td>0.804</td>
</tr>
<tr>
<td>Natural population growth</td>
<td>0.222</td>
<td>0.854</td>
<td>0.244</td>
<td>0.818</td>
<td>0.601</td>
<td>0.821</td>
<td>0.209</td>
<td>0.779</td>
</tr>
<tr>
<td>( R^2/S_{comb} )</td>
<td>0.866</td>
<td>0.966</td>
<td>0.846</td>
<td>0.972</td>
<td>0.830</td>
<td>0.965</td>
<td>0.817</td>
<td>0.956</td>
</tr>
<tr>
<td>High-tech</td>
<td>0.026</td>
<td>0.856</td>
<td>0.003</td>
<td>0.867</td>
<td>0.044</td>
<td>0.863</td>
<td>0.013</td>
<td>0.858</td>
</tr>
<tr>
<td>Export 2</td>
<td>0.000</td>
<td>0.877</td>
<td>0.001</td>
<td>0.885</td>
<td>0.001</td>
<td>0.914</td>
<td>0.001</td>
<td>0.930</td>
</tr>
<tr>
<td>Secondary school</td>
<td>0.418</td>
<td>0.906</td>
<td>0.168</td>
<td>0.847</td>
<td>0.002</td>
<td>0.764</td>
<td>0.204</td>
<td>0.717</td>
</tr>
<tr>
<td>Tertiary school</td>
<td>0.021</td>
<td>0.872</td>
<td>0.040</td>
<td>0.866</td>
<td>0.697</td>
<td>0.835</td>
<td>0.002</td>
<td>0.804</td>
</tr>
<tr>
<td>Natural population growth</td>
<td>0.201</td>
<td>0.854</td>
<td>0.180</td>
<td>0.818</td>
<td>0.748</td>
<td>0.821</td>
<td>0.820</td>
<td>0.779</td>
</tr>
<tr>
<td>( R^2/S_{comb} )</td>
<td>0.863</td>
<td>0.967</td>
<td>0.861</td>
<td>0.952</td>
<td>0.830</td>
<td>0.966</td>
<td>0.806</td>
<td>0.955</td>
</tr>
<tr>
<td>High-tech</td>
<td>0.006</td>
<td>0.856</td>
<td>0.003</td>
<td>0.867</td>
<td>0.059</td>
<td>0.863</td>
<td>0.009</td>
<td>0.958</td>
</tr>
<tr>
<td>Export 3</td>
<td>0.000</td>
<td>0.838</td>
<td>0.002</td>
<td>0.879</td>
<td>0.009</td>
<td>0.920</td>
<td>0.000</td>
<td>0.919</td>
</tr>
<tr>
<td>Secondary school</td>
<td>0.813</td>
<td>0.906</td>
<td>0.168</td>
<td>0.847</td>
<td>0.000</td>
<td>0.764</td>
<td>0.899</td>
<td>0.717</td>
</tr>
<tr>
<td>Tertiary school</td>
<td>0.017</td>
<td>0.872</td>
<td>0.030</td>
<td>0.866</td>
<td>0.715</td>
<td>0.835</td>
<td>0.032</td>
<td>0.804</td>
</tr>
<tr>
<td>Natural population growth</td>
<td>0.221</td>
<td>0.854</td>
<td>0.200</td>
<td>0.818</td>
<td>0.744</td>
<td>0.821</td>
<td>0.199</td>
<td>0.779</td>
</tr>
<tr>
<td>( R^2/S_{comb} )</td>
<td>0.866</td>
<td>0.968</td>
<td>0.844</td>
<td>0.952</td>
<td>0.853</td>
<td>0.966</td>
<td>0.807</td>
<td>0.951</td>
</tr>
</tbody>
</table>

Notes: Export 1: exports of goods and services, constant USD; Export 2: exports of goods and services, current USD; Export 3: merchandise exports, current USD 2; \( R^2 \): adjusted R squared.
<table>
<thead>
<tr>
<th></th>
<th>High-tech, export, secondary school</th>
<th>High-tech, export tertiary school</th>
<th>Natural pop. growth, export secondary school</th>
<th>Natural pop. growth, export tertiary school</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>Sig</td>
<td>$R^2$</td>
<td>Sig</td>
</tr>
<tr>
<td>GDP per capita</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>0.845</td>
<td>0.002, 0, 0</td>
<td>0.854</td>
<td>0.002, 0, 0</td>
</tr>
<tr>
<td>1978</td>
<td>0.864</td>
<td>0.010, 0, 0</td>
<td>0.894</td>
<td>0.030, 0, 0</td>
</tr>
<tr>
<td>1992</td>
<td>0.881</td>
<td>0.001, 0, 0</td>
<td>0.846</td>
<td>0.030, 0, 0</td>
</tr>
<tr>
<td>2007</td>
<td>0.784</td>
<td>0.050, 0, 033</td>
<td>0.813</td>
<td>0.018, 0, 0</td>
</tr>
<tr>
<td>GNI per capita</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>0.835</td>
<td>0.003, 0, 0</td>
<td>0.856</td>
<td>0, 0, 0</td>
</tr>
<tr>
<td>1978</td>
<td>0.864</td>
<td>0.020, 0, 0</td>
<td>0.897</td>
<td>0.001, 0, 0</td>
</tr>
<tr>
<td>1992</td>
<td>0.874</td>
<td>0.001, 0, 0</td>
<td>0.837</td>
<td>0.020, 0, 0</td>
</tr>
<tr>
<td>2007</td>
<td>0.822</td>
<td>0.002, 0, 0</td>
<td>0.857</td>
<td>0.001, 0, 0</td>
</tr>
</tbody>
</table>

Note: $R^2$: Adjusted R squared.
Factors affecting long-term economic growth-consistency

Table 3 presents various optional specifications of the model (using MLR), such that all (or overwhelming majority) the explanatory variables are statistically significant for both dependent variables. When utilising MLR modelling tool, the modelling professionals tend to try all the relevant explanatory variables in the first regression run, and then (if some explanatory variables come out insignificant) they try to find combination of the smaller amount of explanatory variables such that all of them will be statistically significant. Table 3 displays different options of combining three explanatory variables and testing whether they maintain statistical significance over the period of 1965–2007. The table provides four different specifications. Of course, more specifications are possible, but four specifications presented here are sufficient to make the point.

The first specification (high-tech, export, and secondary school) came out successful (explanatory variables significant for every year), although secondary school was a poorly performing variable in Tables 1 and 2. The second specification (high-tech, export, and tertiary school) is also successful where all explanatory variables came out significant for all years. The third specification (natural population growth, export, and secondary school) is successful since all explanatory variables are significant for all years (except natural population growth, which is insignificant only in 2007). Also in this specification, secondary school generates excellent results despite being poorly performing variable in Tables 1 and 2. The fourth specification (natural population growth, export, and tertiary school) generates similar results to the third one.

Thus, Table 3 (based on MLR) generate some results that contradict the conclusions of Tables 1 and 2. We will address this issue in detail in the next two sections.

10 Results

This study has consisted of a very large amount of regression runs – we wanted to make sure that the study is as robust as possible.

We tested our model over extended period – from 1965 to 2007, and utilised two different types of measurements for ‘standard of living’ (GDP and GNI) and three different types of measurements for ‘exports per capita’. In addition, we conducted the study utilising two completely different modelling tools (MLR vs. SR), thus we generated as much as possible information for achieving maximum robustness for the study.

10.1 Unambiguous results

Variable export per capita came out highly significant in various specifications of MLR, as well as based on SR – for all the years covered by this study. Despite utilising three different measurements of exports (numerically substantially different from each other), the variable came out consistently significant. Based on SR, the variable generated in most cases higher measure of significance ($S_{Y,X}$) in comparison to other variables. Both MLR and SR derived consistent results.

High-tech also came out successful explanatory variable. Based on MLR, of all the different specifications and the four different years, this variable came out insignificant only once (Table 2, third specification, year 1992). In addition, based on SR, high-tech is highly significant variable for all the years under study. MLR derived mostly consistent results and all SR results were consistent.
Since exports per capita and high-tech are both proxies of international competitiveness in the original theoretical model in this study, the results definitely validate international competitiveness as a major factor facilitating long-term economic growth.

10.2 Ambiguous results

Tertiary school is relatively successful explanatory variable, but with several cases of being insignificant that are difficult to explain. In Table 1, this variable appears insignificant in all three specifications for the year 1992, and in Table 2, it is insignificant in two out of three specifications for the year 1992. For all other years, this variable is significant in both tables. In Table 3 tertiary school appears in two specifications, and is consistently significant for all the years. Based on SR (which utilises all five explanatory variables combined), tertiary school came out significant in both, Tables 1 and 2 for all the years under study. Thus, as far as the significance of Tertiary school, the results of Table 3 and the SR of Tables 1 and 2 are consistent. The only inconsistency, are the results of MLR in Tables 1 and 2 for year 1992. It is difficult to explain this inconsistency on logical grounds, and it is most likely a result of technical issues (see above) arising due to lack of independence among the explanatory variables. While we found some MLR results to be inconsistent, all SR results came out to be consistent.

Secondary school is an explanatory variable, which is additional proxy for human capital and for the degree of social progress. Based on Table 1, this variable came out insignificant for all three specifications (MLR) for all years except 1992. Similar results appear in Table 2. On the other hand, based on Table 3, secondary school is consistently significant in two specifications, in all four years. How one can reconcile such contradictions? This is a good example of difficulties experienced by modellers who utilise MLR. In addition, it should not come as a surprise, given the technical issues associated with MLR as were explained above.

On the other hand, according to SR, secondary school is significant for 1965 and 1978, then declines to partial significance in 1992 and 2007, while in 2007 it is approaching the borderline of insignificance. In the early years covered by our study, there was a substantial gap between the secondary education enrolments in wealthy countries as compared to poor countries. However, over the years the enrolment into secondary education institutions has been increasing in the underdeveloped countries, which narrows the gap between developed and underdeveloped countries in this respect. Hence, we expect that the importance of this variable will decline and eventually will fade away. While MLR derived inconsistent results, the SR derived mostly consistent results, and can logically explain fading away phenomenon.

High-tech, tertiary school and secondary school are the three proxies of human capital. Despite the contradictory results of secondary school, the strong results of the high-tech and the tertiary school definitely support human capital as important factor facilitating long-term economic growth.

Natural population growth is another variable, which generated inconsistent results. In Table 1 we can see that natural population growth is significant in 1965 for all specifications of MLR, and then becomes insignificant for other years. In Table 2, natural population growth is insignificant in all regression runs of MLR (all specifications, all years). On the other hand, according to Table 3, the variable natural population growth is significant for 1965, 1978, and 1992 and only drops to insignificance in 2007 (for both
Factors affecting long-term economic growth-consistency

specifications). These results are in line with the results of SR, where natural population growth is significant for 1965, 1978 and 1992, but becomes partially significant (still close to being fully significant) in 2007 in both – Tables 1 and 2. The latest results are also possible to explain on logical grounds: the reason is that many poor countries have been experiencing a gradual decline in the rate of natural population growth. As the difference between the rate of natural population growth between the wealthy countries and the poor countries decreases over time, the variable loses its statistical significance in its ability to explain the gap in standard of living. While MLR derived inconsistent results, the SR method derived mostly consistent results and can logically explain fading away phenomenon.

The variables tertiary school, secondary school, and natural population growth are all proxies for the theoretical variable ‘degree of social progress’. The results above could be interpreted, that two of the proxies (secondary school and natural population growth) were definitely important proxies for the degree of social progress during the early years under study, but in the later part of the period under study, these two variables are in the process of losing their explanatory capability. However, the variable tertiary school has been mostly significant variable (few insignificant cases of MLR runs are most likely due to technical issues discussed above), and therefore validates ‘degree of social progress’ as being relevant factor facilitating long-term economic growth. However, the results also point to the need to identify additional proxy variables for this factor, because of the fact that two of the presently used proxies are in the process of fading away.

It can be summarised, that despite some technical difficulties inherent in MLR, by combining MLR and SR, the extensive and robust testing conducted in this study strongly support international competitiveness, human capital and the degree of social progress as definitely important factors facilitating long-term economic growth.

11 Technical evaluation

Tables 1–3 demonstrate contradictions of results generated by MLR. Some variables came out significant in some specifications and insignificant in others. Only variable Exports per capita came out significant in all specifications. One must remember that theoretically there is only one model specification, which is correct. Other model specifications are incorrect, even if in some cases they generate (with some luck) correct results. The model is subject to misspecification bias. Now the question is which one of the many specifications used in Tables 1–3 is the correct one? Choosing one given specification means we include some variables and exclude other variables. Unfortunately, there is no reliable indicator to make such choice. In addition, there can always be a possibility that all of the specifications attempted in a given study are incorrect. There could be factors, which based on the present perception are not considered important due to our lack of knowledge. There could be some important factors without data, since no one has attempted in the past to measure them. This is the real world situation where numerous modellers and decision makers must operate under severe degree of uncertainty. One of the objectives of this study was to demonstrate the issues involved and to illustrate the difficulties and contradictions.

SR, on the other hand, involved much fewer regression runs, allowed to combine all the explanatory variables into the same function, and did not generate contradictions
among the results. All the results conformed common sense regarding the general behaviour of the data as was explained above (e.g., behaviour of secondary school, natural population growth). The stability and common sense of the results allowed using SR as an additional modelling tool that can help to confirm or reject various model conclusions based on MLR.

In addition to testing and evaluating the model, our objective in this study was also to perform comparison of two quantitative modelling tools MLR vs. SR. The main advantage of MLR is that it is familiar and standardised method, widely practiced by modelling professionals, and when applied appropriately, can generate precise function with predictable and relatively small estimating error due to randomness. The main disadvantage of MLR is that under severe conditions of uncertainty, when the complete and precise specification of a model is unknown, the resulting model could generate distorted results due to the model misspecification bias, thus generating misleading results and therefore being very risky for use for some very critical real world decisions. Thus we can summarise, that if for whatever reason we cannot assure correct model specification, then MLR becomes unreliable modelling tool – even though modelers using MLR attempt to mitigate this problem by utilising add factors. However, using add factors is a tedious, time consuming and essentially patching process which, however, in some cases succeeds to make some models workable. In this case, questionable reliability requires frequent re-evaluation of add factors.

The SR method is not designed to generate precise (excluding random error) function. To the contrary, it assumes in advance that uncertainty and lack of measurable data does not allow us in some cases to specify precise and complete model. However, in this method, one can utilise partial knowledge of a given domain, and to generate partial model, which is logically consistent. If the membership function is in line with human logic and common sense, so will be the results generated by SR, although the generated function is imprecise. For some applications, where generating precise mathematical function is essential, obviously SR is not the right tool to apply. However, for the purpose of supporting decision making, where correctness of the conclusions is more important than precision, SR is definitely a preferable tool.

From the standpoint of computational convenience for the users, both methods are convenient. From mathematical standpoint, SR is simpler than MLR, but MLR has numerous standard computerised programs which are very simple to use.

It seems that under conditions where uncertainty is relatively minor, and the only imprecision expected is the phenomenon of randomness, then, of course, MLR is the right tool to use and there is no point to apply tools such as SR. However, under severe conditions of uncertainty, where in addition to problems of randomness, there are also problems of vagueness and ambiguity as far as the correct and precise model specification, then SR becomes a relevant and very powerful tool (as could be seen from Tables 1–2). It does not mean that MLR becomes a priori useless. It rather seems, that decision makers, researchers and model builders can benefit by combining the two methods. In particular, we can use SR to validate whether MLR model is correctly specified. In addition, if the SR disqualifies a model constructed with MLR, then we can prevent misspecification and avoid critical mistakes. In such case, less precise, but undistorted results of SR are preferable.
12 Conclusions

In this study, we have conducted extensive quantitative testing of several factors that are supposedly associated with higher standard of living: ‘international competitiveness’, ‘human capital’ and ‘social progress’. Since to the best of our knowledge there are no direct measurements of such factors, we utilised proxy variables instead. The study covered a period from 1965 to 2007.

In order to demonstrate the robustness of the model, we utilised data based on different measurement methodologies for standard of living (dependent variable) and for ‘exports per capita’ (one of the explanatory variables).

The results generated by SR validated significance of all the proxy variables, while showing some weakening of two of them toward the end of the testing period. The results point out to the need to replace the variable ‘secondary education’ as a proxy variable in the future.

The results of MLR were less definite. They definitely support international competitiveness and human capital as important factors, while ‘degree of social progress’ became ambiguous throughout the period.

Testing of the model allowed us to perform comparison between SR and MLR as practical quantitative modelling tools. SR tool generated more stable and consistent results in comparison to linear regression (due to large number of possible specifications of MLR models). The SR is a powerful modelling tool with stable and consistent results, and we can use it in combination with the conventional regression tools, or on its own, according to the objectives of decision makers.

References


