Being a Good Econometrician is not Enough!¹

By Mariusz Maziarz

In this article, I present a new argument against McCloskey’s rhetoric of economics and discuss the Reinhart-Rogoff affair. First, I operationalize the epistemic advices given by McCloskey’s rhetoric of economics. Second, I conduct an econometric exercise and produce contradictory models of the relation between individualist/collectivist orientation of a society and economic development. Third, I analyse Reinhart and Rogoff’s recently widely discussed ‘Growth in a Time of Debt’ and its criticism. These two case studies show that conducting a research in accordance with a McCloskian definition of ‘goodness’ leads to accepting contradictory hypotheses. Therefore, the rhetoric of economics fails to assist in the problem of theory appraisal. Two solutions to the difficulty that arises are offered. Either a realist point of view should be applied or a constructivist approach should be enriched by elaborating on a method of theory choice.

Keywords: good economist, McCloskey, scientific realism, critical realism, Reinhart-Rogoff, rhetoric of economics

1. Introduction

The choice of theory in an empirical manner is one of the main issues that any economic methodology has to deal with (Hausman 1989). My aim is to present an argument against McCloskey’s (1998) rhetoric of economics by demonstrating on two examples that her methodology is not helpful in deciding which of contradictory econometric models is better (to avoid using the word true). Furthermore, I consider the recently widely discussed affair of Reinhart and Rogoff’s ‘Growth in a Time of Debt’ and its criticism (Herndon et al. 2014) and conclude that both researches are ‘good’ (in McCloskian understanding of this term).

Recently, two research programs dominated debate in philosophy of economics. On the one hand, scientific realists state that models are used by economists as a representation of actual (or possible) reality (Mäki 2009). Due to the complexity of economic reality, building a (theoretical) model is only possible by means of isolation and idealisation of the studied phenomenon. On the other hand, in accordance with the rhetoric approach, one analyses models as metaphors that are constructed in order to persuade (McCloskey 1998), hence verisimilitude is not their fundamental feature.

In spite of recent attempts to fuse rhetoric with realist analysis (e. g. Hardt 2014, Mäki 1988), these two research programs are contradictory in several aspects. First, the economist’s goal is defined differently. Namely, according to the realist point of view, economists should

¹ This work was supported by the National Science Centre, Poland under Grant number 2015/19/N/HS1/01066.
explain (Salmon 1998). In contrast, McCloskey writes that the researcher’s aim is to persuade (McCloskey 1998).

Second, the epistemology is different, too. The correspondence theory of truth (big-T Truth according to McCloskey (1995)) claims that falsity or truth of a statement or belief is determined by its relation to the (economic) world. Realists\(^2\), in general, define truth in terms of correspondence (Peter 2001) or in terms of essesimilitude (Niiniluoto 2013), i.e. correspondence with only those important aspects of reality (essences or ontic bases) that a model focuses on. Mäki (1995) claims that McCloskey supports only the coherence theory of truth (the small-t one), which states that an explanation or an idea is true if it is coherent (i.e. consistent) with other beliefs held. McCloskey (1995) rejects this interpretation and elaborates her position based on a refusal to choose between the correspondence and coherence definitions of truth.

The same point of view, i.e. supporting both the correspondence and coherence theories of truth, seems to be popular among practising economists. For instance, Knight, when answering the question ‘What is Truth in Economics’ (1940), highlighted the importance of conversation and at the same time was a realist: The essential point for our purpose is that knowledge of external reality presupposes “valid” intercommunication of mental content, in the sense of the knowledge, opinion, or suggestion, among the members of a knowing group or intellectual community. A conscious, critical social consensus is of the essence of the idea of objectivity or truth. McCloskey seems to be aware of the difficulty of judgment whether a hypothesis corresponds with reality and elaborates on her own account of theory choice: a theory should be preferred alternatives if it is more persuasive (McCloskey 1998), cf. the next section. Furthermore, in spite of being a realist (in terms of believing in the existence of external reality at least (Mäki 1995)), she supports a coherence theory of truth as epistemic guidance (McCloskey 1995)\(^3\).

Third, realism and rhetoric differ in methodology as well. Realists state that mechanisms underlying observable reality (Mäki2013) or capacities (Cartwright 1989) are in principle discoverable by isolating small fragments of the world or by idealizing it. In spite of

\(^2\) There certainly is no single concept of realism or one definition of a theory of truth, thus the following differentiation between realism and rhetoric of economics is simplified.

\(^3\) The difficulties in interpreting McCloskey’s (1995) views on truth can be solved by distinguishing between the ontological and epistemic definitions, as Collier (1994) and Lawson (1997: 237) suggested. According to the critical realism tradition, truth corresponds (ontologically) with reality. However, since the relation of correspondence is epistemically inaccessible, other features (such as, for instance, coherence) should be employed to the problem of theory appraisal.
the recent voices defending modernist methodology (Blaug 1993), the methodology of mainstream economics seems now to be much more liberal than the concept criticized by McCloskey. However, she argued against any concept of methodology because it spoils scientific conversation, and says therefore let’s get rid of them (McCloskey 1989).

Instead of methodology, McCloskey advises a researcher to be good, i.e. she encourages honesty, clarity and tolerance. Her understanding of the term ‘good economist’ is described in Section 2. Section 3 consists of an econometric exercise aimed at assessing the relation between individualism/collectivism and the pace of economic development. Then, I apply the (small-m) methodology advised by McCloskey to the problem of choosing one of the two contradictory empirical models. Afterwards (in Section 4), I discuss the Reinhart-Rogoff affair and show that both supporters and opponents of the threshold hypothesis are ‘good’ in the McCloskian understanding of goodness. Therefore, both case studies illustrate that the rhetoric of economics does not provide a framework for choosing one of contradictory hypotheses. In Section 5, I discuss possible solutions to the problem of theory choice. Finally, I present conclusions and suggest areas of further research.

2. Who is a ‘Good Economist’?

In this section, I attempt to operationalize McCloskey’s methodological recommendations. It should be highlighted that my aim is by no means to review or summarize McCloskey’s critique in general. Taking into account difficulties in interpreting her views described in literature (cf. Mäki 1995; McCloskey 1995), operationalizing the rhetoric of economics will certainly simplify this approach. It would otherwise be impossible to express McCloskey’s methodological advices in a way that would make a comparison of two hypotheses possible. In her famous book, McCloskey (1998) criticized the modernist methodology of economics and wrote that the rules of scientific research limit the researcher’s potential and creativity: Constraints, after all, constrain (McCloskey 1998: 159).

Hausman (2013) argued that Deidre McCloskey denies that there are any non-trivial methodological standards that economics must meet. Her resistance to (big M) Methodology made Garnett (2005: 239) write that while McCloskey is careful to distinguish her rhetoric...
from anything goes relativism, her vision of science bears a notable similarity to the methodological and epistemological anarchism of Paul Feyerabend. In a similar vein, Boylan and O’Gorman (2004: 46) acknowledged that there are intriguing parallels between economic rhetoric and epistemological anarchism. However, contrary to the author of Against Method (1993), McCloskey attempted to formulate a few standards of economic conversation which she labelled (small m) method and those economists who follow these advices “good”.

Instead of disapproved modernist methodology\(^6\), she advised economists to be good: *What is good for science now (...) is good scientists, in most meanings of "good"* (McCloskey 1998: 186). McCloskey enumerated a set of practical rules called *method with a small m* (McCloskey 1998: 159). What to do when the data are biased or when it is hard to think of a reason for the observed phenomenon, how to write scientific prose and how to avoid mistakes of statistical significance are only a few examples of the method (McCloskey 1998: 159). An economist who follows these rules writes articles that present novel and significant results, is persuasive and pays attention to other minds present in the discussion.

The first factor that differentiates a good economist from a bad one is novelty of her/his research and contribution to a conversation (McCloskey 1998: 162). What is more, good economists appropriately use statistical significance tests, i.e. they differentiate between economic and statistical significance (McCloskey 1998: 136, McCloskey 1985). According to McCloskey’s point of view, economic significance is a term coined to describe research on relation that is important for academia or policy purposes. On the other hand, saying that an estimated coefficient is statistically significant means (only) that uncorrelated variables are unlikely to lead to an obtained estimate.

One of the accusations formulated against modernism by McCloskey is that it demands (often the impossible) for empirical verification: *The problem comes, and the modernist shouting begins, with the words “empirical and “evidence.” Should it all be “objective,” “experimental,” “positive,” observable”? Can it be? I doubt it* (McCloskey 1998: 158). In accordance with the rhetoric of economics, right theories (not to say true ones) are found to be more persuasive: *the people speaking in a conversation of science are often worth listening to when a scientific assertion is at issue. (...) Small-t truth is about social*

---

\(^6\) It should be noted that the current economic methodology, in spite of a few voices aimed at defending modernist approach (e.g. Blaug 1993), is more liberal than the concept criticized by McCloskey.
agreement, not God’s mind (McCloskey 1995). Therefore, McCloskey (1998) expressed that a theory should be preferred to alternative ones if it is more persuasive.

McCloskey identifies a good economist with a good rhetor as s/he uses language and does it with self-awareness and attention to other minds present in a discussion (McCloskey 1998: XIX; 5). The so-called Sprachethik is based on the following rules: Don't lie; pay attention; don't sneer; cooperate; don't shout; let other people talk; be open-minded; explain yourself when asked; don't resort to violence or conspiracy in aid of your ideas (McCloskey 1998: 160).

In summary, the method advised by McCloskey is based on being ‘good’, i.e.: reaching novel and economically significant conclusions, producing persuasive theories and hypotheses, and paying attention to other minds present in the discussion. Does the methodology of the rhetoric of economics provide a sufficient framework for solving the problem of theory choice? Below, I describe two case studies (my own econometric exercise and the Reinhart-Rogoff affair) and apply the rules described above that a good economist should use in her/his research to choose between contradictory hypotheses.

3. An Econometric Exercise

Taking to heart McCloskey’s advice on spending more time in the archives (i.e. conducting empirically based researches), I decided to start researching the relation between collectivist and individualist orientation of society and economic development from the point of view of econometric analysis. There are two measures of this dimension of culture that cover most European Union countries. The chronologically first one was developed by Hofstede (1980) who analysed the database of employee value scores collected by IBM. IDV indicator, one of the six dimensions of culture constructed by Hofstede (1980), covers over 70 countries and measures the strength of the individualist orientation in a society. On the other hand, Diener et al. (2000) based their indicator on World Values Survey conducted in 42 countries. IC measures the degree to which a culture is collectivistic. Since individualism and collectivism are usually considered as two opposites of the same dimension of culture (Triandis and Suh 2002), both indicators are aimed at measuring the same dimension of culture and they (to some degree\(^7\)) succeed.

\(^7\) IC and IDV are positively correlated. The correlation coefficient calculated for EU member states included in both IC and IDV databases equals 0,485 (p=0,026) (own calculation).
Mining European macroeconomic data and variables describing individualist/collectivist orientation and running a few hundred regressions (fortunately not two million as Sala-I-Martin (1997) did), I got two empirical models that result in contradictory hypotheses. Taking into account a well-known statement, [*] imply producing a model that comes up with a desired result is a test of the cleverness of the author, not a test of the relevance of the theory ([Blaug 2002]), I can boast about obtaining the desired result twice. Both models’ specifications are based on the findings that Barro (1996) got when studying growth determinants. Due to a smaller sample, I limited the number of variables from eight to the three most significant ones and added a measure of individualist/collectivist orientation.

Model 1 (cf. Table 1) shows that the pace of economic development in 21 European Union member states[^8] is positively related to life expectancy and the level of collectivism and negatively to government consumption and schooling[^9]. Values of coefficients of the White test (p=0,0103) and the Jarque-Bera test (p=0,32) indicate that homoscedacity of the error term, skewness and kurtosis of the data equals 0, accordingly. In addition, the coefficients of Model 1 are not only statistically significant but economically significant[^10], too (cf. Appendix 1: Data).

Table1: Least squares estimation, 21 instances, dependent variable: GDP_growth_02_12

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. dev.</th>
<th>t-Student</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC</td>
<td>-0,19767</td>
<td>0,0949761</td>
<td>-2,0813</td>
</tr>
<tr>
<td>gov_consumption</td>
<td>-0,0391033</td>
<td>0,0221254</td>
<td>-1,7673</td>
</tr>
<tr>
<td>life_expectancy</td>
<td>0,0666898</td>
<td>0,0126215</td>
<td>5,2838</td>
</tr>
<tr>
<td>schooling</td>
<td>-0,0165835</td>
<td>0,00927101</td>
<td>-1,7888</td>
</tr>
</tbody>
</table>

| Arith. mean of the dep. v. | 1,619325 | Std. dev. of the depend. v. | 0,583809 |
| Sum of sq. of the residuals | 3,942063 | Std. dev. of the residuals | 0,481545 |
| Determination coef. R2 | 0,936298 | Adjusted R2 | 0,925057 |
| F(4, 17) | 62,46718 | p-value of the F test | 6,13e-10 |
| Log-likelihood ratio | -12,23312 | AIC | 32,46623 |
| Schwarz bayes. crit. | 36,64432 | HQC | 33,37299 |

[^8] The IC measure developed by Diener et al. (2000) describes the orientation on individualism/collectivism of these 21 member states: Austria, Belgium, Bulgaria, Czech Republic, Denmark, Estonia, Finland, France, Hungary, Ireland, Italy, Latvia, Lithuania, Netherlands, Poland, Portugal, Slovenia, Spain, Sweden, United Kingdom.

[^9] The result may be surprising unless one is aware of Plumper and Schneider’s (2007) hypothesis that schooling is one of the tools of unemployment policy.

[^10] One point (of a hundred-point scale) growth of the IC variable relates to slower economic growth of circa 1,82 per cent point a year, i.e. 19 per cent points faster GDP growth during 10 years.
Then, I exchanged the IC variable for IDV and slightly modified the sample. Contrary to Model 1, which is estimated on the sample of 21 EU member states included in the Diener et al. (2000) research, Model 2 is based on the sample of all Eurozone countries except Luxembourg. It (cf. Table 2) shows that individualism is positively related to economic development, which contradicts Model 1. The regression coefficient of the IDV variable is economically significant: a one-point (of a hundred-point scale) raise of individualism correlates to an increase of over 7 percentage point of economic development within 20 years. The goodness of fit of Model 2 is lower than that of the former one (cf. for example the determination coefficient $R^2$ or information criteria). However, did not McCloskey herself argue against the rhetoric of statistical significance (McCloskey 1998: 112-138)? Furthermore, the shortcoming may be caused by a less numerous sample (or is it just a rhetoric argument?).

Table 2: Least squares estimation, 17 instances, dependent variable: GDP_growth_92_12

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. dev.</th>
<th>t-Student</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDV</td>
<td>0.0722975</td>
<td>0.0537234</td>
<td>1.3457</td>
<td>0.20139</td>
</tr>
<tr>
<td>gov_consumption</td>
<td>-0.633362</td>
<td>0.242772</td>
<td>-2.6089</td>
<td>0.02164 **</td>
</tr>
<tr>
<td>life_expectancy</td>
<td>0.457507</td>
<td>0.155682</td>
<td>2.9387</td>
<td>0.01152 **</td>
</tr>
<tr>
<td>schooling</td>
<td>-0.156299</td>
<td>0.0714104</td>
<td>-2.1887</td>
<td>0.04747 **</td>
</tr>
<tr>
<td>Arith. mean of the dep. v.</td>
<td>4.107574</td>
<td>Std. dev. of the depend. v.</td>
<td>5.357582</td>
<td></td>
</tr>
<tr>
<td>Sum of sq. of the residuals</td>
<td>261,2983</td>
<td>Std. dev. of the residuals</td>
<td>4,483287</td>
<td></td>
</tr>
<tr>
<td>Determination coef. R2</td>
<td>0.649775</td>
<td>Adjusted R2</td>
<td>0.568953</td>
<td></td>
</tr>
<tr>
<td>F(4, 13)</td>
<td>6.029736</td>
<td>p-value of the F test</td>
<td>0.005705</td>
<td></td>
</tr>
<tr>
<td>Log-likehood ratio</td>
<td>-47.34777</td>
<td>AIC</td>
<td>102,6955</td>
<td></td>
</tr>
<tr>
<td>Schwarz bayes. crit.</td>
<td>106,0284</td>
<td>HQC</td>
<td>103,0268</td>
<td></td>
</tr>
</tbody>
</table>

The econometric research described above justifies two contradictory hypotheses:

H1: Collectivism fosters economic development.

H2: Individualism fosters economic development.

Sala-I-Martin (1997) faced the same difficulties when he was researching empirical determinants of growth. The aim of empirical literature on economic growth is to find variables that are correlated with the pace of economic development. One estimates regressions of the form of:

$$Y_t = A + a_i \cdot X_i + \varepsilon$$  \hspace{1cm} (1)

where:

$Y_t$ - pace of economic development

$A$ – constant
Theories of growth are not developed well enough to help in choosing the right model specification (Sala-I-Martin 1997). This observation applies to the problem of determining whether it is individualism or collectivism that fosters economic growth. Sala-I-Martin (1997) explicitly points out that if one runs different regressions, s/he can find two equations where a particular variable’s coefficient is significant in only one of them: *iff one starts running regressions combining the various variables, variable x₁ will soon be found to be significant when the regression includes variables x₂ and x₃, but it becomes nonsignificant when x₄ is included*. As long as a theory is insufficient, the question whether variable xᵢ is really correlated with growth stays open.

Since individualism and collectivism constitute a single cultural dimension (Earley and Gibson 1998; Ho and Chiu 1994; Schwartz 1990), arguing in favour of the two hypotheses in one paper violates the coherence theory of truth, the theory (among others) supported by McCoskey. In order to make the contradiction more visible, I reformulated the above hypotheses. If X₁ of equation 1 is a measure of individualist/collectivist orientation and the higher level of individualism, the higher the value of X₁, then:

H1: \( a₁ < 0 \)
H2: \( a₁ > 0 \)

There is no real number that satisfies both inequalities, therefore it is justified to assert that only one of the hypotheses can be true (there is a third possibility, i.e. collectivism and individualism may be unrelated to the pace of economic growth). Otherwise, unless the choice is made, the findings of the econometric research are contradictory. Attempting to be a ‘good’ economist, I should choose one of the hypotheses. Is it possible to decide which of them is true, more correct, better or, at least, should be preferred on the grounds of the rhetoric of economics?

If I conducted my research in accordance with the Methodology criticized by McCloskey, the choice would be easier, because, for instance, I could decide on the basis of a comparison of the determination coefficient of the models or the statistical significance of the coefficient \( a₁ \) of the measures of the individualist/collectivist orientation. But, both econometric models show economic significance (due to epistemic reasons, at least) of the relation between collectivism and individualism (respectively) and economic growth. In accordance with Model 1, one-point decrease in the IC indicator correlates to an increase of
19 percentage points in GDP growth during 10 years. On the other hand, the coefficient of the IDV variable in Model 2 states that a one-point increase in it correlates to an increase of 7 percentage points in GDP growth during 20 years. In spite of the differences, both numbers are economically significant. Additionally, both models are novel in the literature of growth determinants. Taking into account McCloskey’s argumentation (1998: 114), one would get a statistically significant coefficient of the IDV variable in Model 2 by raising the number of countries included in the sample of Model 2\(^{11}\). Yet a ‘good’ rhetoric differentiates economic significance from statistical significance and pays attention to the former (McCloskey 1998: 136, McCloskey 1985). Therefore, these two models may be considered persuasive metaphors.

The next advice of McCloskey is to pay attention to other minds present in the discussion. However, taking it into account is not conclusive either. It is easy to be persuaded by previous researches that argue in favour of hypothesis 2. For example, one can appeal to the authority of Smith\(^{12}\) (2007, Book 4, Chapter 2), who claimed in his famous ‘An Inquiry into the Nature and Causes of the Wealth of Nations’ that individual behaviours driven by egoism and self-interest maximize everyone’s wealth. In addition, some recent economists support this hypothesis. For example, Moris et al. (1994) conclude their research with the following statement: it appears that entrepreneurship declines the more collectivism is emphasized. Second, Peterson (1980 p. 69-70) asserts that a growing collectivist orientation hinders economic growth as belonging requires a lower level of orientation toward achievement among the people. Hofstede (1991: 76) makes a similar argument.

On the other hand, one can find the opposite point of view among a number of researchers. For instance, Allik and Realo (2004) argue that the level of social capital is positively related to collectivism. Ball (2001) presents evidence in favour of hypothesis 1: collectivist orientation alleviates effects of market failures. Moreover, collectivism is known to raise the level of interpersonal trust and (therefore) reduce transaction costs (Huff and Kelley 2005). Likewise, collectivist orientation improves cooperation within groups (Wagner 1995).

In conclusion, being a good econometrician (as defined by McCloskey, cf. Section 2) is not enough to decide which of two hypotheses is better and should be supported or advocated by a good economist. If one takes the rhetoric of economics seriously, s/he should

---

\(^{11}\) Expanding sample reduces standard deviations of a coefficient’s estimation.

\(^{12}\) At least unless one is aware of Smith’s point of view presented in *The Theory of Moral Sentiments*. 
accept both hypotheses because of their economic significance, persuasiveness and the support they find in literature. However, in this case the rhetoric of economics is not helpful in choosing one of contradictory hypotheses.

4. A Better-Known Example

It is easy to disregard the above argument against taking the rhetoric of economics too seriously by claiming that data mining aimed at constructing models that support contradictory hypotheses is not what economists usually do. However, the recent affair of the widely commented paper ‘Growth in a Time of Debt’ (Reinhart and Rogoff 2010) criticized by Herndon, Ash and Pollin (2014) is evidence of the contrary.

Concluding whether the attack was justified exceeds the scope of this paper. Nonetheless, the necessity of a brief summary of the discussion is highlighted by Shuchman’s (2013) opinion on the article: *it is perhaps the most quoted but least read economic publication of recent years*. Moreover, I would like to show that the other two drawbacks pointed out by Herndon (et al. 2013), apart from the spreadsheet error, resulted from different subjective choices made during the research. Reinhart and Rogoff (2010) presented their empirical research on the relation between the ratio of government debt to GDP and the pace of economic growth. The analysis based on forty-four countries spanning circa two hundred years leads to the conclusion that above 90 percent (government debt/GDP ratio), median growth rates fall by one percent, and average growth fall considerably. (Reinhart and Rogoff 2010).

Even considering the research apart from causal interpretation (which is literally rejected by the authors), the above finding is certainly economically significant due to being important for academia purposes (as McCloskey (1998: 159) would call it) since Reinhart and Rogoff’s (2010) analysis is based on a new dataset collected by the authors and leads to a novel finding. Moreover, ‘Growth in a Time of Debt’ turned out to be very persuasive. In fact, Krugman (2013) called this paper famous and described as surely the most influential economic analysis of recent years. Owing to publication during the financial crisis and at a time when the preference of the most prominent policy makers changed in favour of austerity, the conclusions of the research immediately got attention not only among academics but influenced the most important economic policy practitioners worldwide, including Manuel Barroso and Olli Rehn, President and Vice-President of the European Commission; German Chancellor Angela Merkel and her Finance Minister Wolfgang Schäuble; the Labor and Tory
governments in the UK; and of course large parts of the political elite in Washington DC (Botsch 2013). For instance, Oli Rehn, EU Commissioner for Economic Affairs, advised austerity policy due to the fact that *serious empirical research has shown that at such high levels, public debt acts as a permanent drag on growth* (Smith 2013).

Then, at the peak of its popularity, Herndon (et al. 2014) published the findings from a replication\(^{13}\) of Reinhart and Rogoff’s paper (2010). The economists working at the University of Massachusetts found three main drawbacks, including a coding error that excluded Australia, Austria, Canada and Denmark, selective exclusion of New Zealand and an unconventional weighting of the summary statistics. They concluded that the average GDP growth is not different when the debt/GDP ratio exceeds 90%. The critique resulted in a huge number of disapproving articles being published mostly in popular press.

Some of the commentators explicitly accused Reinhart and Rogoff of fudging and flubbing empirical work in order to get the result they desired (Okalow 2013; Wray 2013). Is this criticism justified? The authors of ‘Growth in a Time of Debt’ surely misspecified a spreadsheet formula. However, the mistake did not influence their result in a major way. Strictly speaking, the Excel error is the least influential of the three drawbacks indicated by Herndon (et al. 2014)\(^ {14}\). On the other hand, the choice of averaging scheme and the exclusion of New Zealand data, contrary to popular point of view on the Reinhart-Rogoff affair, are methodological decisions rather than mistakes. Therefore, the Reinhart-Rogoff affair possibly exemplifies the emerging contrary results phenomenon (ERR\(^ {15}\)) rather than resulting from the spreadsheet error.

Reinhart and Rogoff (2013a) wrote in an errata to ‘Growth in A Time of Debt’ that the selective exclusion of New Zealand (and Spain) was caused by data novelty and more than one time series being accessible, cf.: (...) *there were still gaps in our public debt data set at the time of this paper, a data set no one else had ever been able to construct before and which we now have filled in much more completely*. (Reinhart and Rogoff 2013a). For example, New Zealand’s GDP estimates were delivered by Angus Maddison’s Database and The New Zealand Historical Statistics. A glance at the two bottom cells of Table 3 explains the exclusion due to arbitrariness of the database choice that strongly affects overall statistics.

---

\(^{13}\) Which was in fact a robust check (Clemens 2015).

\(^{14}\) The difference caused by the miscoded formula equals from 0.1 (in the case of the basket of low debt) to 0.3 percentage point in the case of the basket of very high debt (Reinhart and Rogoff 2013b).

\(^{15}\) Cf. Goldfarb (1997) for a deeper analysis of the emerging recalcitrant result (ERR) phenomenon and causes of its occurrence.
Table 3: Average and median growth in the four baskets (Reinhart and Rogoff 2013a).

<table>
<thead>
<tr>
<th></th>
<th>RR(2010) from paper—see Summary_with_coding_error</th>
<th>30% or less</th>
<th>30-60%</th>
<th>60-90%</th>
<th>above 90%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Long Sample</td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td>Post war</td>
<td>3.7</td>
<td>3.0</td>
<td>3.4</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Post war</td>
<td>4.1</td>
<td>2.8</td>
<td>2.8</td>
<td>-0.1(0.0)</td>
</tr>
<tr>
<td></td>
<td>Coding corrected—see Summary corrected</td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td>Long Sample</td>
<td>3.6</td>
<td>3.1</td>
<td>3.2</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>Post war</td>
<td>4.0</td>
<td>3.0</td>
<td>2.5</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Coding corrected—Spain filled 1960-1980—see Summary_Spain Full Sample</td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td>Long Sample</td>
<td>3.8</td>
<td>3.1</td>
<td>3.2</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>Post war</td>
<td>4.2</td>
<td>3.0</td>
<td>2.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

The third objection raised by Herndon et al. (2014) concerns the unconventional weighting of summary statistics. In addition, Reinhart and Rogoff (2010) did not describe or justify their methodology, which made Herndon, Ash and Pollin call the averaging arbitrary and unsupportable. Reinhart and Rogoff (2010) calculated average GDP growth for each country within a group without taking into account the number of years a country in question belonged to one of the four groups.

Are the accusations of arbitrariness and unsupportable methodology justified? On the one hand, the applied averaging scheme equalled the influence of instances on the mean values without taking into account how many years a particular country belonged to one of the four groups. Herndon (et al. 2014) exemplified this problem with a case of United Kingdom and United States: real GDP growth in the UK averaged 2.4 percent per year during 19 years that the UK appeared in the highest public debt/GDP category while real
GDP growth for the US averaged -2.0 percent per year during the 4 years that the US appeared in the highest category. Therefore, Herndon et al. (2014) chose to apply weighted arithmetic mean, cf. Table 4. On the other hand, the methodology can be defended in the following ways. First, Reinhart and Rogoff (2013b) advocate their point of view and justify what they believe was a perfectly natural (...) and hardly unconventional methodology that was applied in order to solve the problem of excessive weight of some countries belonging to one of the four groups differentiated due to their level of debt/GDP ratio: [t]he post war Advanced Economy experience would quickly reduce to the experiences of Greece and Japan. (Reinhart and Rogoff 2013b.) Second, they argue that their approach had been applied to other researches where overweighting a limited number of countries was undesired (Reinhart and Rogoff 2013b). Third, as the critics themselves acknowledge, public debt seems to be serially correlated (Herndon et al. 2014).

Table 4: An excerpt from Table 3: Published and replicated average real GDP growth, by public debt/GDP category(Herndon et al. 2014).

<table>
<thead>
<tr>
<th>Method/Source</th>
<th>Public debt/GDP category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Below 30 percent</td>
</tr>
<tr>
<td>Corrected results</td>
<td>Country-year weighting, all data</td>
</tr>
</tbody>
</table>

Replication elements

<table>
<thead>
<tr>
<th>Separate effects of RR calculations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet error only</td>
</tr>
<tr>
<td>Selective years exclusion only</td>
</tr>
<tr>
<td>Country weights only</td>
</tr>
</tbody>
</table>

RR Published Results

| RR [2010a] Figure 2 (approximated) | 4.1 | 2.9 | 3.4 | -0.1 |
| RR [2010b] Appendix Table 1       | 4.1 | 2.8 | 2.8 | -0.1 |

The authors of ‘Growth in A Time of Debt’ applied an averaging method that was found unconventional by their opponents; however, in accordance with McCloskey’s rhetoric of economics one cannot argue that the method used by Reinhart and Rogoff is nonstandard and (therefore) inappropriate, because the standard must be argued (McCloskey 1998: 114). Summing up, both articles described are very persuasive and present findings of high economic significance. What about other minds present in the discussion?

The critics (especially those who publish in newspapers) of Reinhart and Rogoff (2010) do not admit that the literature on the threshold in the relation between debt to GDP ratio and GDP growth is decidedly divided. For instance, Cecchetti (et al. 2011) analysed a
new dataset containing the levels of government, non-financial corporate and household debt in 18 OECD countries from 1980 to 2010. The economists working for Bank for International Settlement concluded their research in a similar way to ‘Growth in a Time of Debt’, i.e. with best guess (...) that there is a threshold at something like 85% of GDP, but the estimate of the impact is extremely imprecise. In addition, they reviewed theoretical mechanisms that possibly connect debt to GDP growth. Furthermore, Minea and Parent (2012) researched a new IMF database covering more than a century from 1890 to 2012 and found evidence for a nonlinear relation and existence of two thresholds at the level of 90% and 115% of debt to GDP ratio. Reinhart (et al. 2012) arrived at the same conclusion having analysed a database spanning from the early 1800s to 2010 and discovered that when the debt to GDP ratio has exceeded 90% for at least five years, GDP growth is more than one percent lower. Checherita-Westphal and Rother (2012) drew a similar conclusion on the basis of an analysis of the euro area covered by their 40-year database that started in 1970.

On the other hand, there are some empirical studies that argue in favour of a linear relation or no relation at all\(^{16}\). For example, Pescatori (et al. 2014) used a novel empirical approach and found no evidence of any particular debt threshold above which medium-term growth prospects are dramatically compromised (Pescatori et al. 2014).

In conclusion, there is a number of both empirical and theoretical studies that, like Reinhart and Rogoff’s (2010), but (assuminally) without any coding errors, argue in favour of hypothesis 1:

\[
\text{H1: High debt/GDP levels (90 percent and above) are associated with notably lower growth outcomes.} \quad \text{(Reinhart and Rogoff 2010)}
\]

However, there are researchers that support the opposite point of view. Similarly to Herndon, Ash and Pollin (2014), Pescatori (et al. 2014) and a vast number of other scholars state that hypothesis 2 should be preferred:

\[
\text{H2: GDP growth at public debt/GDP ratios over 90 percent is not dramatically different than when debt/GDP ratios are lower.} \quad \text{(Herndon et al. 2013)}
\]

\(^{16}\) Although it certainly is a difficult task to report conducting an empirical research and finding no evidence.
As long as both studies conducted by Reinhart and Rogoff (2010)\textsuperscript{17} and by Herndon, Ash and Pollin (2013) fulfil the requirements of McCloskey’s method\textsuperscript{18} (i.e. persuasiveness, economic significance and paying attention to other minds present in the discussion), then a ‘good’ economist should accept both hypothesis 1 and hypothesis 2, which inevitably leads to a contradiction.

5. Possible Solutions

The Rainhart-Rogoff affair taught us that every economic analysis should be replicated and pass a number of robust checks. The conclusions also apply to my research. The attempt of operationalizing the McCloskian method should be repeated and compared with these findings. Moreover, additional case studies should be conducted. If other philosophers of economics arrive at similar results, i.e. if further investigations confirm that the rhetoric approach is not helpful in the problem of theory choice since the excessively broad definition of ‘goodness’ (not to use the word truth) leads to accepting contradictory hypotheses, then a reflection on McCloskey’s ideas or on the philosophy of economics in general is needed.

There are, at least, two possible solutions to the problem that arose. First, one can defend the rhetoric of economics by acknowledging that scientists, on the basis of the same data but different methods considered ‘good’, reach opposite conclusions. In this case, one accepts that measurement in economics is not objective and conclusions of econometric research cannot be generalized. For instance, the two hypotheses based on my econometric exercise are not contradictory only if IC and IDV do not measure the same feature of culture. In fact, Chen and West (2007) argues that IDV, despite being the most popular indicator of individualism that is used in economics, is more connected to employees’ approach to work than to individualist orientation of a society. Due to such a criticism of IDV, IC was developed (Diener et al. 2000).

The difficulties arise when one takes into consideration the ‘Growth in a Time of Debt’ controversy. The case of New Zealand shows that even such well-defined variables

\textsuperscript{17} McCloskey would certainly oppose the point of view that a research containing such an obvious mistake as the spreadsheet error committed by Reinhart and Rogoff (2010) is good, but the miscoded formula influenced the results in a minor way. In other words, the discrepancy between ‘Growth in a Time of Debt’ and its replication resulted from different averaging schemes and the methodological decision on New Zealand’s data. In fact, even the corrected findings show that there is a nonlinear difference between the third and fourth group of countries, i.e.at the level of 90% debt to GDP ratio, cf. Table 2 (above).

\textsuperscript{18} Some of the abovementioned critics doubt Reinhart and Rogoff’s honesty. However, it is impossible to prove their error being committed on purpose.
used in econometric modelling as GDP differ in estimated values. Furthermore, conclusions from econometric studies contradict each other due to choice of samples. For example, Reinhart and Rogoff (2010) described two periods (i.e. the post-war sample and the long one). On the other hand, Herndon (et al. 2014) focused their criticism on the former, which may raise suspicions that the choice of data was caused by the fact that the findings based on the post war sample support their point of view that there is no threshold in the relation between debt to GDP ratio and pace of economic development. Nevertheless, data choice is not the only problem faced by Reinhart and Rogoff (2010) and Herndon (et al. 2014) since their contrary findings are to a large degree an outcome of the different averaging schemes. Panizza and Presbitero (2013) reviewed empirical literature and concluded that the thresholds are immune neither to small changes in data (i.e. period and sample taken into considerations) nor to empirical techniques.

In this case, the fact that there are different groups of economists that consider diverse methods as ‘good’, reach contradictory conclusions and present them as true demands further explanation. For instance, Fleck (1979) faced a similar contradiction investigating development of medical theories on syphilis. The early twentieth century microbiologist elaborated on his own point of view on the philosophy of science. Namely, a scientist’s opinions are shaped by the thought style one belongs to (defined in terms of a scientist’s discipline, educational background and discussions with co-workers). Fleck concluded his book by underlining that two different thought styles usually hold divergent (and sometimes contradictory) hypotheses as true: scientific observation differs when two different thought styles are involved (Fleck 1979: 133). This way of reasoning saves McCloskey’s methodology but does not help to solve the problem of choosing one of contradictory hypotheses if the choice is to be made (for instance for economic policy purposes). Therefore, the constructivist approach to philosophy of economics should elaborate on a method of theory appraisal.

Second, the realist approach can be applied to the problem of deciding which of contradictory hypotheses is better, true or, at least, should be supported. By “the realist approach” I mean the two realist research programs present in the philosophy of economics. Even though scientific realism (mostly advocated by Mäki in its minimal form (2011)) and critical realism (employed by Lawson (1997) to criticizing the mainstream economics) differ
to a high degree\textsuperscript{19}, \textit{critical realism certainly shares many central tenets with Mäki’s realist vision of economics such as the emphasis on observer-independent truth and the importance of finding causal mechanisms} (Kuorikoski and Ylikoski 2012: 256). The methodologists who support the realist approach (apart from the abovementioned, for instance John Mill, Daniel Hausman and Nancy Cartwright) agree on the importance of investigating underlying causal mechanisms or tendencies rather than regularities which are essential for science viewed from the instrumentalist and empiricist perspectives.

The two case studies discussed above show that the objections against econometrics formulated, among others, by Darnell and Evans (1990) are still up to date. Since empirically driven attempts to discover economic relations lead to uncertain or possibly misleading findings that strongly depend on methodological decisions (e.g. choosing averaging method, data sample etc.), supporting econometric research with theoretical investigations\textsuperscript{20} is possibly more fruitful and may give a better understanding of the modelled phenomenon (Stinchcombe 1991). For instance, an analysis of causal mechanisms connecting individualism/collectivism and economic development such as, for example, more efficient market relations and higher innovativeness of collectivistic cultures shows that hypothesis 1 should be preferred, i.e. collectivism fosters economic development (Maziarz 2015). Similarly, the Reinhart-Rogoff affair can be solved by focusing on theoretical investigations instead of searching for empirical regularities. The inverse association between public debt and economic growth is widely known. In order to explain the threshold hypothesis, it is helpful to consider the concept of debt intolerance introduced by Reinhart, Rogoff and Savastano (2003), according to which default risk (and hence risk premium) are nonlinear functions of debt. In the light of arguments coined by Giordano and Tommasino (2011), debt thresholds met by different countries are located at divergent levels of the debt to GDP ratios, which supposedly undermines the cliometric project in this field since the empirical results are sensitive to methodology and sample choice. Bell, Johnston and Jones (2015) concluded their econometric research by indicating that debt thresholds experienced by different economies are varied indeed.

\textsuperscript{19} In fact, these research programs contradict each other in some points. For instance, Lawson explicitly criticizes the methodology of isolation (advised by scientific realists), arguing that it justifies omitting important layers of reality (Hodge 2008).

\textsuperscript{20} Theoretical investigations can be understood either according to scientific realism, as building models aimed at isolating the important causal relations, or as Lawson and other critical realists would prefer, as focusing on mechanisms underlying observable reality.
It should be highlighted that supporting econometric findings with mechanistic evidence certainly does not offer objective standards that will determine which of two contradictory hypotheses is right. However, comparing results of econometric observation with analysis focused on causal mechanisms (either inferred from theoretical models by means of isolation or researches aimed at analysing deeper layers of reality) can help in producing new evidence that will possibly shed light on the verisimilitude of the considered hypotheses.

6. Conclusions

Above, I showed that the rhetoric of economics does not set a methodological framework for econometricians to choose one of two contradictory hypotheses. Lack of methodological constraints leads an economist loyal to McCloskey’s point of view to accept two contrary hypotheses on the basis of their persuasiveness, economic significance and review support. There are two possible solutions to the difficulty that arises. The rhetoric of economics can be defended by accepting that scientists belonging to different thought styles accept different methods and obtain contrary results that are ‘good’ for them. Such a strategy dismisses my argument against the rhetoric approach, but does not offer a framework for choosing one of the contradictory hypotheses. The questions how to choose a hypothesis that should be preferred and what makes different economists find contrary hypotheses as ‘good’ stay open. Today, realist philosophy of economics, in the light of the argumentation presented, seems to be more fruitful in the area of theory choice, which is one of the most important goals of any methodology of science.

References


http://www.nytimes.com/2013/04/19/opinion/krugman-the-excel-depression.html?r=1, April 18, 2013 (4.03.2015).


64. Reinhart C., Rogoff K., (2013a), *Errata: “Growth in a Time of Debt”*,
http://www.nytimes.com/interactive/2013/04/17/business/17economix-
response.html?_r=1&, April 17, 2013 (5.03.2015).
Completely Misses the Point*, Forbes,
http://www.forbes.com/sites/realspin/2013/04/18/that-reinhart-and-rogoff-committed-
71. Smith J., (2013), *From Reinhart&Rogoff’s own data: UK GDP increased fastest when
debt-to-GDP ratio was highest – and the debt ratio came down*, “Prime Economics”,
73. Triandis H., Suh E., (2002), *Cultural influences on personality*, “Annual Review of
Personality”, 53, 133-160.
75. Wray R., (2013), *Why Reinhart and Rogoff Results are Crap*, EconoMonitor,
http://www.economonitor.com/lrwray/2013/04/20/why-reinhart-and-rogoff-results-
are-crap/, April 20, 2013 (4.03.2015).
al. (ed.) Economics for Real: Uskali Mäki and the Place of Truth in Economics, 241-
251.