ASERS

Theoretical and Practical Research in Economic Fields

Biannually
Volume III
Issue 2(6)
Winter 2012

ISSN 2068 – 7710

Winter 2012
Volume III, Issue 2(6)

ASERS Publishing
http://www.asers.eu/asers-publishing
ISSN 2068 – 7710

Contents:

Call for Papers
Summer_Issue 2013

Theoretical and Practical Research in Economic Fields

Many economists today are concerned by the proliferation of journals and the concomitant labyrinth of research to be conquered in order to reach the specific information they require. To combat this tendency, **Theoretical and Practical Research in Economic Fields** has been conceived and designed outside the realm of the traditional economics journal. It consists of concise communications that provide a means of rapid and efficient dissemination of new results, models and methods in all fields of economic research.

Theoretical and Practical Research in Economic Fields publishes original articles in all branches of economics – theoretical and empirical, abstract and applied, providing wide-ranging coverage across the subject area.

Journal promotes research that aim at the unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking. It explores a unique range of topics from the frontier of theoretical developments in many new and important areas, to research on current and applied economic problems, to methodologically innovative, theoretical and applied studies in economics. The interaction between empirical work and economic policy is an important feature of the journal.

Theoretical and Practical Research in Economic Fields, starting with its first issue, it is indexed in RePEC, IndexCopernicus, EBSCO, ProQuest, CEEOL databases.

The primary aim of the Journal has been and remains the provision of a forum for the dissemination of a variety of international issues, empirical research and other matters of interest to researchers and practitioners in a diversity of subject areas linked to the broad theme of economic sciences.

All the papers will be first considered by the Editors for general relevance, originality and significance. If accepted for review, papers will then be subject to double blind peer review.

Invited manuscripts will be due till June 15th, 2013, and shall go through the usual, albeit somewhat expedited, refereeing process.

Deadline for submission of proposals:	15 June 2013
Expected publication date:	July 2013
Website:	www.asers.eu/journals/tpref/
E-mail:	tpref@asers.eu

To prepare your paper for submission, please see full author guidelines in the following file: TPREF_Full_Paper_Template.doc, then send it via email at tpref@asers.eu.

ASERS Publishing

Volume III Issue 2(6) Winter 2012

DOI: 10.2478/v10261-012-0005-2

INSIDE FINANCE CONSTRAINTS ON THE GROWTH OF ITALIAN SMALL MEDIUM SIZED ENTERPRISES

Cristiana DONATI
Department of Economics, Second University of Naples, **Italy**
cristiana.donati@unina2.it
Giuseppe CINQUEGRANA
ISTAT, Italian Statistic Institute
gcinquegrana@istat.it
Domenico SARNO
Department of Economics, Second University of Naples, **Italy**
domenico.sarno@unina2.it

Abstract:

Our study confirms that the financial constraints of the small and medium size enterprises' (SME's) growth tend to appear as an excess of sensibility of the investment expenditures on the firm's cash flow. Through the application of dynamic panel data techniques to an extended version of Eulero's investment equation of a sample of Italian SMEs, the analysis shows that the growth of small firms subsample in backward regions of Italy is more constrained by inside finance than that of firms in more developed regions. This is because the typical information opacity of SMEs is worsened here by the unsatisfactory development of financial markets. Moreover our analysis ascertains that the small firms can significantly relax the constraints if they establish a close relationship with the banks, making it easier for the banks to access the firm's information.

Keywords: firm growth, financial constraints, relationship lending.

JEL Classification: E22, G31, G32.

1. Introduction

It is generally accepted that the availability of finance is one of the main factors affecting the ability of firms to grow. Especially in small and young firms, growth appears constrained by the quantity of internally generated resources. Furthermore, where financial markets are poorly developed, the gap between outside and inside finance widens since firms find it more difficult to access outside finance. The obstacles to firms tapping outside finance can be partly overcome by improving bank access to the company's information, hence establishing a close relationship.

Our study aims not only to ascertain that the growth of small firms in backward regions is more financially constrained by inside finance than that of firms in more developed regions, but also that close relationships between firms and banks raise the ability of firms to finance their growth with outside resources. Our analysis falls within the field of investment literature which is known to deal with problems related the financing of firm growth and the effects of financial constraints as measured by the investment/cash flow relationship. Excess sensitivity of investment expenditures to cash flow means that a firm's growth is financially constrained because it is strictly dependent on the ability to generate internal resources for its own financing.

For our investigation we apply dynamic panel data techniques to an extended Euler investment equation. We are interested in analyzing the issues related to sensitivity of the investment expenditures to cash flow with reference to the first half of the last decade. Nevertheless, our analysis goes back to the previous years in order to introduce some cyclical evaluations and to capture some dynamics. Therefore, our dataset is formed out of two distinct samples of Italian small and medium sized enterprises obtained from surveys taken by the Italian manufacturing firms and published by Italian private banks in accordance with the Italian Ministry of Industry. The *dataset* for the period 1998-2006 is divided into two balanced samples referring to the six-years' period 1998-2003 and 2001-2006. In this period, the Italian banking system emerged from major reorganization managed by the monetary authority. This process explicitly aimed to make it possible for the national banks to face the increasing competition from the European banks following the integration of EU financial markets.

The consolidation of the national banking system occurred through an upsizing process and M&A operations to the detriment of regional banks in southern Italy. This caused the full disappearance of the large regional banks which had previously played a key role in supporting the growth of the backward regions in southern Italy. This reorganization led to an improvement in efficiency but it ended up with the financial requirements for regional growth depending on smaller local banks.

This period coincided with the end of the positive trend of sustained export-driven growth. There is an inversion of the cycle, leading to a sharp fall in exports and production. This is when the real financial difficulties started for the Italian SMEs, prior to the explosion of the financial crisis in the years to come.

The remainder of the paper is organized as follows. In section 2, we review the main questions about the nexus between financial constraints, relationship lending and firm growth. Section 3 is devoted to explaining the methodological approach and the empirical data. In section 4 after discussing some features both of the capital structure and of relationship banking of the sample firms, we then present the findings of the econometric exercises. Some concluding remarks are contained in the final section. Finally, there is an appendix showing Euler's investment equation.

2. Financial constraints, relationship lending and firm growth

The literature on the capital structure-growth nexus deals with the problems about the sources of financing the production and the financial constraints of the firm's growth. Our analysis is concerned not only with the framework of the Modigliani-Miller propositions according to which capital structure does not matter for a firm's growth but also with the hierarchy hypothesis according to which external finance is not a perfect substitute for internal finance and inside resources are preferred in order to finance firm growth.

Starting with the seminal work by Fazzari *et al.* (1988), this literature has sought to ascertain whether there is a positive relationship between a firm's investment expenditures and its cash flow. According to this approach, great sensitivity of a firm's investment to inside finance indicates that there are financial constraints to the firm's growth. Much was later written to confirm this relationship, estimating empirical models where the investment function is adjusted by proxies of the capital structure, especially by cash flow variables[1]. Since a positive relationship between investment spending and cash flow can prove the existence both of financial constraints or of good opportunities for firm's growth, most analyses have confirmed this relationship, showing that the investment spending in the sample of the firms classified as *ex-ante* "financially constrained" is more sensitive to cash flow (amongst others, Devereux, Schiantarelli, 1989; Hoshi *et al.*, 1991, Oliner, Rudebusch, 1992; Schaller, 1993; Himmelberg, Petersen, 1994; Gilchrist, Himmelberg, 1995; Fazzari *et al.*, 2000)[2].

This approach has been strongly questioned by the latest empirical research beginning with that of Kaplan and Zingales (1997). According to this analysis, the relationship between investment and cash flow does not necessarily prove that financial constraints are binding. On the contrary, capital expenditure will be systematically sensitive to cash flow because the user cost of outside finance is always higher. Therefore, sensitivity to cash flow will be higher for "financially non-constrained" firms than for the financially constrained because the former hold larger internal resources. Other works have confirmed this conclusion (amongst others, Kadapakkam *et al.* 1998; Cleary, 1999; Kaplan, Zingales, 2000; Gomes, 2001; Ati, 2003; Moyen, 2004).

The matter of the "linearity" of the investment-cash flow relationship is largely unresolved[3]. Nevertheless there are many reasons suggesting that small firms face higher financial constraints because the opacity of the relationship of the firm with the financial markets raises difficulties to access outside resources. Many empirical studies have confirmed this thesis, showing that the growth of small firms is more sensitive to inside finance compared to larger firms (amongst others, Oliner Rudebusch, 1992; Westhead, Storey, 1997; Cress, Olofsson, 1997; Audresch, Elston, 2002). By the same token it can be said that the bottlenecks of the resource flows devoted to finance growth are highly likely both if the firm's performance is negative and if the financial markets are not fully developed. Consequently, the dependence of a firm's growth upon inside finance becomes even

[1] The paper by Hubbard (1998) reviews this literature.

[2] In order to distinguish between ex-ante constrained and non-constrained firms several dummies for financial decisions have been used. In the initial works involving samples of larger firms the criterion discriminating between constrained and unconstrained firms lies in the dividend policy. The choice is inspired by finance theory according to which dividend payments are subordinated to investment policy; consequently firms with good opportunities distribute low dividends in order to finance their investments if they are financially constrained compared to firms with large funds and paying higher dividends.

[3] The controversy continues. In the last work by Fazzari *et al.* (2000) the conclusions proposed by Kaplan and Zingales are contested; they say that the sample used in the analysis by Kaplan and Zingales is too small. In reply the latter (Kaplan, Zingales, 2000) restate their arguments and recall that the results of Cleary's work (1999) are obtained using a larger sample and are consistent with their theory.

stronger (Bagella *et al.*, 2001; Becchetti, Trovato, 2002; Bond *et al.*, 2003; Sarno, 2005, 2008; Oliveira, Fortunato, 2006; Becchetti *et al.*, 2009).

There appears to be a broad consensus that relationship lending is the best practice to relax financial constraints. Since the work of Diamond (1984), greater benefits which mitigate the information asymmetries have been assigned to the relationship lending approach as opposed to transaction-based lending. Through relationship lending the bank establishes a long-term relationship with the firm: the bank now gains access to information about the firm while the firm enjoys better access to outside financial resources.

A large flow of information about the firm arises from its utilizing a wide range of financial services offered by the bank. This information cannot be observed by, or transferred to, other banks, and the bank granting exclusive loans to the firm becomes the exclusive owner of such information. The free rider problems deriving from the public nature of the information are avoided and it follows that the bank will bear all the risks and at the same time will gain the benefits arising from its financial decisions. The close relationship enables the bank to support the growth of the firm with regard to its financial needs while, for the firm, benefits arising from the relationship generally consist in an increase in credit availability (Petersen, Rajan, 1994; Berger, Udell, 1995; Cole, 1998; Boot, 2000) or a decrease in the interest rates and the collateral (Petersen, Rajan, 1994, Berger, Udell, 1998). Furthermore, the relationship ensures greater flexibility in the bank's function as an intermediary that can subsidize the firm in adverse events and can be reimbursed in favorable years (Greenbaum *et al.*, 1989; Boot, Thakor, 1994).

From relationship lending there may also arise some disadvantages. The firm can be informationally "captured" by the bank (the so-called hold-up problem)[4]. The exclusive relationship involves monopolistic power by the bank. It can exploit this power by charging increasing interest rates on new loans or rationing additional borrowing. In this regard it can be shown that relationships with more than one bank can reduce its monopoly power (Von Thadden, 1995; Ongena, Smith, 2000) and also ensure greater availability of outside financial resources when there is a credit squeeze (Detragiache *et al.*, 2000).

The conclusions of the empirical analysis are ambiguous. The net gain of relationship lending seems to arise for the firm when the benefits of the informational advantage are not completely balanced by the costs of exploiting monopoly power. Recent theoretical developments suggest that the efficiency of the relationship is strictly dependent on bank competition as well (Boot, Thakor, 2000; Dinç, 2000). According to this analysis, competitive pressure in the local credit market drives the bank to use financing relationships strategically to exploit information advantages. By contrast, the incumbent banks are unable to preserve their position when this informational advantage is unimportant and the profitability of the incumbents is diminishing (Hauswald, Marquez, 2006; Zarutskie, 2006).

3. Data and methodological approach

The data for the empirical analysis were obtained from the surveys of Italian manufacturing SMEs. In the past such surveys were carried out every three years and conducted through both interviews and balance sheet data. The surveys contain information related to several sections with regard to company employment, R&D expenditures, innovations and investment, internationalization, markets and finance; balance sheet data comprise both reclassified revenue statements and asset and liability statements.

The firm sample for the survey is representative of the size and geographical composition of the universe of manufacturing firms. It consists of firms with more than 10 employees and is stratified by the productivity index (value added per employee). The total sample is defined according to the Neyman formula with reference to the individual strata, and the cross-industry composition is determined in proportion to the universe[5].

We defined two balanced firm samples related to the two six-year periods 2001 - 2006 and 1998 - 2003. They contain the observations referring to the same firms in the adjacent surveys, which mean the 1998 - 2000 and 2001 - 2003 surveys and 2001 - 2003 and 2004 - 2006 surveys, respectively. Next, we dropped the large

---

[4] For the purpose of this analysis the opposite case in which the bank is captured by the firm (so-called *soft budget constraint problem*) does not appear relevant to us.

[5] The surveys were carried out by the research centers of various public and private banks and supported by the Italian Ministry of Industry. The first surveys were conducted by the public bank Mediocredito Centrale, the last survey by the private Unicredit Bank. Although some changes were made, the surveys retained their usual structure. In the last survey (2003-2006) the sample was extended to the service industry.

firms with more than 250 employees. Consequently, the two closed samples were formed by SMEs with 10-250 employees and include 1134 observations for 1998-2003 and 823 observations for 2001-2006[6].

We then identified the firms belonging to the backward regions of Italy, on the one hand, and the firms with a close relationship with the main bank, on the other. In the former case, we established that firms operating in backward regions were those with plants in southern Italy, the so-called *Mezzogiorno* (MEZ)[7]. This area is known to have significantly lower overall development conditions compared with the more advanced developed regions in northern and central Italy. There were 135 such firms (11.9% of total observations) in the 1998-2003 sample, and 99 (12.0% of total observations) in the 2001-2006 sample.

Next, we identified the firms with a stable relationship (STAB) with the main bank. In this regard, we utilized information from the "Finance" section of the surveys. We consider firms with a close banking relationship those that show with reference to both surveys:

- a debt share with the main bank equal to or greater than 30% of total debt;
- a relationship with the main bank dating back 15 years or more[8].

According to these criteria, 188 firms had a stable relationship with the main bank (16.6% of total observations) among the former sample, and 141 (17.1% of total observations) among the latter. Thus the information from the surveys as much as the balance sheet data was used to create the dataset for econometric analysis.

We provided estimates in accordance with Arellano-Bond's Dynamic Panel Data method (DPD) which is able to ensure a satisfactory solution to the endogeneity problem arising from the correlation between the fixed effects and the independent variables. This method involves the transformation of all the variables into first order differences in order to drop the fixed effects. Next, it suggests application of the Generalized Method of Moment (GMM) and inclusion of valid instruments for every moment. Therefore the transformed variables are not generally correlated with the fixed effect starting at time $t=2$ (if the start time is $t=0$). From this time the lagged values can be used as instrumental variables for the GMM estimate.

The choice of the empirical model with which to verify the investment-cash flow relationship characterizes the different approaches. Since a positive relationship can be interpreted as evidence of good opportunities for the firm, initial analyses resort to Tobin's Q theory. These empirical specifications suggest controlling for the opportunities through the Q ratio and hence estimating the standard relationship between capital expenditures and the Q measure augmented by cash flow variables. Thus excess sensitivity of investment spending to cash flow indicates that more funds from inside resources are made available for investment when the firm is unlikely to make provision for its own needs from outside finance. Many objections can be raised against this approach; for example, if Tobin's Q is not a good proxy of a firm's opportunities, then excess sensitivity of capital expenditure to cash flow does not necessarily indicate that financial constraints are binding[9].

The alternative approach is proposed by the Euler equation that is the first order condition of the optimization problem of the inter-temporal income flows of the firm. This approach is able to prevent many questions arising from the Q approach because it is founded on the hypothesis of perfect functioning of the capital market. When this hypothesis fails to hold, then the imperfections in the capital markets arise and the firms face financial constraints. Many of the cited works choose the approach of the Euler equation[10].

[6] The upper limit corresponds with the employment criterion fixed by the EU for SMEs. However, the definition of SME by European statistics is more complex because it also considers levels of sales.

[7] The Mezzogiorno comprises the southern area of Italy and is formed by eight administrative regions: Abruzzo, Molise, Campania, Calabria, Puglia, Basilicata, Sicily and Sardinia.

[8] We alternatively proved a higher share of the main bank's debt equal to or greater than 50% of total debt, but the econometric results are unchanged.

[9] According to Gilchrist and Himmelberg (1995), there are three reasons giving rise to skepticism: a) Tobin's Q contains less information about the younger, smaller and fastest growing firms because the markets are unable to gather information; b) if Q is not varying among firms, then the investment-cash flow relationship can result in a higher sensitivity to firm's revenues rather than the existence of financial constraints; c) the relationship should result in a swifter reaction of the younger and smaller firms with regard to variations in investment opportunities (Gilchrist, Himmelberg, 1995; pp. 544-545).

[10] Many objections can also be raised against the Euler equation approach. For example, it is unable to compare results from different studies because it is a reduced form model. Moreover, the estimates are excessively sensitive to the empirical specification of the model, particularly for samples of smaller firms. Finally, it imposes restrictions for every period and fails to consider that financially unconstrained firms today can be constrained tomorrow.

4. Empirical model

For the empirical analysis we follow the Bond-Meghir model (Bond, Meghir, 1994). This model involves Euler's investment function arising from the dynamic optimization of the present value of the expected net earnings function with the symmetric squared adjustment cost. The net earnings function is constrained by the capital accumulation function. If the condition of perfect competition holds, the constrained optimization function means that we can write Euler's investment equation without financial constraints. Empirical estimation of the investment function makes it possible to test the *ex ante* conditions in order to ascertain whether there are constraints; if the ex-ante conditions do not hold, then it cannot be excluded that financial constraints are binding (see the APPENDIX).

From Euler's investment equation the following empirical version of the investment equation can be derived:

$$(I/K)_{t,i} = \beta_1(I/K)_{t-1,i} + \beta_2(I/K)^2_{t-1,i} + \beta_3(CF/K)_{t-1,i} + \beta_4(Y/K)_{t-1,i} + \beta_5(D/K)^2_{t-1,i} + d_t + \eta_i + u_{t,i}$$

where I is the investment expenditure, K the capital stock, CF the cash flow, Y the sales, D the total debt, d and η the time and individual effects, respectively, and u the stochastic term.

We use this version of the empirical equation to investigate two different questions. First, we are interested in mapping the regional differences concerning the financing of investment expenditures. In this regard we assume that between the investment functions of firms in the various regions there are no technological differences except for the financing composition of their expenditure. We will test our hypothesis that the contribution of inside finance for firms in backward southern regions is greater compared to firms based in other Italian regions. In order to capture this effect we introduce in the previous empirical model an interaction variable between the cash flow and the dummy MEZZ; this latter variable has a value of 1 for southern Italian firms and is equal to 0 otherwise.

Second, we will verify the hypothesis according to which the presence of close relationships of the firm with the main bank significantly relaxes financially binding constraints. In order to test the difference with regard to the sensibility of investment spending to cash flow, we introduce into the previous empirical model an interaction variable between the cash flow variable and the dummy STAB. The latter variable is able to distinguish the firms with relationship banking from other firms according to previously set criteria; it assumes a value of 1 when firms have close relationships with their main bank, and 0 otherwise.

Table 1 contains the main features of the sample firms with reference to size and profitability, capital structure and a number of key factors characterizing relationship banking. They are presented so as to outline the localization and the relationship with banks of the sample firms. With regard to the former (see columns [a] and [b]), it can be appreciated that firms in the *Mezzogiorno* are smaller and less profitable than those elsewhere in Italy. Moreover, they have higher debt both on sales and on assets, a higher share of short-term maturity, but debt share on equity is lower. Finally, they are younger and hence their relationship with the main bank is less old than other Italian firms in spite of the fact that both the debt share of the main bank and the multiple relationships are not different. With reference to the latter (see columns [c] and [d]), it can be seen that firms enjoying stable relationships with the main bank are smaller and older. Their ROE is no different compared to other firms besides the ROI is significantly lower while the weight of the debt in the capital structure is systematically greater. Finally, the scenario depicted for the two three-year periods confirms that the positive trend of export-driven production of Italian firms is close to an end and that this has reduced the weight of the debt, presumably raising the flows of inside resources available to finance both firms' current activities and investments.

Hence the estimating equation can be represented as follows:

$$(I/K)_{t,i} = \beta_1(I/K)_{t-1,i} + \beta_2(I/K)^2_{t-1,i} + \beta_3(CF/K)_{t-1,i} + \beta_4(Y/K)_{t-1,i} + \beta_5(D/K)^2_{t-1,i} + \delta MEZZ^*(CF/K)_{t-1,i} + \gamma STAB^*(CF/K)_{t-1,I} + d_t + \eta_i + u_{t,i}$$

The variables are expressed as logarithms and are determined as follows: gross fixed investment (I) is obtained directly from the surveys[11]; capital stock (K) is equal to annual net fixed assets; cash flow (CF) is calculated as the sum of gross earnings and the depreciation of the fixed assets; net sales (Y) is equal to net revenue; total debt (D) is equal to the sum of annual short-term liabilities and medium-long term liabilities.

[11] In most studies the investment data are obtained as "library value"; instead, we use data on the fixed investment obtained directly from the inquiries. The interviews indicate the amount of net fixed investment (plant and equipment, hardware and software, information and innovation technology) in the years of the survey.

Table 1. Mean features of sample firms (mean of median values)

	1998-2003 years				2001-2006 years			
	[a] Italian Firms	[b] *Mezzogiorno* Firms	[c] Firms with relationship	[d] Firms without relationship	[a] Italian firms	[b] *Mezzogiorno* firms	[c] Firms with relationship	[d] Firms without relationship
FIRM SIZE								
sales ('000 €)	13,2	13,4	11,4	14,1	14,6	13,6	12,5	15,5
employment (units)	82,2	67,5	77,0	84,5	80,3	73,2	78,3	82,5
FIRM PROFITABILITY								
return on equity	15,6	5,7	12,6	16,2	11,2	3,8	6,4	13,0
return on investment	5,1	3,8	5,4	5,1	4,7	3,0	4,3	4,7
FIRM CAPITAL STRUCTURE								
Total debt on sales (%)	24,2	31,1	25,8	23,8	19,9	29,5	24,9	18,3
bank debt on total debt (%)	77,9	81,2	81,4	77,4	83,6	83,5	86,3	82,6
Short-term bank debt on total debt (%)	50,0	53,3	54,1	48,8	49,1	52,6	53,8	47,4
Total debt on equity (%)	392,3	354,9	424,0	384,4	267,6	252,6	322,5	256,4
Total debt on assets (%)	112,4	95,5	122,1	107,8	74,2	87,7	101,5	67,8
FIRM-BANK RELATIONSHIP								
Age (years)	28	18	31	27	31	21	34	30
Main bank debt on Total debt (%)	20	20	40	10	20	20	40	10
Number of banks (units)	5	5	6	5	5	5	6	5
Age of main bank relationship (years)	20	15	25	20	20	15	25	20
Lenght of branch from head main bank (KM)	5	9	6	4	5	9	6	4

Notes: Statistics are expressed as mean values of the individual median values related to two six years. In the column [a] there are the statistics of the firms of the whole sample; in the column [b] there are the statistics of *Mezzogiorno*'s firms. In the column [c] there are the statistics related to the firm that have a close relationship with the main bank while in column [c] there are statistic of the firms that do not have a close relationship with the banks. The *Mezzogiorno* is the area of South Italy formed by the administrative regions of Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia and Sardegna. All the regions with the exception of Abruzzo are in the area classified as Objective 1 and are benefiting of economic development and cohesion policies of the UE.

Table 2. Statistics and correlation matrices

2003-1998 YEARS						
STATISTICS						
		Average	S.D.	Min.	Max.	
(I_K)		0,9	1,3	0,0	4,6	
$(I\_K)^2$		2,5	4,8	0,0	21,2	
(CF_K)		3,6	1,5	0,0	17,0	
(Y_K)		6,5	1,1	0,0	11,5	
$(D\_K)^2$		37,0	13,0	0,0	121,0	
CORRELATION MATRIX						
		(I_K)	$(I\_K)^2$	(CF_K)	(Y_K)	$(D\_K)^2$
(I_K)		1,000				
$(I\_K)^2$		0,968	1,000			
(CF_K)		0,047	0,074	1,000		
(Y_K)		0,005	0,042	0,544	1,000	
$(D\_K)^2$		-0,030	0,006	0,414	0,901	1,000
2006-2001 YEARS						
STATISTICS						
		Average	S.D.	Min.	Max.	
(I_K)		9,3	11,2	0,0	32,9	
$(I\_K)^2$		211,4	353,8	0,0	1085,7	
(CF_K)		42,9	41,7	0,0	133,3	
(Y_K)		611,2	446,5	163,3	1549,2	
$(D\_K)^2$		142424,4	182044,7	9569,4	577864,8	
CORRELATION MATRIX						
		(I_K)	$(I\_K)^2$	(CF_K)	(Y_K)	$(D\_K)^2$
(I_K)		1,000				
$(I\_K)^2$		0,959	1,000			
(CF_K)		0,204	0,215	1,000		
(Y_K)		0,147	0,182	0,620	1,000	
$(D\_K)^2$		0,067	0,102	0,397	0,826	1,000

The estimates are obtained following the Arellano-Bond method for dynamic panels according to the first difference variables are involved as instrument of the GMM estimates[12]. In the empirical model dummy variables are introduced for temporal effects.

In Table 2 statistics and correlation matrices of both six-year samples are shown. Estimates of the investment equations are contained in Table 3. Besides the F test, for each equation the Sargan test and the AR tests are reported. The former test of over-identification verifies that the instrument number is not excessive. The latter tests investigate the autocorrelation between the independent variable and the fixed effects; in this regard it

[12] For the estimate we use the STATA 8.0 package that contains an opposite procedure for the dynamic panel estimate according to the Arellano-Bond method.

may be expected that the AR[1] test does not exclude the presence of autocorrelation, while the contrary should hold for the AR[2] test. These tests always meet expectations.

Columns [1] and [3] show the estimates of the investment equations related to overall firms for the periods 1998-2003 and 2001-2006, respectively. The conditions of the models generally hold, but the former equation performs better than the latter. The coefficients of the lagged dependent variable are positive, whereas the coefficients of its lagged value are negative and lower than unity[13]. The coefficients related to the debt variables are negative and lower than unity; they are compatible with the hypothesis of the presence of taxes and distress costs in the former equation, but they are not significant in the latter equation. The coefficients of sales variables are positive; the accelerating effect is higher in the former equation and dramatically lower in the latter. Both these outcomes indicate the prevailing trend toward the deceleration of the debt weight and the rise of inside resource devoted to finance firm growth due to sustained performance of the previous years. At the same time, this improvement in the capital structure of the firm is counterbalanced by a negative trend of production caused by intensified competitive pressure on international markets and by a dramatic decrease in national exports.

Finally, the cash flow coefficients are negative with reference to all the sample firms, indicating that financial constraints are not important. Columns [2] and [5] report the estimates of the investment equations including the interaction cash flow variable MEZZ*(CF/K) devoted to capturing the impact of internal finance on capital expenditures in southern Italian firms. The estimates confirm the previous results because the coefficient remains negative; it may be noted that while the variations of the other coefficients are not significant, the cash flow value is significantly higher. Furthermore, the cash–flow coefficient becomes positive and higher when it refers to the interaction variable of southern firm observations. Since this variable measures the difference of the impact of inside finance on investment in southern firms, it may be stated that while the coefficients referring to the overall sample are negative, equal to -0.17 and -2.81 for the two six-year periods respectively, the one referring to the southern firms is positive and approximately equal to +0.21 and +2.54, respectively in 1998-2003 and 2001-2006. This indicates excess sensitivity of investment expenditures to cash flow for southern firms[14].

Finally, the estimates including the interaction variable related to the presence of relationship banking STAB*(CF/K) are reported in columns [3] and [6]. As can be seen, while the previous results are confirmed with reference to the remaining coefficients, the cash flow coefficients related to the overall firm become positive and they remain significant: the values are equal to +0.14 for the former six-year period coefficient and +0.12 for the latter six-year coefficient. However, this excess sensitivity of investment expenditure is mitigated by the effect of relationship banking. The coefficients related to the interaction variable are negative; the value in the former equation is -1.17, that of the latter equation -1.04. The net effect is approximately the same in both cases. Therefore, relationship lending appears to relax the pressure on inside resources, improving the liquidity conditions of the firms. According to our empirical outcomes it can be said that the benefits of the relationship tend to significantly overcome the effects of the financial constraints arising from the typical information opacity of the SMEs.

Table 3. Investments-*cash flow* relationship estimates

	1998-2003 years			2001-2006 years		
	[1]	[2]	[3]	[4]	[5]	[6]
$(I / K)_{-1}$	0,755	0,778	0,781	0,316	0,326	0,289
	[.191]***	[.193]***	[.205]***	[.060]***	[.073]***	[.088]***
$(I / K)^2_{-1}$	-0,128	-0,136	-0,134	-0,006	-0,007	-0,006
	[.047]***	[.048]***	[.051]***	[.002]***	[-.002]***	[.003]**
$(CF / K)_{-1}$	-0,063	-0,145	0,141	-0,018	-0,071	0,121
	[.032]**	[.048]***	[.072]**	[.008]**	[-.017]***	[.039]***
$(Y / K)_{-1}$	0,469	0,486	0,483	0,003	0,005	0,006

[13] The coefficient of the lagged dependent variable in the first equation is no different from unity (Student's t is equal at -1,2), but the same coefficient in the second equation related to 2001-2006 years is significantly different from unity (Student's t is -19.2).

[14] The southern firm coefficient is the overall sample coefficient and the interaction variable coefficient.

	1998-2003 years			2001-2006 years		
	[1]	[2]	[3]	[4]	[5]	[6]
	[.212]**	[.214]**	[.227]**	[.001]**	[.002]***	[.002]**
$(D / K)^2_{-1}$	-0,048	-0,048	-0,050	-0,000	-0,000	-0,000
	[.019]**	[.019]**	[.020]**	[.000]	[.000]	[.000]
$EZZ*(CF / K)_{-1}$		0,327			1,116	
		[.140]**			[.178]***	
$TAB*(CF / K)_{-1}$			-1,169			-1,044
			[.366]***			[.173]***
F	262,4***	301,2***	267,2***	22,7***	18,3***	12.5***
Sargan Test	191,4***	183,2***	156,8***	2924,9***	1940,7***	1029,3***
AR(1)	-15,9***	-15,6***	-14,7***	-18,4***	-9,36***	-6.8***
AR(2)	1,4	0,7	-1,3	-1,3	-0,9	-1.4
n° Obs	4188	4188	4188	2868	2868	2868

Notes: Variables are: I=Investment, K=Capital Stock, CF=*cash flow*, Y=sales and D=Total Financial Debt; MEZZ is dummy with unity value if the firm is belonging in *Mezzogiorno*'s regions and zero value otherwise; STAB is the dummy with unity value if the firm is involving a closed relationship with main bank and zero value otherwise. The estimates are obtained through the GMM method; standard errors are in brackets; *,**,*** are indicating statistical significance of the coefficients at 90%, 95% and 99%, respectively.

Conclusions

It is commonly believed that SMEs face greater obstacles in obtaining the necessary outside resources to finance their growth. This paper provided some further evidence in this regard with reference to a sample of Italian SMEs. We adopted the well-known approach according to which the sensitivity of the investment expenditures to inside finance may suggest that financial constraints to company growth are binding. We ascertained through the estimate of Euler's investment equation the sensitivity of investment on firm's cash flow variables and we also investigated two related questions: to what extent financial constraints are more binding for firms in backward regions and then whether relationship lending can significantly mitigate their effects.

Our conclusions are as follows. First, with reference to our sample firms there is no confirmation from econometric analysis that investment expenditures show excess sensitivity to company cash flow. Second, we found sensitivity of investment to inside finance instead for firms in the backward regions of southern Italy. In this case there is no sound reason which justifies the ambiguity in the economic literature as regards interpretation of such sensitivity, namely that it can prove the existence of both financial constraints and good investment opportunities. In this regard we provided evidence that firms in backward regions perform less well, are financially weaker and therefore have fewer opportunities compared to firms in more developed regions. Our empirical analysis confirmed that in backward regions a firm's growth is constrained by inside financial resources, or that it is more dependent on inside finance than elsewhere. Finally, we proved that the information advantages arising from relationship lending considerably relax asymmetry effects and significantly mitigate financial constraints, reducing the sensitivity of investment expenditure to cash flow. This empirical outcome appears consistent with the events related to the latest financial crisis in which small local banks enjoying close relationships with firms played an important role in the SME growth.

References

[1] Alti, A. 2003. How sensitive is investment to cash flow when financing is frictionless? *Journal of Finance*, 58: 707-722.

[2] Audretsch, D.B., and Elston, J.A. 2002. Does firm size matter? Evidence on the impacts of liquidity constraints on firm investment behavior in Germany*, International Journal of Industrial Organization,* 20: 1-17.

[3] Bagella, M., Becchetti, L., and Caggese, A. 2001. Financial constraints on investment: a three-pillar approach based on a priori identification, *Research in Economics,* 55: 219-254.

[4] Becchetti, L., Castelli, A., and Hasan, I. 2009. Investment - cash flow sensitivities, credit rationing and financing constraints, *Small Business Economics,* 35: 467-497.

[5] Berger, A.N., and Udell G.F. 1995. Relationship Lending and Lines of Credit in Small Firm Finance, *Journal of Business,* 68: 351-381

[6] Berger, A.N., and Udell, G.F. 1998. The economics of small business finance: the role of private equity and debt markets in financial growth cycle, *Journal of Banking and Finance*, 22: 613-673.

[7] Bianco, M. 1991. Vincoli finanziari e scelte reali delle imprese italiane: gli effetti di una relazione stabile con una banca, in Angeloni I., Conti V., Passacantando F. (a cura di), "*Le banche ed il finanziamento delle imprese*", Bologna, Il Mulino: 23-60.

[8] Bond, S., and Meghir, S. 1994. Dynamic investment models and the firm's financial policy, *Review of Economic Studies,* 61: 197-222.

[9] Bond, S., Elston, J., Mairesse, J., and Mulkay, B. 2003. "Financial factors and investment in Belgium, France, Germany and UK: a comparison using company panel data*, Review of Economics and Statistics,* 85: 153-165.

[10] Boot A.W.A. 2000. Relationship banking: what do we know, *Journal of Financial Intermediation,* 9: 7-25.

[11] Boot A.W.A, and Thakor, A.V. 1994. Moral hazard and secured lending in an infinitely repeated credit market game, *International Economic Review,* 35: 899-920.

[12] Boot, A.W.A, and Thakor, A.V. 2000. Can relationship banking survive competition?, *Journal of Finance,* 2: 679-713.

[13] Clearly, S. 1999. The relationship between investment and financial status, *Journal of Finance* 54: 673-692.

[14] Cole, R. 1998. The importance of relationships to the availability of credit, *Journal of Banking and Finance,* 22: 959-977.

[15] Cressy, R., and Olofsson C. (eds). 1997. European SME financing: an overview, Special Issue of Small Business Economics, *Small Business Economics,* 9: 87-96.

[16] Detragiache, E., Garella, P., and Giuso, L. 2000. Multiple vs. Single Banking Relationships: Theory and Evidence, *Journal of Finance,* 55: 1133-1161.

[17] Devereux, M., and Schiantarelli, F. 1990. Investment, financial factors and cash flow: evidence from U.K. panel data in HUBBARD G.R. (ed.) – *Asymmetric Information, Corporate Finance and Investment*, University of Chicago Press, Chicago.

[18] Dinç, S.I. 2000. Bank reputation, bank commitment and the effects on competition in credit markets, *Review of Financial Studies,* 13: 781-812.

[19] Fazzari, M.S., Hubbard, G.R., and Petersen, B.C. 1988. Financing constraints and capital investment, *Brooking Papers of Economic Activity*, 1998: 141-195.

[20] Fazzari, M.S., Hubbard, G.R., and Petersen, B.C. 2000. Investment cash flow sensitivities are useful: a comment on Kaplan and Zingales, *Quarterly Journal of Economics,* 115: 695-705.

[21] Gilchrist S., Himmelberg C.P. (1995). Evidence on the role of cash flow in reduced - form investment equations, *Journal of Monetary Economics*, 36: 541-572;

[22] Giunta, A., and Sarno, D. 2009. Firm's financing and industrial structure in the less developed regions of the south Italy, *Journal of Applied Economic Science,* 4(10): 509-525.

[23] Gomes, J.F. 2001. Financing investment, *American Economic Review,* 91: 1263-1285.

[24] Greenbaum, S., Kanatas, G., and Venezia, I. 1989. Equilibrium Loan Pricing under the Bank-Client Relationship, *Journal of Banking and Finance,* 13: 221-235.

[25] Hauswald, R., and Marquez, R. 2006. Competition and Strategic Information Acquisition in Credit Markets, *Review of Financial Studies,* 19: 967-1000.

[26] Himmelberg, C.P., and Petersen, B. 1994. R & D and Internal Finance. A Panel Study of Small Firms in High Tech Industries, *Review of Economics and Statistics,* 76: 38-51.

[27] Hoshi, T., Kashyap, A., and Scharfstein, D. 1991. Corporate Structure, Liquidity and Investment: Evidence from Japanese Panel Data, *Quarterly Journal of Economics,* 106: 33-60.

[28] Hubbard, G.R. 1998. Capital-Market Imperfections and Investment, *Journal of Economic Literature,* 36: 193-225.

[29] Kapapakkam, P.-R., Kumar, P.C., and Riddick L.A. 1998. The Impact of Cash Flows and Firm Size on Investment: the International Evidence, *Journal of Banking and Finance,* 22: 293-320.

[30] Kaplan, S.N., and Zingales, L. 1997. Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints, *Quarterly Journal of Economics,* 112: 169-215.

[31] Kaplan, S.N., and Zingales, L. 2000. Investment-Cash Flow Sensitivities are not Measures of Financial Constraints, *Quarterly Journal of Economics,* 115: 707-712.

[32] Moyen, N. 2004. Investment-Cash Flow Sensitivities: Constrained versus Unconstrained Firms, *Journal of Finance,* 59: 2061-2092.

[33] Oliner, S., and Rudebusch, G. 1992. Sources of the Financing Hierarchy for Business Investment, *Review of Economics and Statistics,* 74: 643-654.

[34] Oliveira, B., and Fortunato, A. 2006. Firm Growth and Liquidity Constraints: a Dynamic Analysis, *Small Business Economics,* 27: 139-156.

[35] Ongena, S., and Smith, D. 2000. What Determines the Number of Bank Relationships? Cross Country Evidence, *Journal of Financial Intermediation* 9: 26-56.

[36] Petersen, M., and Rajan, R. 1994. The Benefits of Lending Relationships: Evidence from Small Business Data, *Journal of Finance*, 49: 3-37.

[37] Sarno, D. 2005. Liquidity Constraint on the Production of Firms in Southern Italy, *Small Business Economics,* 25: 133-146.

[38] Sarno, D. 2008. Capital Structure and Growth of the Firms in the Backward Regions of the South Italy, *Applied Financial Economics,* 18: 293-305.

[39] Sarno, D. 2008. Financial and Legal Constraints to Firm Growth: the Case of Italy, *Journal of Applied Economic Sciences,* 3: 293-305.

[40] Schaller, H. 1993. Asymmetric Information, Liquidity Constraints and Canadian Investment", *Canadian Journal of Economics,* 26: 552-574.

[41] Von Thadden, E.L. 1995. Long-Term Contracts, Short-Term Investment and Monitoring, *The Review of Economic Studies,* 62: 557-575.

[42] Westhead, P, and Storey, D.J. 1997. Financial Constraints on the Growth of High Technology Small Firms in the United Kingdom, *Applied Financial Economy,* 7: 197-201.

[43] Zurutskie, R. 2006. Evidence of the Effects of Bank Competition on Firm Borrowing and Investment, *Journal of Financial Economics,* 81: 503-537.

APPENDIX

Bond-Meghir's model can be represented by an investment function derived from the dynamic optimization of the present value of the expected cash flow with symmetric squared adjustment cost function. The constrained optimization problem is

$$Max\ E_t \left[\sum_{j=0}^{\infty} \beta_{t+j}\, \Pi_t(\cdot) \right]$$

where E is the conditional expectation on the information available at time t, β is the nominal discount factor between t and $t+j$ and Π is net earnings[15].

The constraint is represented by the capital accumulation function

$K_{t+1} = (1-\delta)K_t + I_t$. *where* K is capital stock, δ depreciation rate and I investment.

The function of net earnings is

$$\Pi_t = p_t F(K_t,L_t) - p_t 1/2bK_t [(I/K)_t - c]^2 - w_t L_t - p^I_t I_t$$

where L is the labour input, p^I is the capital price, p is the output price, w is the labour price. $F(K_t,L_t)$ is the production function with constant returns to scale while $1/2bK_t[(I/K)_t\text{-}c]^2$ indicates the adjustment cost function that is linearly homogeneous in K and L variables. The hypothesis of perfect competition conditions holds; therefore p (the price of the firm's output) is dependent on the output through demand elasticity $(\varepsilon>1)$, that is constant.

The derivatives of the previous equation with respect to I and to K are, respectively,

$(\partial\Pi_t/\partial I_t) = -bap_t (I/K)_t + bcap_t - p^I_t$ and $(\partial\delta\Pi_t/\partial K_t) = ap_t (Y/K)_t - ap_t (\partial\Pi_t/\partial K_t) (L/K)_t + bap_t (I/K)_t^2 - bcap_t (I/K)_t$

where $Y=F\text{–}G$ is the net sales of the firm while $a=[1-(1/\varepsilon)]>0$[16].

The Euler equation without financial constraints is

$$(I/K)_{t+1}=c(1-\varphi_{t+1})+(1+c)\varphi_{t+1}(I/K)_t-\varphi_{t+1}(I/K)_t^2-\varphi_{t+1}/ba(CF/K)_t+\varphi_{t+1}/baJ_t+ \varphi_{t+1}/b(\varepsilon-1)\ (Y/K)_t - [(1+r_t)v_t/b(1-\delta)a]\ (D/K)_t^2 + v_{t+1}$$

where, $\varphi_{t+1}=(1+\rho_{t+1})/(1-\delta)$, $(1+\rho_{t+1})=(1+r_t)(p_t/p_{t+1})$, ρ_{t+1} is the real discount rate, $(CF/K)_t=(p_tY_t\text{–}w_tL_t)/(p_t K_t)$ is the ratio between the real *cash flow* and the capital stock, $J_t=(p^I_t/p_t)\ \{1-p^I_{t+1}(1-\delta)/[(1+r_t)\ p^I_t]\}$ is the user cost capital, $(D/K)_t^2=(p^I_t/p_{t+1})[(D_t/p^I_tK)_t^2$ is the debt ratio while v_{t+1} is the error term.

We assume that the real discount rate $[\varphi_{t+1}]$, net sales and the debt ratio coefficients are constant over time and across firms, and are therefore parameters.

The hypothesis of the model is satisfied if it happens *ex post* that the forward investment rate coefficient $[(1+c)\varphi_{t+1}]$ is greater than or equal to 1, the forward squared investment rate coefficient $[-\varphi_{t+1}]$ is negative and lower than 1 in absolute value.

Moreover, it may happen that the coefficient $[\varphi_{t+1}/ba]$, that is the same for the forward *cash flow* rate and for the user cost of the capital, is negative; the forward net sales coefficient $[\varphi_{t+1} / b(\varepsilon -1)]$ is positive (or equal to 0 if the perfect competition hypothesis holds). The expectation about the sign of the debt variable coefficient $[v_t]$ is not certain: if the Modigliani-Miller propositions hold, then it is equal to 0, while in the opposite case, it is positive. However, if there are taxes and bankruptcy costs, then it is negative.

The empirical specification of the investment function is as follows

$$(I/K)_{t,i} =\beta_1(I/K)_{t-1,i}+\beta_2(I/K)^2_{t-1,i}+\beta_3(CF/K)_{t-1,i}+\beta_4(Y/K)_{t-1,i}+\beta_5(D/K)^2_{t-1,i}+d_t+\eta_i+u_{t,i}$$

where d_t are time effects, η_i individual effects and $u_{t,i}$ is a stochastic time term for individual observations.

---

[15] β is equal to $1/(1+r)$ where r is the expected yield. The operator E is the expectation conditional on the information available at initial period t; the expectation is related to the interest rate, the input and output prices and the technology.

[16] The derivative on K is based on the assumption that Y_t is homogeneous on (K_t,L_t). Moreover, the labor marginal productivity $(\partial F/\partial L)$ can be substituted by the first order condition (w/ap); this allows us to avoid specifying the parametric form of the production function.

ASERS Publishing

Volume III Issue 2(6) Winter 2012

DOI: 10.2478/v10261-012-0006-1

SUSTAINABILITY AND HUMAN DEVELOPMENT: A PROPOSAL FOR A SUSTAINABILITY ADJUSTED HUMAN DEVELOPMENT INDEX

José PINEDA
Human Development Report Office
United Nations Development Program (UNDP), **United States**
jose.pineda@undp.org

Abstract:

This paper proposes sustainability - adjusted human development index (SHDI) in which countries' achievements in human development are penalized if there is over-exploitation of the environment. The human development approach has been a powerful framework in the past for advancing the measurement of human progress, particularly the human development index (HDI), which is a capabilities index aiming to capture to what extent people have the freedom to live substantively different lives. Today, this approach can help us make more explicit the profound connections between current and future generations' choices by offering a framework for understanding sustainability that connects inter- and intra-generational equity with global justice. The empirical analysis shows that there are important global sustainability challenges ahead since there are 90 (out of 185) countries with per capita CO_2 emissions above the planetary boundaries. There are 19 countries that lose at least one position in their HDI ranking after adjusting for sustainability. Between these countries, however, the countries that experienced the largest drop in ranking were 102 positions for the United States, 39 positions for China, and 22 positions for the Russian Federation.

Keywords: sustainability, HDI, human development.

JEL Classification: O1, O15, O5, Q5.

1. Introduction

The HDI, produced by the Human Development Report Office of UNDP, has contributed to global discussions to best measure human progress. Since its inception, it was recognized that the concept of human development is larger than what can be measured by the index. This creates policy challenges, since there may be situations in which HDI progress masks deterioration in other key aspects.

The evidence presented by Hughes et al (2012) suggests that, if no action is taken, the current and future environmental threats could jeopardize the extraordinary progress experienced in the HDI in recent decades. Moreover, projection-scenarios exercises done by Hughes et al (2012) suggest that, in an adverse "environmental disaster" scenario —envisioning vast deforestation and land degradation, dramatic declines in biodiversity and accelerated extreme weather events— the global HDI would be at least 15 percent below the projected baseline. Consequently, if no measures are taken to halt or reverse current trends, the environmental disaster scenario could lead to a turning point before 2050 in developing countries—their convergence with rich countries in HDI achievements begins to reverse.

The idea of this paper is to propose a sustainability-adjusted HDI (from now on SHDI) in which countries' achievements in human development are penalized, to reflect the over-exploitation of the environment and its relative intensity.

2. What can we learn from trends in measures of sustainability?

2.1. Aggregate measures

There is an on-going conceptual debate on how to define sustainability —mostly grouped either under weak sustainability or strong— which have implications for the measurement and assessment of sustainability trends. The main difference between both concepts of sustainability is that weak allows for substitutability across all forms of capital, while strong acknowledges that sustainability requires preserving so-called critical forms of natural capital (Neumayer, 2011). In fact, depending on which concept of sustainability is adopted the loss of the natural environment can be compensated by increased levels of other forms of capital, physical capital for example. This conceptual debate also makes it difficult to have a broadly acceptable quantitative measure of sustainability. Here we review some of the aggregate measures that are most in use, but for a comprehensive review of sustainability measures and indicators see Jha, and Pereira (2011).

Green national accounting is an approach that adjusts measures such as gross domestic product or savings for environmental degradation and resource depletion. This has been done under the System of Environmental-Economic Accounts (SEEA) framework, which contains the internationally agreed standard concepts, definitions, classifications, accounting rules and tables for producing internationally comparable statistics on the environment and its relationship with the economy.

One important aggregate measure under this category is the World Bank's Adjusted Net Savings (ANS), also known as Genuine Savings, which takes the rate of savings, adds education spending and subtracts for the depletion of energy, minerals and forests as well as for damage from carbon dioxide emissions and pollution. Based on the theory developed in Hamilton and Clemens (1999), the ANS aims to measure the change in present and future well-being, by showing the true rate of savings in an economy after taking into account how the economy invests and consumes all its assets (human, natural and man-made). The measure could be used as an indicator of future consumption possibilities. Ferreira, Hamilton and Vincent (2008) use a panel data for 64 countries (1970-82) and empirically show a significant positive correlation –after adjusting by population growth- between past per capita genuine savings and future changes in per capita consumption. This measure is consistent with the weak sustainability framework, since it implies that the different kinds of capital are perfect substitutes, so that financial savings, for example, can replace a loss of natural resources or lower human capital.

The Adjusted-Net Savings measure has been criticized by many authors like Neumayer (2004, 2010, 2011), mainly because of the human capital investment and of the natural capital depreciation measures. The human capital investment (measured by current education expenditures) has been argued to be probably overestimated, because human capital is lost when individuals die. Also, health does not enter the calculus, which, according to Dasgupta (2007), makes the notion of human capital used inadequate.

The depreciation of natural capital from extraction of natural resources is calculated as the price of the resource minus the average cost of extraction (as an approximation of the marginal cost) times the resource extraction volume. According to Neumayer (2010), there are preferable methods to compute the natural resource rents, like the one described in El Serafy (1981), which includes future capital gains when valuing the depreciation of exhaustible resources. Neumayer (2010) argues that this method is preferable to the one used by the World Bank, mainly because it does not depend on the assumption of efficient resource pricing; it takes into account the country's reserves of natural resources, so that a given extraction volume has different implications for sustainability depending on the total stock available. For example, valuing natural resources at market prices can overestimate the sustainability of an economy that produces them as the resources become scarcer and thus more expensive. For more detailed discussion see Teignier-Baqué (2010). Nonetheless, Hamilton and Ruta (2009) show that the approach by El Serafy is likely to lead to artificially to low asset values and therefore to low values for the depletion of the assets, resulting in an over-estimation of the social welfare (higher ANS).

The CO_2 emission damages are valued at US$20 per metric ton of carbon in the ANS, following Frankhauser (1995). This, according to Dasgupta (2007) and others, is clearly an underestimate of the actual damage. The UNDP's Human Development Report 2007-2008, for instance, considers that an adequate carbon price would be on the range US$60-100, and the Stern Report concludes that is above US$100. As Frankhauser (1994) admits, the US$20 per metric ton of carbon value is only a rough order-of-magnitude assessment of the actual marginal costs of greenhouse gas emissions, and "care should be exercised when interpreting the figures". Tol (2008) reviews a number of studies and shows that many of them find higher costs than Frankhauser (1995).

This is particularly problematic given the uncertainty embodied in the measurement of greenhouse gas emissions and their monetary valuations. For instance, Garcia and Pineda (2011) using Tol (2008) meta-analysis showed that the number of countries considered unsustainable using adjusted net savings in 2005 would rise from 15 to 25 if we use a more comprehensive measure of emissions that includes methane and nitrous oxide as well as carbon dioxide and acknowledged monetary valuation uncertainties.

Two examples under the strong sustainability framework are the Ecological Footprint (EFP) - a measure of the annual stress people put on the biosphere— and the Environmental Performance Index.

As Neumayer (2011) explains, the carbon emissions constitute the main element in the Ecological Footprint of many countries, and in fact there is a strong and statistically significant cross-country correlation (0.85) between the per capita volume of carbon emissions and the value of the EFP. Van den Berth and Verbruggen (1999), argue that the conversion of consumption categories into land area is incomplete and that it uses a set of weights which do not necessarily correspond to social weights because they do not reflect scarcity changes. Other problems, they argued, are that the EFP denotes land area as something that is hypothetical, since the world's EFP can exceed the world's total available productive land. According to Neumayer (2011) another important objection related to the energy or carbon footprint, which constitutes the main component of the EF for many countries, is that there are much less land-intensive ways of sequestering or avoiding carbon emission from burning fuels than (hypothetical) reforestation.

From all of the aggregate measures of sustainability, only two are available for a large number of countries over a relatively long period of time: the World Bank's Adjusted Net Savings and the Global Footprint Network's Ecological Footprint. Another more recent measure is the Environmental Performance Index, developed at Yale and Columbia Universities. The EPI measures environmental performance using a set of policy targets, which are based on international treaties and agreements, standard developed by international organizations and national governments, the scientific literature and expert opinion. This composite index uses 25 indicators to establish how close countries are to established environmental policy goals — a useful policy tool, built from a rich set of indicators and providing a broad definition of sustainability. But the measure's data intensity (requiring 25 indicators for more than 160 countries) inhibits construction of a time series so we will exclude it from the analysis of trends. Another important limitation of the EPI for international comparison is that some of its data is modeled.

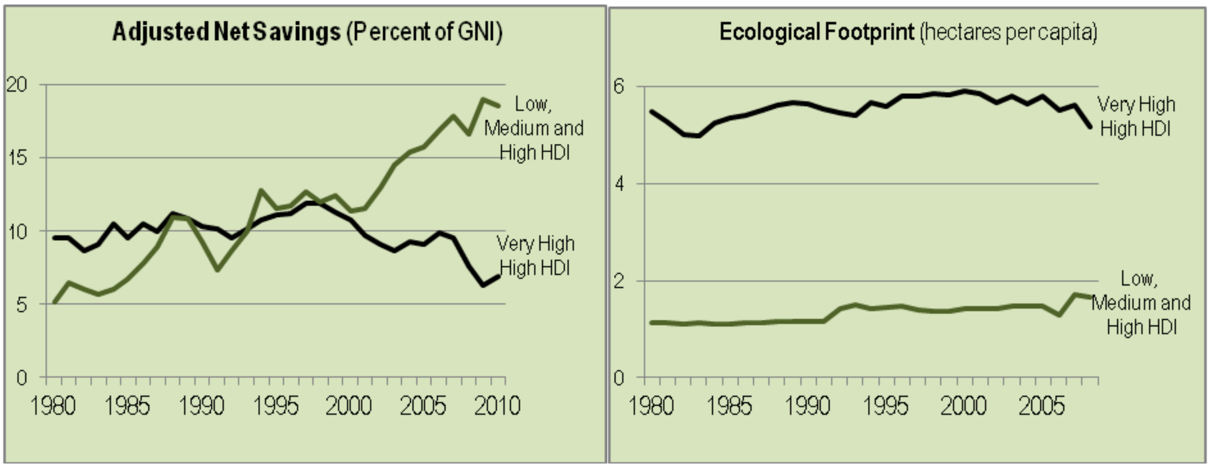


Figure 1. Aggregate measures of sustainability Adjusted net savings and ecological footprint.

Source: Based on data from World Bank (2012), Ecological Footprint Network and own calculations.

As we can see from Figure 1, the Adjusted Net Savings measure is positive for all groups according to the HDI, which means that the world is (weakly) sustainable. However, while the trend for low, medium and high HDI countries suggests that their sustainability (measured by this indicator) has improved over time, the trend of the very high HDI countries is declining.

In contrast, the sustainability trend that emerges from the ecological footprint shows that the world is increasingly exceeding its global capacity to provide resources and to absorb wastes. Given the calculations presented in the 2011 HDR, if everyone in the world had the same consumption level as people in very high HDI countries, with the current technologies, we would need more than three Earths to withstand the pressure on the environment. Current patterns of consumption and production are unsustainable at the global level and imbalanced regionally. And the situation is worsening, especially in very high HDI countries.

2.2. Specific indicators

Patterns of carbon dioxide emissions over time constitute a good, although imperfect, proxy for the environmental impacts of a country's economic activity on climate. Evidence from the 2011 HDR showed that emissions per capita are much greater in very high HDI countries than in low, medium and high HDI countries combined. It also showed that there are significant differences across groups with different HDI achievements. Today, the average person in a very high HDI country accounts for more than four times the carbon dioxide emissions and about twice the emissions of the other important greenhouse gases (methane, nitrous oxide) than a person in a low, medium or high HDI country.

Results from the 2011 HDR also showed a strong positive association between the level of HDI (especially its income component) and carbon dioxide emissions per capita. This positive relationship was also found in terms of changes over time. Countries with faster HDI improvements have also experienced a faster increase in carbon dioxide emissions per capita. This hints at the fact that the recent progress in the HDI has been associated with higher emissions putting at risk its sustainability. The discussion about the relationship between the environmental threats due to carbon dioxide emissions and achievements in human development should take into account a historical perspective, since the stock of carbon dioxide trapped in the atmosphere is a product of historical emissions. Today's concentrations are largely the accumulation of developed countries' past emissions. With about a sixth of the world's population, very high HDI countries emitted almost two-thirds (64 percent) of carbon dioxide emissions between 1850 and 2005, with the United States representing about 30 percent of total accumulated emissions.

Climate change - with effects on temperatures, precipitations, sea levels and vulnerability to natural disasters - is not the only environmental problem. Degraded land, forests and marine ecosystems pose chronic threats to well-being, while pollution has substantial costs that appear to rise and then fall with increasing levels of development. The 2011 HDR showed that nearly 40 percent of global land is degraded due to soil erosion, reduced fertility and overgrazing. Between 1990 and 2010 Latin America and the Caribbean and Sub-Saharan Africa experienced the greatest forest losses, while desertification threatens the dry-lands that are home to about a third of the world's people. Some areas are particularly vulnerable - notably Sub-Saharan Africa.

The 2011 HDR also showed that since 1970, global carbon dioxide emissions have increased 248 percent in low, medium and high HDI countries and 42 percent in very high HDI countries. The global growth of 112 percent can be broken down into three drivers: population growth, rising consumption and carbon-intensive production. Rising consumption (as reflected by GDP growth) has been the main driver, accounting for 91 percent of the change in emissions, while population growth contributed 79 percent. The contribution of carbon intensity, in contrast, was a reduction of 70 percent, reflecting technological advances. Hence, when added the individual contributions we are able to explain the 100 percent of the total growth, and results show to forces inducing more emission and only one force reducing it. In other words, the principal driver of increases in emissions is that more people are consuming more goods - even if production itself has become more efficient, on average. Although the carbon efficiency of production has improved by 40 percent, total carbon dioxide emissions continue to rise. Average carbon dioxide emissions per capita have grown 17 percent over 1970–2007.

Patterns of carbon dioxide emissions vary widely across regions and stages of development. While very high HDI countries account for the largest share of world carbon dioxide emissions, low, medium and high HDI countries account for more than three-fourths of the growth in carbon dioxide emissions since 1970. East Asia and the Pacific is the largest contributor by far to the increase in these emissions (45 percent), while Sub-Saharan Africa contributed only 3 percent, and Europe and Central Asia, 2 percent. We have data for a shorter period for methane and nitrous oxide, but in these cases too, the contribution of the East Asia and the Pacific region is particularly pronounced. Trade enables countries to shift the carbon content of the goods they consume to the trading partners that produce them. Several countries that have committed to cutting their own emissions are net carbon importers, including Germany and Japan, as are countries that have not signed or ratified global treaties, such as the United States.

In a recent study Peters *et al.* (2011) examined the "virtual carbon trade" flows, by defining a country's carbon consumption as the difference between the tons of greenhouse gases it emits ("carbon production") and the net carbon content of its imports and exports. Their estimates highlight a sizeable transfer of carbon from the poor world to the rich world", so the authors argue that "the rich world has been 'off shoring' or 'outsourcing' its emissions" to developing countries. However, divergences between the production and consumption of carbon cannot be ascribed solely to the "outsourcing" of carbon-intensive production from developed to developing economies. Relatively large carbon exports largely reflect countries' natural resource endowments, rather than a "leakage" of carbon-intensive manufacturing away from developed economies. Furthermore, the virtual carbon trade data suggests that carbon- and energy-exporting countries are also more likely to allow domestic energy prices to lag behind world energy prices, in order to subsidize domestic energy consumption resulting in lower levels of energy efficiency.

3. Incorporating sustainability into the measurement of human development

3.1. Existing alternatives

UNDP's Human Development Index is one of the most prominent indicators of well-being. However, the HDI does not take into account sustainability variables in a broader sense. Recent academic work has mainly focused on examining the potential for 'greening' the HDI so as to include environmental and resource-consumption dimensions. These works have yielded various proposals for extending the HDI to take sustainability and environmental aspects into account.

Shreyasi Jha (2009) proposed modifying the income dimension of the HDI which reflects the use of natural resources by using a more inclusive measure of wealth per capita, that includes natural capital. In this regard, the author proposes three viable alternatives: replace GDP with Net National Production; use World Bank's Total Wealth indicator; or replace GDP with a measure for Green Net National Product.

De la Vega and Urrutia (2001), on the other hand, present a pollution-sensitive human development index. This indicator incorporates an environmental factor, measured in terms of CO2 emissions from industrial processes per capita with the standard measure of human development. This composite measure penalizes the income component by taking into account the environmental costs arising from such output.

Morse (2003) proposes an environmentally sensible HDI, equal to the sum of the HDI plus the integral environmental indicator, which is the average of an indicator of the environmental state of country and an indicator of the environmental evaluation of human activities. The author emphasizes that any greening of the HDI should make sure that the basic HDI remains unmodified.

Constantini (2005) proposes to calculate a composite Sustainable Human Development Index as the simple average of the four development components: education attainment, social stability, sustainable access to resources (Green Net National Product), and environmental quality.

Other efforts include Dewan (2009) Sustainable Human Development (SHD) – in which the developmental goal is to achieve higher human development for the maximum number of people in present and future generations. Dahme *et al.* (1998) Sustainable Human Development Index -an extension for the HDI which is produced by using total material requirement- sums all material inputs (a-biotic raw materials, biotic raw materials, moved soils, water and air) required to produce a country's national output. Ramanathan (1999). Environment Sensitive HDI -a product of HDI and Environment Endangerment Index (EEI) - is computed with data on deforestation, number of rare, endangered or threatened species, a greenhouse gas emissions index and a chlorofluorocarbon emissions index.

3.2. Sustainability adjusted human developed index (SHDI)

The capability approach sees human life as a set of 'doings and beings' or 'functionings', which are constitutive of a people's being, and an evaluation of a person's well-being has to take the form of an assessment of these constituent elements. In this approach, a functioning is an achievement of a person: what he or she manages to do or to be, while a capability reflects the various combinations of functionings he or she can achieve, reflecting his or her freedom to choose between different ways of living.

This conceptual approach is very different from a utilitarian approach, since the later may fail to reflect a person's real deprivation, which is not the case for the capability approach. For example, a thoroughly deprived person might not appear to be badly off in terms of the mental metric of utility, if the hardship is accepted with no-grumbling resignation, even though he or she may be quite unable to be adequately nourished, decently clothed, and minimally educated and so on.

The HDI, an index which tries to capture capabilities, is conceptually different from a social welfare function. The key difference is that a social welfare function is designed to be maximized, while a capabilities index is meant to give a measure of the extent to which people in different countries have accesses to substantively different lives. In this sense, the capabilities approach contrasts with traditional theories of social justice, such as utilitarianism, which postulate the maximization of utility as the final goal of human action. The capabilities approach is a partial theory of well-being, which does not ambition to establish a complete description of the entire components of a good life. Instead, a capabilities index aims to tell us the extent to which people have the freedom to live substantively different lives.

Neumayer (2004) stated that sustainability is the requirement to maintain the capacity to provide non-declining well-being over time. Sustainability, unlike well-being, is a future-oriented concept. Hence, he suggested that it is better to use separate indicators to trace these two concepts and not one. We propose an approach for which indicators are calculated separately for each country, and later, based our conceptual approach which connects present and future choices; they are combined on our Sustainability Adjusted HDI.

In the Annex 2 we present tables and graphical analysis of the relationship between 6 sustainability indicators, 2 aggregate (ANS and EFP) and 4 specific indicators (per capita CO2, per capita fresh water withdrawals, percentage of extinct species over total and percentage of land with permanent crops), and the HDI.

3.2.1. Linking present and future choices

Today, we are facing an increasing need for improvements in the measurement of human progress that would not only capture the scope of the choices available to the current generation but also the sustainability of these choices. In other words, we need a measure that is able to connect present choices to future choices. Sen (2009) argues the need to achieve "sustainable freedom", which implies the preservation of human's freedom and capabilities today without "comprising capabilities of future generations to have similar or even more freedom". As was already mentioned, the basic purpose of development is to enlarge people's choices. However, as Anand and Sen (2000) explain, the basic idea of human development involves equal rights applied to all. Universalism considers unacceptable any form of discrimination based on class, gender, race, community, and also generation. A more utilitarian view can be found in Roemer (2009), who proposes that an ethically attractive approach to sustainability is one in which today we choose a consumption path that maximizes the level of the worst-off generation. The justification, he argues, is that since the birth date of a person is arbitrary, no generation should be better off than any other unless it comes without lowering the utility of the worst-off generation. This implies that future generations should receive the same kind of attention than the current generation. The same idea can be found in the Human Development Report 1994: "There is no tension between human development and sustainable development. Both are based on the universalism of life claims".

Drawing upon the Universalist principle, people should not only care about the choices that are open to them (as measured by the HDI), but also about how they were procured, and their impact on the choices available to future generations globally. An index capturing capability should focus on the measurement of human achievement and freedom in a reflective – rather than mechanical – way Sen, A. (1989). As part of this reflective process invaded in the index, achievements today should also be valued taking into account its potential impact on future generations.

Thus, progress in human development achieved at the cost of the next generations should be viewed less favorably than progress achieved in a sustainable way. It is critical that this connection is fully integrated into the analysis and measurement of human progress. One of the main dimensions affecting the connection between the choices of current and future generations is the environment, but not the only one. For example, the savings and investment decisions of current generations will affect the possibilities for command over resources by the next generations; it is also well known that parents' education has a significant positive impact on the likelihood of their children being more educated, healthier, and with a future higher command over resources. Parents have an enormous influence on their children's education for several reasons, but most importantly because they are their children's first teachers (Gratz, 2006). They also affect children aspirations, since children with more highly educated parents developed higher aspirations for their own education and on average attained more education by age 19, which in turn related to higher levels of adult educational attainment Dubow *et al.* (2009). However, as we will see later in this paper, the existence of global sustainability thresholds and externalities (within and between generations), generates a particular relevance for environmental considerations when we explicitly connect present and future generation's choices.

3.2.2. National and global sustainability and the existence of tipping points

The previous analysis implies that inter-generational equity should be measured in a way that goes beyond national borders. When measuring progress at the country level, we should care about the potential negative effect of current generation's actions on the possibilities available to future generations globally.

For the analysis of sustainability it is crucial to distinguish between the local, national and global dimensions. Measures of global sustainability examine the aggregate, although the effects of policies may vary greatly by location not only between countries but within countries as well. For example, as Dasgupta (2009) discusses, the world's poorest people often have no substitutes when their local resource base is degraded, so even if they live in a country considered sustainable, the conditions in which these disenfranchised groups live may not be. While recognizing that the local level is essential in the human development approach as well as for policy-making, the present analysis focuses on the global level owing to the pressing need to find a measurement tool that integrates both inter-generational and global equity. Most of existing aggregate measures of sustainability, as already discussed on section 2 of this paper, typically lack of this integrated framework; since they mostly focus on the country level, without taking into account the complexity of the global challenges that we are facing on this shared planet. They also tend to focus only on adjusting economic or environmental indicators in ways that do not necessarily reflect non-linearities and tipping points, and which assume near-perfect substitutability of all types of capital or not substitutability at all.

Given the need of a general framework in which the concept of human development could be enhanced in a shared planet -not only today but tomorrow- we take a global perspective of sustainability, aiming to capture up to what extent our current life style is compromising future generations' human development. It is important also to clarify that our vision is not presented as necessarily contradictory with any other particular view of sustainability, but rather as an approach that is closer and more coherent with the human development paradigm.

The impact of a particular country to the global sustainability of the earth can be measured by taking into account the relative damage that the country's actions impose on the whole world, or, in other words by including the externalities of such country's action. Most existing approaches to sustainability, particularly those that use resource accounting such as the Adjusted Net Saving, have a country focus which does not allow them to internalize the global implications of countries' behavior. In fact, such an approach does not analyze the reasons why a particular country is depleting its assets, nor does it take into account that it is as important to sustain the stock of capital as how to (globally) sustain it (Neumayer (2011, 2010)). The human development approach is a better guidance of what is important to sustain and how it should be sustained, by putting people at the centre of the analysis now and in the future through the lens of the "universalist" principle.

There is an increasing consensus about the seriousness of the threats that humanity is facing in terms of global sustainability. As the Report of the United Nations Secretary-General's High-level Panel on Global Sustainability emphasized, awareness is growing on the fact that there is an increased danger of surpassing "tipping points" beyond which environmental changes accelerate, and become self-perpetuating, making it difficult or even impossible to reverse. The existence of these threats supports a vision of non-substitutability across all forms of capital, as the strong sustainability approach argues with respect to the role natural capital plays in absorbing pollution and providing direct utility in the form of environmental amenities (Neumayer, 2010). They also support a vision in which a global perspective of sustainability is taken into consideration and not just the sustainability of individual countries in isolation.

This analysis aims for a greater integration of science into all levels of policymaking on sustainable development, as it has been the call from the Report of the United Nations Secretary-General's High-level Panel on Global Sustainability. The analysis of planetary boundaries developed by Rockström *et al.* (2009) is an important example of scientific work in this field. This approach argues that the anthropogenic pressures on the Earth System have reached a scale where abrupt global environmental changes can no longer be excluded. It proposes an approach to global sustainability based on definitions of planetary boundaries within which humanity can be expected to live safely. Transgressing one or more of these (nine) planetary boundaries may be deleterious or even catastrophic due to the risk of crossing thresholds that will trigger non-linear, abrupt environmental change within continental- to planetary-scale systems. The Millennium Ecosystem Assessment and the Intergovernmental Panel on Climate Change (IPCC) are also important references that assess environmental challenges on human well-being based current knowledge, scientific literature, and data.

3.3. The loss function

In our analysis, we use a pragmatic approach between a single composite indicator and a dash-board. Indicators of sustainability are calculated separately for each country and then integrated into a single indicator, but the interpretation can be easily decomposed. The indicators to be used should preferably reflect the planetary boundaries that have been identified, for which given the current scientific understanding, there are seven quantifications: climate change; ocean acidification; stratospheric ozone; biogeochemical nitrogen cycle and phosphorus cycle; global freshwater use; land system change; and the rate at which biological diversity is lost. The two additional planetary boundaries for which they have not yet been able to determine a boundary level are chemical pollution and atmospheric aerosol loading. Because of data limitations in terms of country coverage but also time coverage, there are only a few areas for which environmental indicators with implications for global sustainability can potentially be identified at the national level for a large number of countries over time, namely carbon dioxide emissions, land use for permanent crops and fresh water withdrawals. We aim at identifying those countries that are exceeding the "*threshold*" or planetary boundary needed to achieve sustainability. As better and more comprehensive data is available other areas of sustainability can be integrated to the analysis. However, not all these indicators can easily be linked to global sustainability by just looking at their national values. This highlights the difficulties to connect the global and local dimensions in our evaluation of sustainability.

The following table illustrates the case for fresh water withdrawals, in which we have many countries exceeding the global threshold that are within their local boundary. For example, Canada has large water resources while Kuwait is water constrained, while the first is exceeding the global threshold and not its local threshold, the second is experiencing the opposite. Also, we can see how the United Arab Emirates is locally constrained by water availability and uses 20.32% of its own water resources, while using 56% of their water global threshold. A similar case could be made for land usage, particularly given the lack of information on the quality of the land that is used.

Given the challenges in the use of indicators related to water and land usage, we will focus the calculations of the loss function on the use of CO_2 emissions, for which data is relatively of good quality, it is collected regularly as a time series, and its connections to global sustainability are better understood in the literature. Of course, CO_2 is one out of many GHGs, but a very important one for which there is data for most countries and for many years.

Table 1. Countries exceeding local/global thresholds of fresh water withdrawals

Country	water per capita usage as % of global threshold	% water/own resources	Country	water per capita usage as % of global threshold	% water/own resources
Turkmenistan	544.2	100.8	Kuwait	35.7	2465
Guyana	244.4	0.7	United Arab Emirates	55.7	2032
Uzbekistan	239.8	118.3	Saudi Arabia	93.1	943.3
Kazakhstan	227.9	28.9	Libyan	75.5	718
Iraq	221.2	87.3	Qatar	25.9	455.2
Kyrgyzstan	209	43.7	Bahrain	29.7	219.8
Tajikistan	190.8	74.8	Yemen	15.7	168.6
United States	171.1	15.6	Egypt	91.9	119
Estonia	151.4	14	Uzbekistan	239.8	118.3
Canada	149.7	1.6	Israel	28.7	101.9

Source: UNDP and World Bank.

The thresholds are taken from Rockström *et al.* (2009), and Meinshausen *et al.* (2009). For CO_2 total accumulated emissions over the next 50 years likely to keep temperature change within 2°C (886 gigatons a year gives 8-37% probability of exceeding 2°C).

Despite the considerable uncertainty and estimated variance around these thresholds in the scientific community, they are an important point of reference and it is important to do extensive sensitivity analysis including as many indicators and incorporating the uncertainties around these thresholds as much as possible.

In section 4, we present results for the lower bound and upper bound of the thresholds. The tighter threshold will be used for the baseline calculations, while the more relaxed will be presented as part of the sensitivity analysis. The upper bound for CO_2 emissions is 1,437 gigatons accumulation for the next 50 with a 29-70% probability of exceeding 2°C. Note that both thresholds used are calculated by taking the global total CO_2 emissions and divide it by 50 years and the level of total global population.

The environmental variable included to calculate the loss function in the SHDI is not to be thought of as adding an extra dimension to the determination of societal well-being in a country. This point of view is in principle warranted by the very nature of the environmental variable under consideration, since this is not a factor that affects the inhabitants of the country alone, but the planet as a whole.

3.3.1. The loss function: fair share and global responsibility

In order to guide policy action, it is of critical importance to combine the best available evidence provided by science with a sound concept of social justice. The issue of climate change has an important dimension of distributive justice. Nevertheless, since there is not a consensus on which is the most appropriate equity principle; it is necessary to specify which equity criteria is applied. The measure proposed should be consistent with this equity criterion, a point that we will discuss in more detail in this section and in Annex 1.

There is a wide variety of criteria that have been used in the climate change literature, such as egalitarianism - equal use right of the environment for every person - sovereignty - equal use right of the environment at the level of nations-, ability to pay –proportionality of costs according economic well-being- and Rawl's maximin - the welfare of the worst-off country should be maximized-. A more detailed discussion can be found in Rose and Kverndokk (2008). We follow a "Rights" approach by proposing a universally equal or "fair" use of the environment, in which everyone has the same right to use the planet's natural capital and the ecosystem services it generates, subject to constraints imposed by planetary boundary considerations. This point has also been made by authors like Raworth (2012): "*Sustainability cannot be achieved without a necessary degree of fairness and justice. It appears therefore necessary to reconcile the social foundations of fairness with the planetary boundaries of a sustainable world"*.

The way in which we incorporate this "*Rights*" approach is by a proper normalization of the indicators, looking for a combination in which resources are used both fairly and sustainably. We express our relevant sustainability indicator in per capita terms (in this case we use per capita CO_2 emissions), and compare the per capita use of the environment of a citizen in a given country to the per capita threshold or maximum fair share according to the planetary boundary. This indicator enables us to capture situations in which the citizens of a country are having an excessive use of the environment by exceeding their fair share of the planetary boundaries. The important point to signal is that everyone in the planet has the right to achieve higher human development but within the limits imposed by the sustainability of our shared planet.

It is also understood that even though each individual has the same right to a fair use of the environment, country level analysis requires an additional consideration for justice depending on the relative size of the country. We call this global responsibility, and we argue that the country's weight in regards to its behavior on the excessive use of the environment should be higher, the larger its population. By incorporating the global responsibility factor we are able to combine both inter- and intra-generational equity considerations. The fair share of the planetary boundary indicates that every individual has an equal right to the environment, including those of future generations, and this is why our use of the environment should stay within these boundaries. The global responsibility increases with the size of the country with respect to the rest of the world. In this sense, it produces a balance between individual actions and a country's responsibility for the state of global sustainability.

If a country's population is exceeding its fair share of the planetary boundaries, its HDI is affected by a loss function, G^i, which is the multiplication of two components, the fair share and the global responsibility, which captures the potential negative effect of current actions of the citizens of a country on the possibilities available to future generations globally.

The loss function, G^i, is bounded between 0 and 1, for each country. The loss is 0 if the country per capita CO2 emissions are below the fair share, while a country that in isolation exceeds the maximum boundary has a loss of 1. The loss function depends on the whole world's situation (by using the thresholds defined by the planetary boundaries), but it gives a particular value for each country according to its level of per capita emissions and its share of the world's population. Also, when the per capita CO2 emissions of a country increase, all other things equal, the loss for such a country cannot decrease. Finally, in order to maintain comparability across countries, if two countries which are exceeding their fair share increase (reduce) their per capita CO2 emissions in the same amount; the relative value of their loss functions remains constant.

To summarize, we propose a Sustainability Adjusted HDI (SHDI), which imposes a loss function to a country's human development achievements given its degree of unfair use of the environment, according to the planetary boundaries, and its share in the global population as a relative size indicator. This is represented in equation 1, where we showed the SHDI for country i. See annex 1 for a mathematical representation of the SDHI.

$$SHDI^i = \left(1 - G^i\right) * HDI^i \quad (1)$$

3.3.2. Interpretation of SHDI

The standard interpretation of the HDI is that it is a capabilities index, thus intended to be a crude measure the size of the set of capabilities of the inhabitants in a country. The question is, then: what does it mean to apply a loss to the HDI of country i by $(1-G_j^i)$? In other words: How is the SHDI in equation (1) to be interpreted given environmental indicator j and country i?

Individuals in a country not only care about the multidimensional choices that are open to them (as measured by the HDI) but also about how those possibilities were procured and the impact that this will have on the choices of future generations. This implies that people care about inter-generational equity (which will now be captured by the SHDI). Thus, human development achievements at the cost of significantly contributing towards global environmental un-sustainability (and then a significant reduction of the choices available to future generations) are viewed less favorably, by the citizens of that country, than those achieved in a sustainable way. Other things equal, the citizens of a country that is within its fair share of the planetary boundaries (and thus not compromising the possibilities for future generations) have more reason to value their achievements in human development. This is a country whose citizens exhibit a higher degree of attention to inter-generational equity, and the prospects for future generations' human development achievements globally. Finally, after the loss function is calculated, the level of human development (and not just the economic activity) plays a role because the loss is multiplied with the HDI to obtain the SHDI. Thus, for two countries with similar CO2 emissions and similar populations but with different levels of HDI, the absolute penalty level will differ.

4. Results

The following tables show a statistical description of per capita CO_2 emissions that we used for the calculation of the loss function of the SHDI. We show the values for the set of countries in the HDI sample transgressing the planetary boundary (at the lower threshold, 2.66 tons of CO_2 per capita), and a secondary threshold that is the value at the upper boundary in the level of uncertainty (less restricting, 4.29 tons of CO_2 per capita).

As we can see from Table 2, for per capita CO_2 emissions there are 90 countries that transgress the lower threshold and 75 countries that transgress the upper bound out of 185. These results are consistent with the fact that CO_2 emissions are one of the three planetary boundaries that - according to Rockström *et al.* (2009) - humanity has already transgressed (along with biodiversity loss and the nitrogen cycle).

The first two column presents basis statistics related to the "Intensity" of emissions (by how much countries are exceeding the global threshold). The fifth and sixth column presents the statistics for global responsibility, while the last column presents the loss due to un-sustainability for those countries exceeding the global threshold (both the lower bound and upper bound). The 90 countries exceeding the more restrictive threshold (lower bound threshold), do so on average by more than 2.4 times. While the 75 countries exceeding the less restrictive threshold (upper bound threshold), do so on average by more than 1.4 times. These differences are reflected in the last column, where the average and maximum loss are almost twice for the more restrictive threshold. Finally, as we can see for the third and fourth column, countries did not differ much in terms of their share of population between groups, with a few countries with a relatively large share of population but the majority is small countries.

Table 2. Intensity, global responsibility and losses due to un-sustainability for per capita emissions

stats	CO_2 Emissions intensity		CO_2 Emissions global responsibility		CO_2 Emissions losses due to un-sustainability	
	above threshold (CO_2 per capita >4.29) (number of times exceeding threshold)	above threshold (CO_2 per capita >2.66) (number of times exceeding threshold)	above threshold (CO_2 per capita >4.29) (share between 0-1)	above threshold (CO_2 per capita >2.66) (share between 0-1)	above threshold (CO_2 per capita >4.29) (between 0-1)	above threshold (CO_2 per capita >2.66) (between 0-1)
mean	1.38	2.43	0.0064	0.0061	0.00567	0.0112971
s.d.	1.8	2.84	0.0282	0.0258	0.01885	0.0388
min	0.009	0.010	0.000	0.000	0	0
max	10.37	17.4	0.2408	0.2408	0.15004	0.273
N (obs.)	75	90	75	90	75	90

Using this information, we were able to generate SHDI for a total of 185 countries. The analysis shows that even though the correlation between the original HDI and the SHDI is very high (0.99), there are significant changes in ranking for some countries.

The effects of adjusting for sustainability using all indicators are higher for very high and high human development groups, which includes some oil producing countries (as can be seen from figure 2). At the lower boundary, there are 90 (out of 185) countries with per capita CO2 emissions above the planetary boundary (which implies a positive penalty).

There are 3 countries for which the penalty is higher than 5% the United States (27.2%), China (23.9%), and the Russian Federation (7.3%). The largest drop in ranking from our sample of 185 countries were 106 positions for the United States, 397 positions for China, and 22 positions for the Russian Federation. In the following table, we present the list of countries with losses in HDI ranking after adjusting for sustainability.

Table 3. Countries positions lost with SHDI (at the lower boundary)

Country	HDI	SHDI	Loss due to Un-sustainability	Rank HDI	Rank SHDI	Number of position lost
United States	0.9099	0.6616	0.2728	4	106	102
China	0.6871	0.5223	0.2399	100	139	39
Russian Federation	0.7553	0.6997	0.0737	65	87	22
Japan	0.9006	0.8581	0.0472	11	25	14
Germany	0.9051	0.8771	0.0310	8	18	10
Saudi Arabia	0.7704	0.7537	0.0216	55	63	8
Iran (Islamic Republic of)	0.7074	0.6935	0.0197	87	94	7
Canada	0.9081	0.8848	0.0257	6	12	6
Ukraine	0.7292	0.7215	0.0106	75	79	4
Korea (Republic of)	0.8972	0.8787	0.0206	14	17	3
Poland	0.8133	0.8038	0.0117	38	41	3
Malaysia	0.7605	0.7546	0.0078	60	62	2
Turkey	0.6991	0.6953	0.0055	91	93	2
South Africa	0.6194	0.6087	0.0173	122	124	2
Netherlands	0.9099	0.9034	0.0071	3	5	2
Kazakhstan	0.7447	0.7365	0.0110	67	68	1
Mexico	0.7700	0.7620	0.0105	56	57	1
Italy	0.8738	0.8600	0.0159	23	24	1
United Arab Emirates	0.8459	0.8377	0.0098	29	30	1

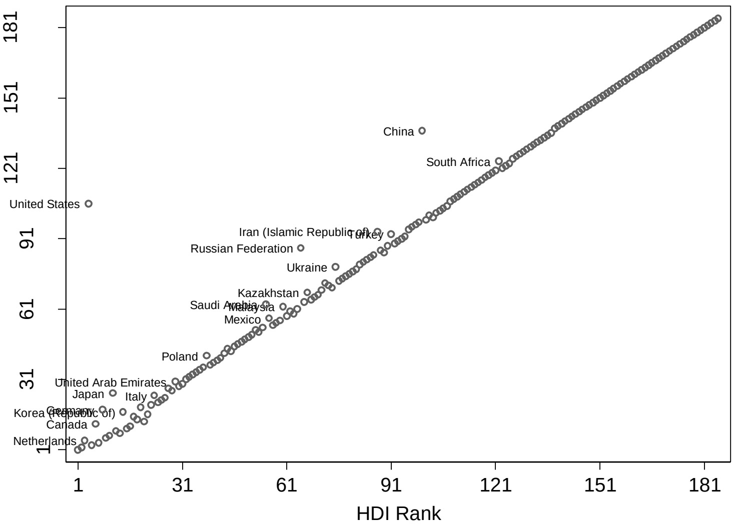


Figure 2. Rank comparison between HDI and SHDI (at the lower boundary)

Our results are based on countries contribution to global sustainability issues; however it is not necessarily the case that the countries that contribute the most to climate change will be the one mostly affected. In fact, as shown in the HDR 2011, the low HDI countries have contributed the least to global climate change, but they have experienced the greatest loss in rainfall and the greatest increase in its variability, with implications significant impact in their human development. These countries are also very likely to experience the largest losses, in terms of lower HDI, in case of an extremely adverse environmental scenario by 2050. When we check the correlation between these expected losses (more from the impact side) and the losses calculated for the SHDI (more from the contribution side), we find it to be low and negative (-0.1057). The country with expected highest loss is

Central African Republic (0.2054), while it has 0 losses for the SHDI. In the opposite situation is the United States, it has a 0.2720 loss for the HDI and a zero expected loss for 2050 due to extreme environmental challenges.

Table 4 presents the results with the less restrictive threshold generating smaller losses, and consequently, smaller variations in rankings for the 75 countries with a positive loss. The largest drop in ranking from our sample of 185 countries were 476 positions for the United States, 12 positions for the Russian Federation, and 9 positions for China and Japan.

Table 4. Countries positions lost with SHDI (at the upper boundary)

Country	HDI	SHDI	Loss due to Un-sustainability	Rank HDI	Rank SHDI	Number of position lost
United States	0.9099	0.7733	0.1500	4	51	47
Russian Federation	0.7553	0.7271	0.0374	65	77	12
Japan	0.9006	0.8808	0.0220	11	20	9
China	0.6871	0.6489	0.0556	100	109	9
Canada	0.9081	0.8955	0.0139	6	10	4
Saudi Arabia	0.7704	0.7613	0.0117	55	59	4
Germany	0.9051	0.8919	0.0145	8	12	4
Malaysia	0.7605	0.7581	0.0032	60	63	3
South Africa	0.6194	0.6145	0.0078	122	125	3
Ukraine	0.7292	0.7263	0.0040	75	78	3
Netherlands	0.9099	0.9067	0.0035	3	5	2
Iran (Islamic Republic of)	0.7074	0.7018	0.0080	87	89	2

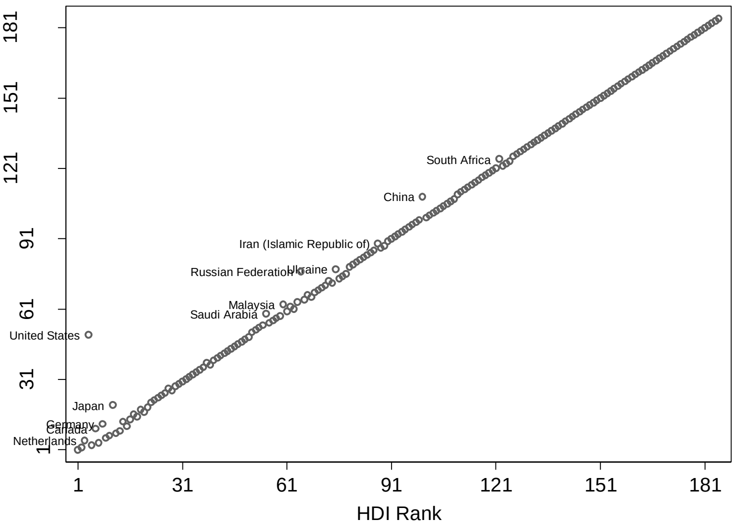


Figure 2. Rank comparison between HDI and SHDI (at the upper boundary)

Conclusions

The current challenges that human progress faces underscore the need to improve our measurement tools. We build upon this in a framework that combines the best available scientific evidence, a human centered development approach, and a social justice criterion in order to connect the choices available to current generations with those that could be available to future generations. The human development approach has been a powerful framework in the past for advancing the measurement of human progress in a multidimensional way. Today, this approach can help us make more explicit the profound connections between current and future generations' choices by offering a framework for understanding sustainability that connects inter- and intra-generational equity with global justice.

This analysis shows that there are important sustainability challenges ahead since there are 90 (out of 185) countries with per capita CO_2 emissions above the planetary boundary (taking into account the more restrictive threshold). There are 19 countries that lost at least one position in the ranking after adjusting for sustainability. Between these countries, however, there are 3 countries for which the penalty is higher than 5%: the United States (27.2%), China (23.9%), and the Russian Federation (7.3%). These countries experience the largest drop in ranking from our sample of 185 countries was 102 positions for the United States, 397 positions for China, and 22 positions for the Russian Federation.

Finally, the relevance of this proposal for a SHDI comes primarily from the fact that it does not try to add more dimensions to the HDI or to use monetary valuations in order to adjust one of its components (mainly income), which has important practical and conceptual limitations, since it does not look at the broader set of capabilities that is captured by the HDI. This approach is not necessarily contradictory with any other particular view of sustainability (in particular those discussed in this paper), but it is closer and more coherent with the human development approach and a capability index such as the HDI.

There are significant data limitations (which must be seriously addressed) in terms of frequency and country coverage, but the results clearly show important policy implications for understanding how to capture sustainability considerations when measuring human development. We particularly consider important the connection between present and future generations within a development framework that is people-centered. We know that this is work in progress and further discussion, both conceptually and empirically (including intensive sensitivity analysis to different functional forms and alternative indicators, in addition to those presented in the annex), will help us to continue the constant search for improving our measures of human progress. So far we consider this to be the starting point of a larger research agenda, but we consider this to be a positive contribution to the broader discussion of sustainability from a human development perspective.

Acknowledgments

This paper has benefited from previous discussions and comments by Khalid Malik, Maurice Kugler, Stephan Klasen, Alan Fuchs, Arthur Minsat, Jingqing Chai, Ricardo Fuentes, Marisol Sanjines, Oscar Garcia, Martin P. Heger, Ramón Pineda and Vanessa Alviarez. I would like to thank Shivani Nayyar for her significant help in the initial part of this work. I am particularly grateful to Eduardo Zambrano and Carmen Herrero for their technical inputs related to the properties of the loss function.

The paper also benefited from the discussion with prominent experts on the sustainability field such us Mathis Wackernagel and Kyle Gracey from Global Footprint Network; Marc Levy from the Center for International Earth Science Information Network (CIESIN); and Kirk Hamilton from the Development Research Group of the World Bank. Finally, I would like to specially thank the excellent research assistance provided by Akmal Abdurazakov, Diana Jimenez, Veronica Postal and Clarissa Santelmo. All errors remain my responsibility. The views expressed herein are those of the author and do not necessarily reflect the views of the United Nations Development Program (UNDP).

References

[1] Anand, S., and Sen, A. 2000. Human Development and Economic Sustainability, World Development, *Elsevier,* 28(12): 2029-2049.

[2] Baqué, M.T. 2010. Integrating Sustainability into the Human Development Framework, *Manuscript UNDP-HDRO*, New York.

[3] Constantini, V. and Monni, S. 2005. Sustainable Human Development for European Countries, *Journal of Human Development*, 6(3): 329-351.

[4] Dahme, K., F. Hinterberger, H. Schutz, and E. K. Seifert. 1998. Sustainable Human Development Index: A Suggestion for Greening the UN's Indicator, *Wuppertal Institute for Climate, Environment and Energy Working Paper.*

[5] Dasgupta, P. 2007. The Idea of Sustainable Development, Sustainable Science, 2: 5-11.

[6] Dasgupta, P. 2009. The Welfare Economic Theory of Green National Accounts, *Environment Resource Economics,* 42: 3-38.

[7] De la Vega, M. C. *et al.* 2001. HDPI: a Framework for Pollution-Sensitive Human Development Indicators, *Environment, Development and Sustainability,* 3: 199-215.

[8] Dewan, H. 2009. Redefining Sustainable Development to Integrate Sustainability and Human Development Goals, *The International Journal of Environmental, Cultural, Economic and Social Sustainability,* 5(4): 147-162.

[9] Dubow, E.F., Boxer, P., and Huesmann, L.R. 2009. Long-term effects of parents' education on children's educational and occupational success: Mediation by family interactions, child aggression, and teenage aspirations. *Merrill-Palmer Quarterly,* 55(3): 224-249.

[10] El Serafi, S. 1981. Absorptive Capacity, the Demand for Revenue, and the Supply of Petroleum, *Journal of Energy and Development,* 7(1): 73-88.

[11] Ferreira, S., Hamilton, K. and Vincent, J.R. 2008. Comprehensive Wealth and Future Consumption: Accounting for Population Growth. *World Bank Econ Rev,* 22(2): 233-248, first published online May 25, 2008.

[12] Frankhause, S. 1994. The Social Costs of Greenhouse Gas Emissions: An Expected Value Approach, *Energy Journal,* 15(2): 157-184.

[13] Frankhauser, S. 1995. *Valuing Climate Change: The Economics of Greenhouse*, London: Earthscan.

[14] Garcia, C., and Pineda, J. 2011. Measuring the Sustainability of Countries: An Uncertainty and Sensitivity Analysis of the Adjusted Net Savings Measure. *Human Development Research Paper 12. UNDP–HDRO*, New York.

[15] Gratz, J. 2006. The Impact of Parents' Background on their Children's Education. *Educational Studies 268: Saving Our Nation, Saving Our Schools: Public Education for Public Good,* 1-12.

[16] Hamilton, K. and Ruta, G. 2009. Wealth accounting, exhaustible resources and social welfare, *Environmental and Resource Economics,* 42:53-64.

[17] Hamilton, K. and Clemens, M. 1999. Genuine savings rates in developing countries. *World Bank Econ Rev,* 13(2): 333-356.

[18] Herrero, C., and Villar, A. 2001. The three musketeers: four classical solutions to bankruptcy problems, *Mathematical Social Sciences,* 42:307-328.

[19] Herrero, C. 2012. On E. Zambrano further thoughts about the proposed SHDI. *Manuscript. Universidad de Alicante, Ivie & IAE.*

[20] Hughes, B.I. *et al.* 2012. Exploring future impacts of environmental constraints on Human Development. *Sustainability,* 4(5): 958-994.

[21] Jha, S. 2009. Environmental Sustainability Measurement: A Critical Review. *UNDP-HDRO. New York.*

[22] Jha, S., and Pereira, I. 2011. Existing Measures of Sustainability: A Review. Background Paper for the 2011 *Human Development Report*. UNDP–HDRO, New York.

[23] Klugman, J., Rodríguez, F., and Hyung - Jin Choi. 2011. The HDI 2010: New Controversies, Old Critiques. Human Development Research Paper 2011/01. *UNDP–HDRO, New York*.

[24] Meinshausen, M. *et al.* 2009. Greenhouse-Gas Emission Targets for Limiting Global Warming to 2°C. *Nature* 458(30): 1158–63.

[25] Morse, S. 2003. Greening the United Nation's Human Development Index, *Sustainable Development*, 11: 183-198.

[26] Neumayer, E. 2011. Sustainability and Inequality in Human Development. *Human Development Research Paper 2011/04.* UNDP–HDRO, New York.

[27] Neumayer, E. 2004. Sustainability and Well-Being Indicators, *United Nations University – World Institute for Development Economics Research,* Research Paper 23/2004.

[28] Neumayer, E. 2010. *Weak versus Strong Sustainability. Exploring the limits of two opposing paradigms*, Cheltenham and Northampton: Edward Elgar Publishing, Third Edition.

[29] Peters, G., Minx, J.C., Weber, C.L., and Edenhofer, O. 2011. Growth in emission transfers via international trade from 1990 to 2008, *Proceedings of the [US] National Academy of Sciences*, 2011.

[30] Ramathan, B. 1999. Environment Sensitive Human Development Index: Issues and Alternatives, *Indian Social Sciences Review*, 1: 193-201.

[31] Raworth, K. 2102. A Safe and Just Space for Humanity. Can We Live within the Doughnut?, *Oxfam Discussion Papers.*

[32] Rockstrom, J. *et al.* 2009. Planetary boundaries: exploring the safe operating space for humanity. *Ecology and Society*.

[33] Roemer, J.E. 2009. The Ethics of Distribution in a Warming Planet, *Cowles Foundation Discussion Paper* 1693, Yale University.

[34] Rose, A. and Kverndokk, S. 2008. Equity and justice in global warming, *Review of Environmental and Resource Economics*.

[35] Sen, A. 1989. Development as capability expansion in Readings in Human Development (2003) edited by Sakiko Fukuda-Parr and A.K. Shiva Kumar pp. 3-16. *Oxford University Press*. New Delhi, India.

[36] Sen, A. 2009. The Idea of Justice. *Cambridge, MA: Harvard University Press*.

[37] Slay, B. 2011. Carbon consumption, transition and developing economies: Sinners, or sinned against?, *manuscript*. UNDP.

[38] Stern, N. 2007. *The Economics of Climate Change – The Stern Review,* Cambridge: Cambridge University Press.

[39] Stiglitz, S., and Fitoussi, J.P. 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress.*

[40] Tol, R. 2008. The Social Cost of Carbon: Trends, Outliers and Catastrophes. Economics: The Open-Access, *Open-Assessment E-Journal,* 2: 2008-25.

[41] Van den Bergh, J. and Verbruggen, H. 1999. Spatial sustainability, trade and indicators: an evaluation of the ecological footprint, *Ecological Economics,* 29(1): 61-72.

[42] Zambrano, E. 2012. "Further thoughts about the proposed SHDI". *Manuscript,* Orfalea College of Business, Cal Poly.

*** United Nations Development Program 1990. "*Human Development Report 1990".* UNDP-HDRO, New York.

*** United Nations Development Program 1994. *"Human Development Report 1994".* UNDP-HDRO, New York.

*** United Nations Development Program 2008. "*Human Development Report 2007/08".* UNDP-HDRO, New York.

*** United Nations Development Program 2011. *"Human Development Report 2011"*. UNDP-HDRO, New York.

*** United Nations Secretary-General's High-level Panel on Global Sustainability 2012. *"Resilient People, Resilient Planet: A future worth choosing"*. New York: United Nations.

Annex 1
Data and mathematical representation of the sustainability adjusted HDI (SHDI)

A.1.1 Data

Carbon dioxide emissions per capita (2008), Annual freshwater withdrawals, percentage of water from own resources (2009) and Adjusted Net Savings (2010) are provided by the World Bank data query17. Land area and permanent crop area (2009) is found in FAO Stats18. The Ecological Footprint (2008) is found in the Global Footprint Network latest report (2011)19. Data regarding extinct and assessed species by country is found in the International Union for Conservation of Nature (IUCN) "Red list"20.

A.1.2 Mathematical representation of the Sustainability Adjusted HDI (SHDI)

This section uses extensively inputs from Zambrano (2012) and Herrero (2012). The world has K countries. For simplicity countries are assigned a number from 1 to K, so that i=1,2, …, K. Total world population is N individuals, where

$$N=\sum_1^K N_i, \tag{1.1}$$

and Ni is the population of country i. Therefore, $\{N_i\}_{i\in k}$ is the country's population. And let us call

$$\theta_i=\frac{N_i}{N} \tag{1.2}$$

For the environmental sustainability indicator j, $\left\{S_j^i\right\}_{i\in k}$, represents the level of use of the environment for indicator j in each country i. $\bar{s}_j$ corresponds to each individual in the planet's 'maximum fair share' according to the planetary boundary for indicator j, that is, the per capita equal share of the global planetary boundary, $\bar{S}_j$, where

$$\bar{S}_j=N^*\bar{s}_j \tag{1.3}$$

We want to create a loss function with respect to the environmental sustainability indicator (or a combination of them). Therefore, let us start with a general definition of what the loss function should comprise.

Definition: A loss function,

$$G_j^i:\langle\{\bar{s}_j,\bar{S}_j\},\{N_i\}_{i\subset k},\left\{S_j^i\right\}_{i\in k},\{\theta_i\}_{i\in k}\rangle\rightarrow[0,1]_{i\in k} \tag{1.4}$$

such that each component of G is weakly increasing in S_i.

This function has three important features:

1. *It depends on the whole world's situation, and gives a particular value for each country.*
2. *It is bounded between 0 and 1, for each country.*
3. *When the pollution of a country increases, all other things equal, the penalty for such a country cannot decrease.*

[17] World Bank, "World Development Indicators", World Data Bank. http://databank.worldbank.org/ddp/home.do.
[18] Food and Agriculture Organization of the United Nations (FAO), FAOSTAT. http://faostat3.fao.org/home/index.html#DOWNLOAD.
[19] Global Footprint Network, "The data tables from the 2010 edition", Data and Results. http://www.footprintnetwork.org/en/index.php/GFN/page/footprint_data_and_results/.
[20] The IUCN Red List of Threatened Species, "Summaries by country", Summary Statistics. http://www.iucnredlist.org/about/summary-statistics.

Now we want some other properties, in order to obtain our desired loss function. With these properties, we specify which countries are going to be positively penalized:

P1. **No penalty for good behavior**. A country that pollutes less than its share minimum fare gets no penalty: If $S_j^i \leq \theta_i \bar{S}_j$ then,

$$G_j^i\left[\langle\{\bar{s}_j,\bar{S}_j\},\{N_i\}_{i\subset k},\{S_j^i\}_{i\in k},\{\theta_i\}_{i\in k}\rangle\right]=0 \tag{1.5}$$

We can call this the exclusion property. Together with the wealth increasing it implies that all countries polluting below their minimum fair share receive no penalty.

P2. **Full penalty for full pollution**. A country that in isolation exceeds the maximum boundary receives full penalty: If $S_j^i \geq \bar{S}_j$, then

Gijsj,Sj,Nii϶k,Sjii∈k,θii∈k=1 (1.6)

This property is similar to the exhaustion property in Herrero and Villar (2001). For countries exceeding the global planetary boundary -and given weak monotonicity- all countries above that level receive full penalty.

P3. **Constant penalty trade-offs**. If two countries, 1 and 2, keeping their emissions in the intervals $[\theta_1\bar{s}_j,\bar{S}_j]$, $[\theta_2\bar{s}_j,\bar{S}_j]$ for indicator j respectively, increment their emissions in the same amount, the relative value of their penalties is constant (independent of the common amount they increase). That is, if $s_1''-s_1'=s_2''-s_2'$, then

$$\frac{G_1''-G_1'}{G_2''-G_2'}=k(1,2) \tag{1.7}$$

This property has been called "Direct Capability", meaning that a country that diminishes (or improves) the environmental variable by an amount of, say "D" when polluting beyond its "fair share", diminishes (improves) its capabilities in direct proportion to "D". P3 is an extension to that principle, but applied to two countries, making explicit a sort of fair treatment in the relationship between the behaviors of the penalties for different countries.

Theorem: *A penalty function satisfies P1. P2 and P3 iff*

$$G_j^i\left[\langle\{\bar{s}_j,\bar{S}_j\},\{N_i\}_{i\subset k},\{S_j^i\}_{i\in k},\{\theta_i\}_{i\in k}\rangle\right]=\max\left\{0,\min\left\{1,\left[\frac{\left[S_j^i-\theta_i\bar{S}_j\right]_+}{\bar{S}_j-\theta_i\bar{S}_j}\right]\right\}\right\}$$
$$=\max\left\{0,\min\left\{1,\left[\frac{\left[s_j^i-\bar{s}_j\right]_+}{\bar{s}_j}\left(\frac{N_i}{N-N_i}\right)\right]\right\}\right\} \tag{1.8}$$

Therefore, we could also represent the loss function G_j^i for indicator j and country i, as the following:

$$G_j^i\left[\langle\bar{s}_j,\{\beta_j^i\}_{i\subset k},\{s_j^i\}_{i\in k}\rangle\right]=\min\left\{1,\beta_j^i*\frac{\left[s_j^i-\bar{s}_j\right]_+}{\bar{s}_j}\right\} \tag{1.9}$$

Given that

$$\frac{\left[S_j^i-\theta_i\bar{S}_j\right]_+}{\bar{S}_j-\theta_i\bar{S}_j}=\frac{\left[S_j^i-\frac{N_iS_j}{N}\right]_+}{\bar{S}_j-\frac{N_i\bar{S}_j}{N}}=\frac{\left[S_j^i-\frac{N_iS_j}{N}\right]_+}{\bar{S}_j\left(1-\frac{N_i}{N}\right)}=\frac{\left[S_j^i-\frac{N_iS_j}{N}\right]_+}{\bar{S}_j}\left(\frac{N}{N-N_i}\right)\left(\frac{N_i}{N_i}\right)=\frac{\left[S_j^i-\frac{N_iS_j}{N}\right]_+}{\bar{S}_j}\left(\frac{N_i}{N-N_i}\right)\left(\frac{N}{N_i}\right)=$$

$$\frac{\left[\frac{s_j^i}{N_i}-\frac{N_i\bar{s}_j}{NN_i}\right]_+}{\frac{\bar{s}_j}{N}}\left(\frac{N_i}{N-N_i}\right)=\frac{\left[\frac{s_j^i}{N_i}-\frac{\bar{s}_j}{N}\right]_+}{\frac{\bar{s}_j}{N}}\left(\frac{N_i}{N-N_i}\right)=\frac{\left[s_j^i-\bar{s}_j\right]_+}{\bar{s}_j}\left(\frac{N_i}{N-N_i}\right) \qquad (1.10)$$

So,

$$\beta_j^i=\left(\frac{N_i}{N-N_i}\right) \qquad (1.11)$$

where c refers to the environmental sustainability indicator used, in this case carbon dioxide emissions per capita, so j=c,w,l; and, the operation $[\quad]_+$ is defined as $[x]_+=\max\{x,0\}$.

The term

$$\frac{\left[s_j^i-\bar{s}_j\right]_+}{\bar{s}_j} \qquad (1.12)$$

measures the degree or intensity of "unfair" or "excessive" use of the environment of the average citizen in each country i (as a proportion of the per capita threshold or maximum fair share). While β_j^i measures the weight given to the average unfair used by country i of the environment (measure by indicator j). Notice that our empirical implementation of the SHDI used only per capita CO_2 emissions, but the framework could be flexible to the use of more indicators.

So, G_j^i is the overall loss function that is imposed to country *i*'s human development achievements given its degree of unfair use of the environment, according to the global planetary boundary for environmental indicator j.

G_j^i is intended to be the answer to the following question: Imagine a country A, with perfect achievements in health, education, and income (thus having an HDI of "1"), and that it is between the global environmental boundaries (thus also having an SHDI of "1"). Compare this to country B, also with perfect achievements in health, education, and income but with a level of, say, its per capita CO_2 emissions are exactly twice the level of per capita maximum fair share. Country B will also have an HDI of "1" but an SHDI of (1*(1- G_c^i)). This is similar for any other indicator on j.

Given the existing research on the planetary boundaries and the available data, we are able to have measures of the fair or unfair use of the global environment.

The intuition for the value of β_j^i is that we can argue the case so that when a country, say country *i*, alone hits the planetary boundary, this will impose unacceptable negative effects on the available choices of future generations and thus in this case the country receives the maximum loss and therefore $G_j^i=1$. This will create two changes in the shape of the loss function for country *i* on environmental dimension j. The first one is that its value is 0 if the country's per capita use of the environment is lower than the fair per capita share (P1. No penalty for good behavior); and the second one is that it has a value of 1 if country's per capita use of the environment is such that it hits or exceeds the planetary boundary (P2: full penalty for full pollution). The intuition could be enhanced by the following figure.

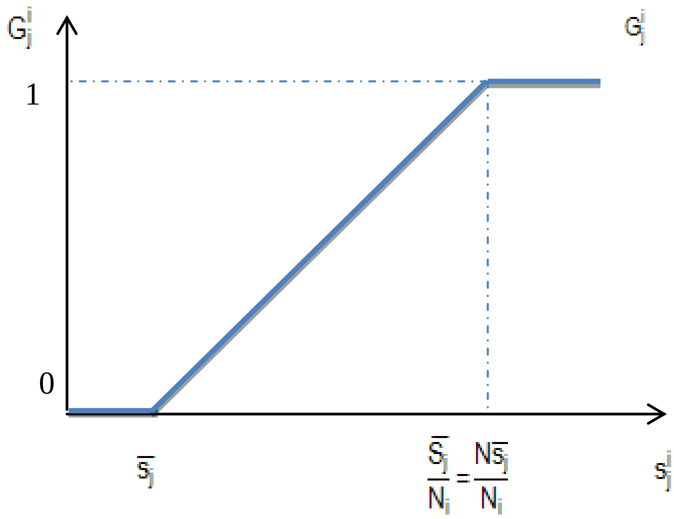


Figure A.1.1. Graphical representation of the loss function G_j^i

We can therefore give the following interpretation of the two components of G_j^i:

The term

$$\frac{\left[s_j^i-\bar{s}_j\right]_+}{\bar{s}_j} \tag{1.13}$$

is called the *fair share* of the environment, given that this is an expression that compares the per capita use of the environment of a citizen in country *i* to the per capita threshold or maximum fair share according to the planetary boundary. This terms capture when a country is having an excessive use of the environment by exceeding its fair share.

The term

$$\frac{N_i}{N-N_i} \tag{1.14}$$

is called the *global responsibility* term, given that this is an expression that gives higher weight to excessive use of the environment behavior, the larger is the population of the country. In other words, the larger a country is with respect to the rest of the world, the larger is its responsibility for the use of the environment from its average citizen.

Including levels of uncertainty in the loss function

Since the planetary boundaries are intrinsically uncertain values, we use the confidence interval that Rockstrom *et al*. (2009) use in their estimations. Therefore, the Figure A.1.1 under two possible thresholds becomes:

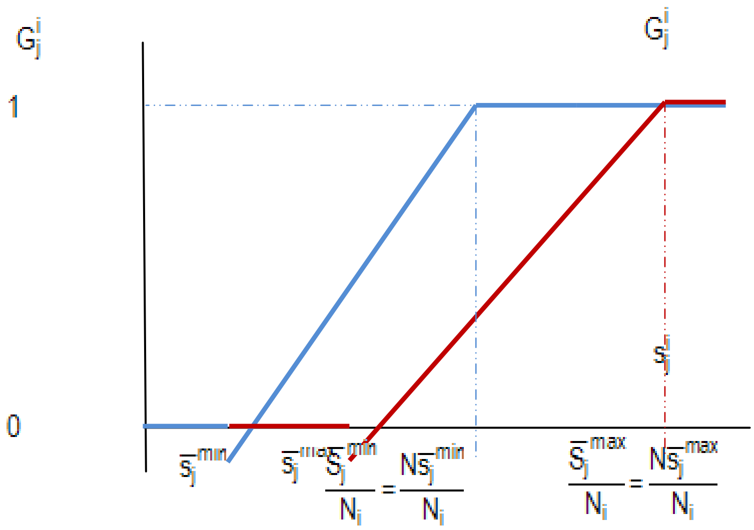


Figure A.1.2. Graphical representation of the loss function G_j^i with a minimum and a maximum planetary boundary

An interesting possibility is to define our loss function as to include the minimum per capita fair share and the maximum global planetary boundary. The graph would therefore become:

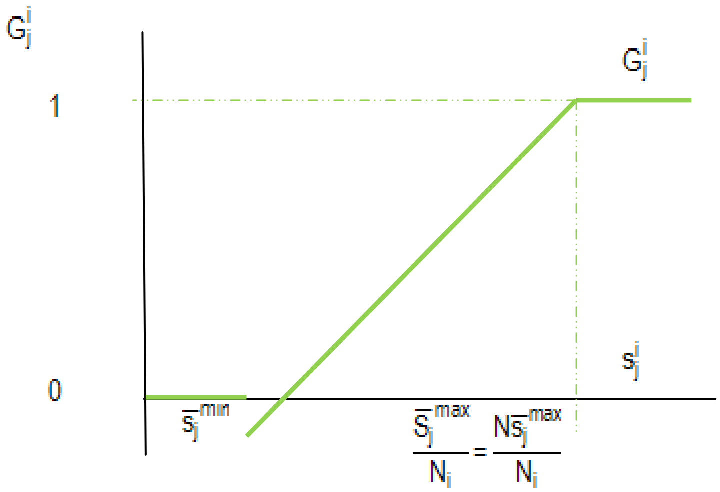


Figure A.1.3. Graphical representation of the loss function G_j^i with the minimum per capita fair share and the maximum global planetary boundary

In this case, the loss function would be defined as:

$$G_i\left[\langle\{s^{min}, s^{max}\}, \{w_i\}_{i\in N}, \{s_i\}_{i\in N}, \{\theta_{ij}\}_{i\in N}\rangle\right] \tag{1.15}$$

And the same former three properties would apply.
The loss function would look like this:

$$G_i\left[\langle\{s^{min}, s^{max}\}, \{w_i\}_{i\in N}, \{s_i\}_{i\in N}, \{\theta_{ij}\}_{i\in N}\rangle\right] = \max\left\{0, \min\left\{1, \left[\frac{[s_i - \theta_i s^{max}]_+}{s^{max} - \theta_i s^{min}}\right]\right\}\right\} \tag{1.16}$$

From this, we can derive the global responsibility term, by setting G_j^i equal to 1. Therefore, when country's i per capita consumption hits the planetary threshold, so for this country $s_j^i = \frac{\bar{S}_j^{max}}{N_i} = \frac{N\bar{s}_j^{max}}{N_i}$. Given this, β_j^i can be defined as follow:

$$1=\beta_j^i*\frac{\left[\frac{\bar{S}_j^{max}}{N_i}-\bar{s}_j^{min}\right]_+}{\bar{s}_j^{min}} \tag{1.17}$$

We can think that the maximum threshold is a value proportional to the minimum:

$\bar{S}_j^{max}=\alpha_j\bar{S}_j^{min}$, so that $\alpha_j=\frac{\bar{S}_j^{max}}{\bar{S}_j^{min}}$

In which $\alpha_j>1$

Therefore,

$$1=\beta_j^i*\frac{\left[\frac{\alpha_j\bar{S}_j^{min}}{N_i}-\bar{s}_j^{min}\right]_+}{\bar{s}_j^{min}} \tag{1.18}$$

So,

$$1=\beta_j^i*\frac{\bar{s}_j^{min}\left[\alpha_j N-N_i\right]_+}{\bar{s}_j^{min}N_i} \tag{1.19}$$

$$\beta_{ji}=\frac{N_i}{\alpha_j N-N_i} \tag{1.20}$$

Calculation of SHDI

We can adjust the HDI by using the loss function G_j^i for indicator j and country i:

$$SHDI_j^i=\left(1-G_j^i\right)*HDI^i \tag{1.21}$$

Giving the data limitations discussed in section 3, we focus the calculation of the loss function G^i for country i only to per capita CO_2 emissions:

$G^i=G^i_{emissions}$ With this loss function, we adjust the HDI for country i as follows:

$$SHDI^i=\left(1-G^i\right)*HDI^i \tag{1.22}$$

Annex 2
Relationship between sustainability indicators and the Human Development Index

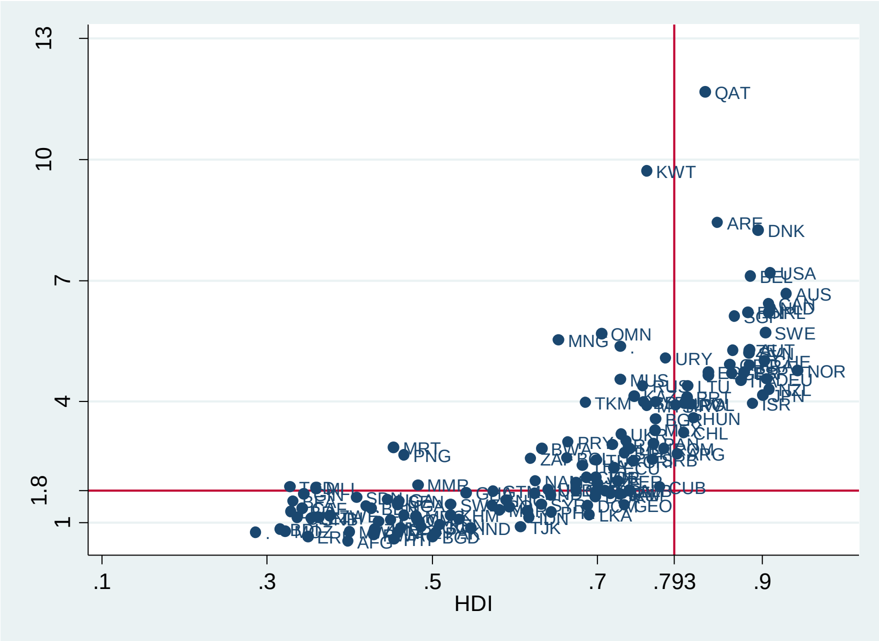


Figure A.2.1. Human Development Index and Ecological Footprint (2008)

Source: UNDP and Global Footprint Network (2011), own calculations.

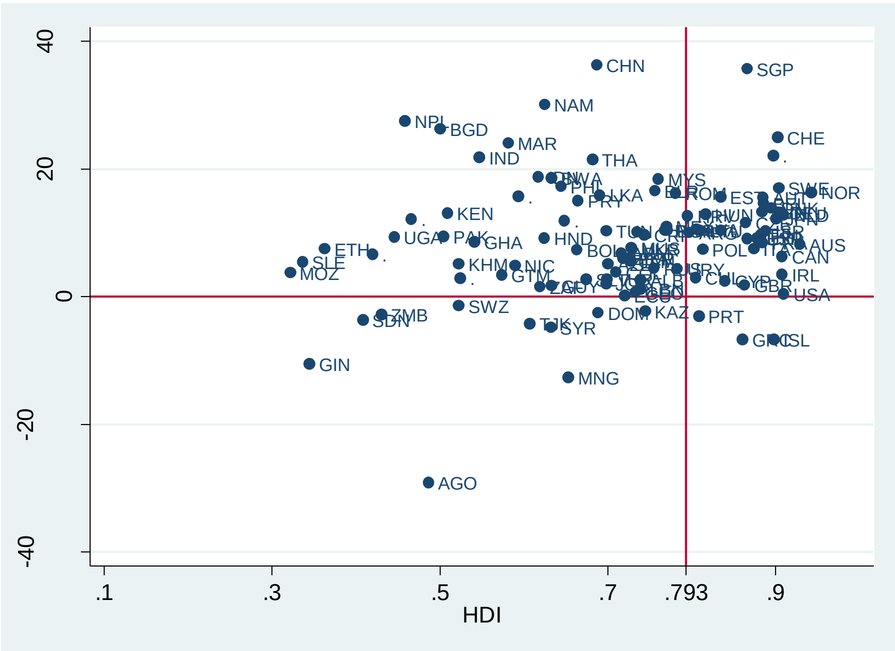


Figure A.2.2. Human Development Index and Adjusted Net Savings (2010)

Source: UNDP and World Bank, own calculations.

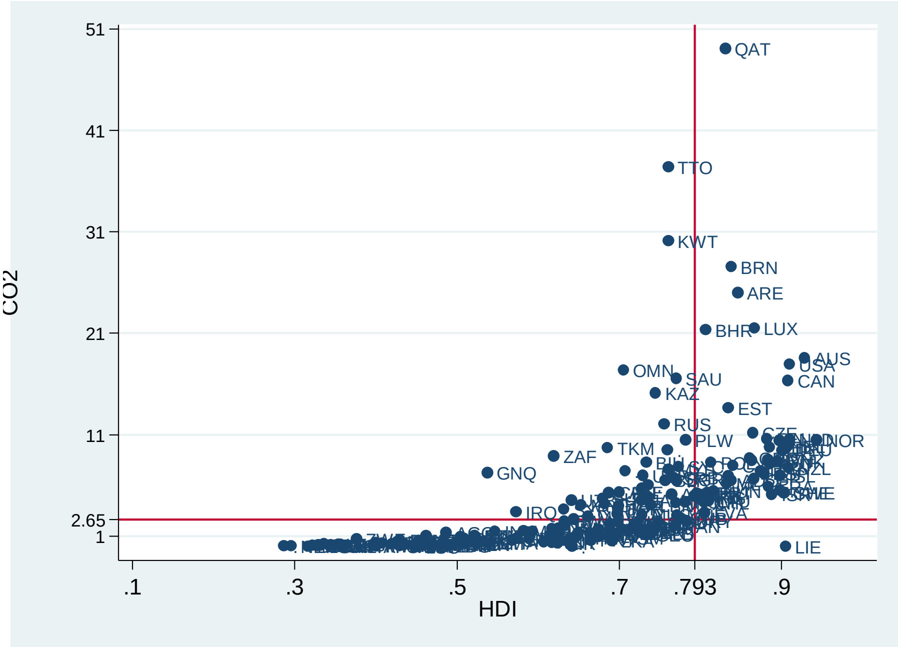


Figure A.2.3. Human Development Index and CO_2 emissions per capita (2008)

Source: UNDP and World Bank, own calculations.

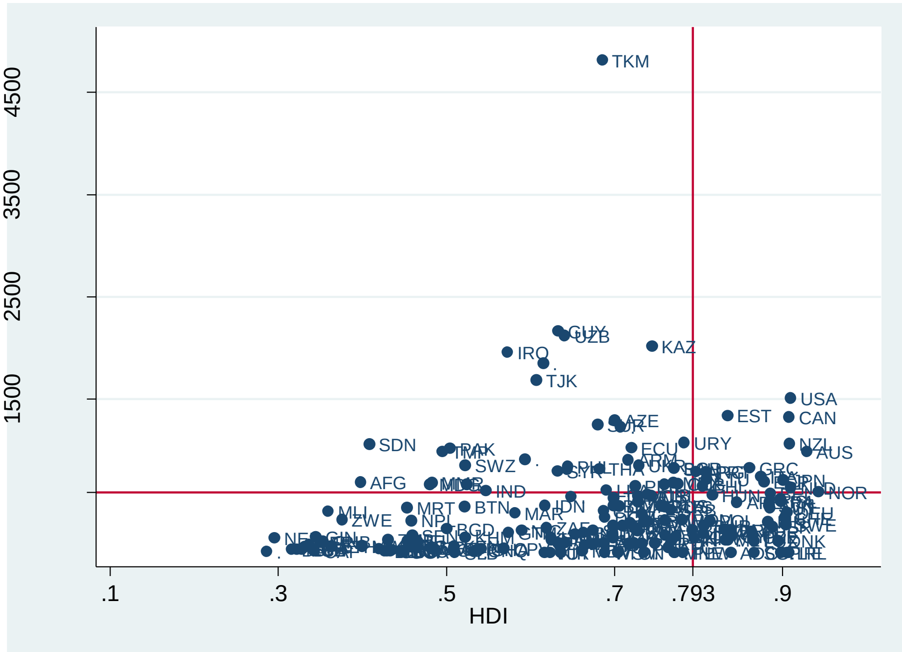


Figure A.2.4. Human Development Index and fresh water withdrawals per capita (2009)

Source: UNDP and World Bank, own calculations.

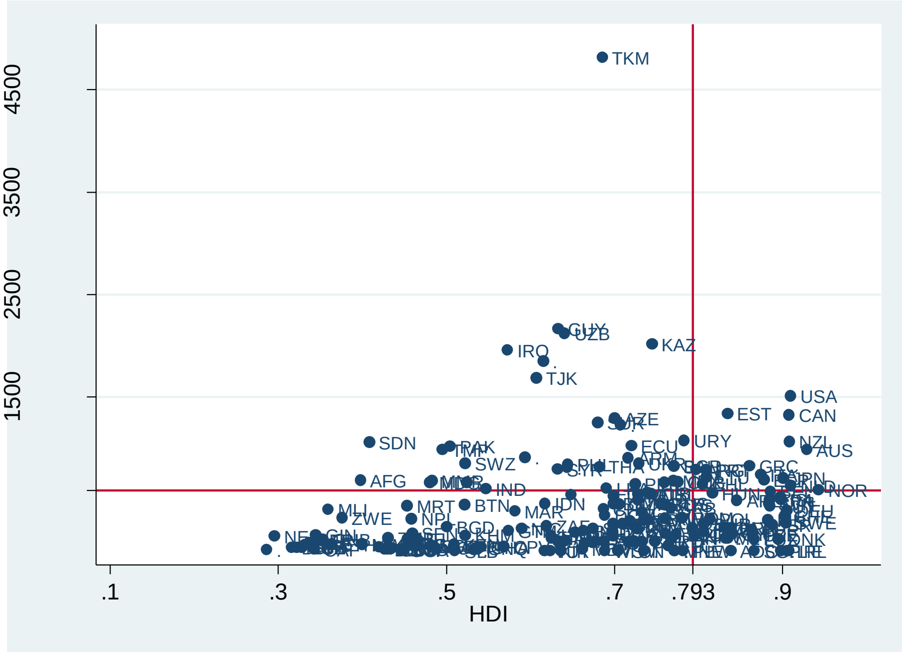


Figure A.2.5. Human Development Index and share of land with permanent crops (2009)

Source: UNDP and FAO, own calculations.

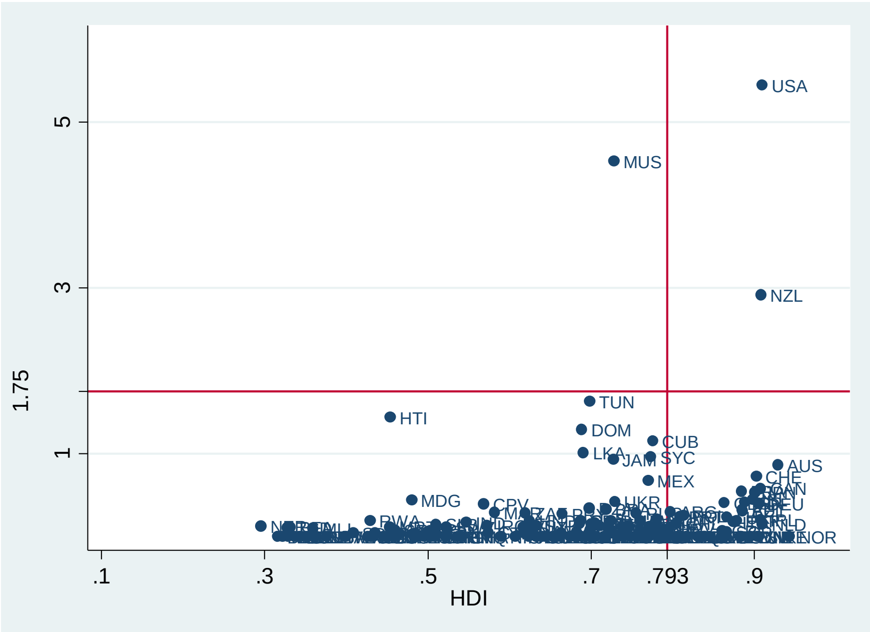


Figure A.2.6. Human Development Index and species extinct as percentage of total species (2010)

Source: UNDP and the IUCN "Red list", own calculations.

Table A.2.1 is similar to the one presented in section 4, which shows a statistical description of the relevant variables. For each one of them, we show the values for the whole set of countries in the HDI sample (column "All") and in its left side, the values for the subset of countries transgressing the planetary boundary (at the lower threshold). In the case of the Ecological Footprint (EFP) the threshold is 1.8, and for the Adjusted Net Savings (ANS) the threshold value is 0, and for the share of extinct species over total we use one standard deviation above the mean.

Table A.2.1. Selected sustainability indicators

stats	EFP		ANS		CO_2		Freshwater		Crop share		Ext share	
	above threshold (>1.8)	All	below threshold (<0)	All	above threshold (>2.66)	All	above threshold (>590.29)	All	above threshold (>15)	All	above threshold (>1.75)	All
mean	4.1	2.9	-6.98	8.64	9.13	4.87	1090.68	467.78	25.1	4.02	4.79	0.25
s.d.	1.91	2.05	7.44	9.6	7.55	6.71	688.53	556.26	10.44	6.82	1.43	0.75
min	1.8	0.54	-29.16	-29.16	2.69	0.02	604.76	8.94	15.28	0	2.92	0
max	11.68	11.68	-1.43	36.26	49.05	49.05	4818.18	4818.18	46.88	46.88	6.25	6.25
N (obs.)	82	140	13	104	90	185	49	172	11	186	4	186

Source: UNDP, World Bank and Global Footprint Network (2011), own calculations.

As we can see from the figures, the only two indicators with a strong positive and statistically significant correlation with HDI are EFP and CO_2 emissions per capita (.75 and .55, respectively). These indicators have the largest share of countries above the threshold, while the share of extinct species over total has the lowest. In fact, their figures look very similar when we just represent the common sample of countries for which both indicators exist (figure A.2.1 has an original sample of 185 countries, here the common sample is only 140 countries).

Figure A.2.7. Human Development Index, CO2 emissions per capita and Ecological Footprint (2008) (Common sample, 140 countries)

Source: UNDP, World Bank and Global Footprint Network (2011), own calculations.

Annex 3
Changes in rank of the top 10 and bottom 10 countries according to the HDI and SHDI ranks

The following tables present the top 10 countries (out of 185) according to the HDI rank and SHDI as well as the change in rankings due to the adjustment from unsustainable environmental behavior. As the tables shown, most of the changes in rankings occur at the upper portion of the distribution, while fewer changes occur at the lower part of it. This result is just consistent with the fact that relatively low human development countries contribute very little to the global environmental un-sustainability.

Table A.3.1. Changes in rank of the top 10 countries after adjusting for sustainability (lower bound)

Country	HDI	SHDI	Loss due to un-sustainability	Rank HDI	Rank SHDI	Change in rank
Norway	0.9430	0.9410	0.0021	1	1	0
Australia	0.9289	0.9106	0.0197	2	2	0
Netherlands	0.9099	0.9034	0.0071	3	5	2
United States	0.9099	0.6616	0.2728	4	106	102
New Zealand	0.9084	0.9073	0.0012	5	3	-2
Canada	0.9081	0.8848	0.0257	6	12	6
Ireland	0.9081	0.9065	0.0018	7	4	-3
Germany	0.9051	0.8771	0.0310	8	18	10
Sweden	0.9038	0.9026	0.0014	9	6	-3
Switzerland	0.9025	0.9016	0.0011	10	7	-3

Table A.3.2. Changes in rank of the top 10 countries after adjusting for sustainability (upper bound)

Country	HDI	SHDI	Loss due to un-sustainability	Rank HDI	Rank SHDI	Change in rank
Norway	0.9430	0.9420	0.0010	1	1	0
Australia	0.9289	0.9188	0.0109	2	2	0
Netherlands	0.9099	0.9067	0.0035	3	5	2
United States	0.9099	0.7733	0.1500	4	51	47
New Zealand	0.9084	0.9079	0.0005	5	3	-2
Canada	0.9081	0.8955	0.0139	6	10	4
Ireland	0.9081	0.9073	0.0008	7	4	-3
Germany	0.9051	0.8919	0.0145	8	12	4
Sweden	0.9038	0.9035	0.0003	9	6	-3
Switzerland	0.9025	0.9023	0.0002	10	7	-3

Table A.3.3.Top 10 countries for HDI and SHDI

Country	Rank HDI	Country	Rank SHDI (lower bound)	Country	Rank SHDI (upper bound)
Norway	1	Norway	1	Norway	1
Australia	2	Australia	2	Australia	2
Netherlands	3	New Zealand	3	New Zealand	3
United States	4	Ireland	4	Ireland	4
New Zealand	5	Netherlands	5	Netherlands	5

Country	Rank HDI	Country	Rank SHDI (lower bound)	Country	Rank SHDI (upper bound)
Canada	6	Sweden	6	Sweden	6
Ireland	7	Switzerland	7	Switzerland	7
Germany	8	Iceland	8	Hong Kong, China (SAR)	8
Sweden	9	Hong Kong, China (SAR)	9	Iceland	9
Switzerland	10	Denmark	10	Canada	10

Results with combined thresholds (minimum per capita fair share and maximum global planetary boundary)

Table A.3.4. Top rank positions lost with SHDI (combined thresholds)

Country	HDI	SHDI	Loss due to un-sustainability	Rank HDI	Rank SHDI	Number of position lost
United States	0.9099	0.7595	0.1652	4	59	55
China	0.6871	0.5941	0.1354	100	126	26
Russian Federation	0.7553	0.7213	0.0451	65	81	16
Japan	0.9006	0.8746	0.0289	11	20	9
Canada	0.9081	0.8937	0.0158	6	11	5
Germany	0.9051	0.8879	0.0190	8	12	4
Turkey	0.6991	0.6968	0.0034	91	95	4
Saudi Arabia	0.7704	0.7601	0.0133	55	58	3
Malaysia	0.7605	0.7569	0.0048	60	63	3
Ukraine	0.7292	0.7245	0.0065	75	78	3
Iran (Islamic Republic of)	0.7074	0.6989	0.0121	87	90	3
Netherlands	0.9099	0.9059	0.0044	3	5	2
South Africa	0.6194	0.6128	0.0106	122	124	2
Italy	0.8738	0.8653	0.0097	23	24	1
Poland	0.8133	0.8075	0.0072	38	39	1
Venezuela (Bolivarian Republic of)	0.7351	0.7326	0.0034	72	73	1

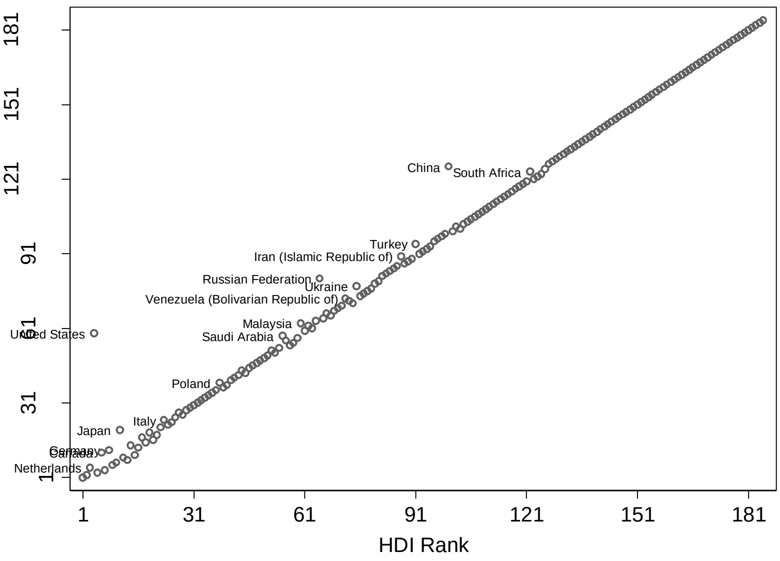


Figure A.3.1. Rank comparison between original HDI and SHDI (combined thresholds)

Volume III Issue 2(6) Winter 2012

DOI: 10.2478/v10261-012-0007-0

BOUNDED RATIONALITY AND PERFECT RATIONALITY: PSYCHOLOGY INTO ECONOMICS

Daniele Schilirò
DESMaS "V. Pareto"
University of Messina
schi.unime@katamail.com

Abstract:

Mathematical algorithms often fail to identify in time when the international financial crises occur although, as the classical theory of choice would suggest, the economic agents are rational and the markets are or should be efficient and behave also rationally.

This contribution tries to highlight some well-known limits of the classical theory of rational choice and compare this theory of choice with bounded rationality, which is a different notion of rationality, and with an approach that seeks to combine economics and psychology, based on experimental data, which established itself as behavioral economics.

The work also examines part of the literature of behavioral finance which has given important contributions in explaining the behavior and the anomalies of financial markets. A final reference is dedicated to neuroeconomics that is gaining more and more ground in the analysis of economic behavior.

Keywords: bounded rationality, procedural rationality, rational choice, behavioral economics, neuroeconomics.

JEL Classification: C60, B52, D81, D83, D86.

1. Introduction

The global financial crisis has created a climate of great uncertainty. People ask why speculation is constantly present in the markets and why individuals (at least some of them) are incapable of curbing speculative instincts to preserve the common good, the stability of the all system rather than the gains of a few. Thus we wonder why mathematical algorithms fail to identify in time when the international financial crises occur if, as the classical theory of choice would suggest, the economic agents are rational and the markets are efficient and behave also rationally.

This contribution highlights some well-known limits of the classical theory of rational choice, underlining the great importance of the notion of bounded rationality, which has in Herbert Simon its most influential theorist. At the same time, this work addresses the relationship between psychology and economics and the influence of psychological factors on economic behavior, focusing on the theoretical explanations provided by behavioral economics. It also examines part of the literature of behavioral finance which has given significant contributions to the analysis of the behavior and of the anomalies of financial markets. A final reference is dedicated to neuroeconomics that is gaining more and more ground in the analysis of economic behavior.

2. Perfect rationality in economics

Economics in its classical conception is seen as a normative theory. In its neo-positivist approach of systemic-formal nature, economics takes the form of nomologic - deductive propositions, which are obtained by reasoning, starting from unproven axioms. With these axioms we deduce the propositions of the theory, which requires the use of logic and mathematics. Thus economics presents itself as a rational science in the sense that its propositions are obtained by means of logic, in a way which is similar to rational mechanics. In economics, moreover, rationality is interpreted in terms of consistency not of substance. We have therefore a syntactic and non-semantic notion of rationality. The agents are rational if they have a coherent criterion of choice. The consistency of the choices implies that the agents are represented by a system of preference. Economics describes the choice as a rational process driven by a single cognitive process that includes the principles of the *Theory of rational choice* and it orders the decisions on the basis of their subjective expected utility. In this view the "*homo oeconomicus"* appears perfectly rational and has a complete knowledge, while his economic choices, guided by rationality, are self contained in the economic sphere without affecting other aspects of the individual such as the emotions or being influenced by the environment.

2.1. The theory of rational choice

In the theory of rational choice (TRC) the first basic parameter which is taken into consideration is the 'preference'. The theory sets several basic axioms on the preference of a rational agent. It follows that an agent who wishes to call himself fully rational must align his preferences to each of these conditions.

Given the expressions of preference relation xPy ("x is preferred to y") and the relationship of indifference xIy ("x is indifferent to y"), these conditions are the following

1. If xPy, then no yPx.
2. If xPy, then no xIy.
3. If xIy, then no xPy and also no yPx.
4. xPy or yPx or xIy, for any two relevant results x and y.
5. If xPy and yPz, then xPz.
6. If xPy and xIz, then zPy.
7. If xPy and yIz, then xPz.
8. If xIy and yIz, then xIz.

Conditions 1 – 3 are usually considered as a single order condition, while the conditions 5 – 8 are called conditions of transitivity. The different ways of numbering the elements in an order of preference are called 'utility functions' or 'utility scales' for the subject. The utility functions are therefore the instrument to characterize the preferences of an individual.

By means of the utility functions it is possible to decline formally the principle of maximization.

If an individual's preferences satisfy appropriate consistency conditions, then it is possible to associate a numerical value to each outcome through a utility function u: E → R.

In the field of analysis of choice under uncertainty, the most important contribution is the expected utility theory, proposed by von Neumann and Morgenstern (1944). The theoretical framework of von Neumann-Morgenstern is generally accepted as a normative model of rational choice. In this model rational agents want to obey its axioms.

According to von Neumann and Morgenstern, individuals generally move in the reality following predetermined patterns of behavior, at the base of which there is the assumption that they always prefer to have a greater wealth than less. The theory studies the preferences underlying consumer behavior under risk, i.e. when the subject is asked to make a decision without knowing with certainty which *ex ante* state of the world will happen, but he knows the probability distribution, that is, it is known to him a list of possible events, each of which he associates a probability of occurrence. This theory assumes that each individual has stable and consistent preferences, and that he makes decisions based on the principle of maximization of subjective expected utility. So given a set of options and beliefs expressed in probabilistic terms, it is assumed that the individual maximizes the expected value of a utility function u (x). The individual uses probability estimates and utility values as elements of calculation to maximize his expected utility function. Thus he evaluates the relevant probabilities and utilities on the basis of his personal opinion but also using all relevant information available.

The expected utility theory is nothing more than a criterion that facilitates choice under risk we get the utility function that can take three forms:

i) is concave when describing the preferences of a risk averse individual;
ii) is convex type when describing the preferences of an individual willing to risk;
iii) is linear when describing the preferences of a risk-neutral individual.

Thus an individual averse, neutral or risk lover has indifference curves convex, linear or concave. There are five axioms which the preferences must satisfy such as to represent them through the expected utility:

1. The first axiom requires that preferences are complete and consistent. Given two or more distributions an individual is always able to indicate that he prefers; he always knows how to choose and place the distributions in some order on the scale of preferences. *The consistency requirement implies that preferences are transitive*. If you prefer the distribution A to B and B is preferred to C, then A is preferred to the distribution of C; if this does not happen it would create circularity and the individual would not be able to choose. The axiom reminds us that intrasitive preferences lead to irrational behavior.

2. The second axiom is that of monotonicity which requires that an individual, given two distributions that have the same consequences, tends to prefer the opportunity that offers the best result with the highest probability.

3. The third axiom of continuity implies that a subject, that is before an alternative by which he can achieve with certainty a given result or to have a probability distribution that gives with probability α the better event and with probability 1– α the worst event, is always able to give a probability α that makes him indifferent between the two alternatives.

4. The fourth axiom of independence is crucial for the formulation of the expected utility, so as to be valid it is necessary that the utilities of each consequence are weighted by their probabilities and summed. The sum

operation is possible only if the utilities are independent. Suppose an individual is indifferent between two events which are certain, then he will also be indifferent to combinations of all these events with any distribution for any probability.

5. The last axiom is that of reduction. Suppose we roll a die: if there is 1 we win otherwise we lose; thus we have a chance of winning equal to one sixth. The prize does not consist in a certain sum of money but in the participation in a lottery whose probability of winning is equal to one third. The axiom of reduction asserts that what matters for the individual is the overall probability of winning equal to 1/6 x 1/3 = 1/18 since the two games are independent, not the pleasure to participate to individual games; the premise also assumes that the individual is always able to calculate such a probability.

The most important result of this theory is the expected utility theorem.

2. 2. Expected utility theorem

Given the five axioms it is possible to build one and only one intervallic function u which has the following properties:

I) $u(x) > u(y)$ if and only if xPy.

II) $u(x) = u(y)$ if and only if xIy.

III) $u[L(a,x,y)] = au(x) + (1 – a)u(y)$.

Any u^* which satisfies I)-III) is a positive linear (affine) transformation of u.

The expected utility theory has been generally accepted as a normative model of rational choice, defining what decisions are rational. If an individual does not maximize his expected utility he is designed to violate in his choices some precise axiomatic principles, which are rationally binding.

This theory has also been applied as a descriptive model of economic behavior (Friedman, Savage, 1948; Arrow, 1971) so as to constitute an important reference model for economic theory.

Finally, what emerges from the analysis of choice under uncertainty is the complexity of the system of choice. In fact, one must take into account three conditions of rationality, namely the existence of a regular system of preferences on the consequences, the rationality of expectations about the consequences of actions, the rationality of the function that determines the system of preference on the actions, on the basis of expectations about the consequences of actions and the consequences to the system of preference (Montesano, 2005; Schilirò, 2011).

3. Psychology into economics. The cognitive dimension

Within the scientific community there has been a growing need to consider adequately the complexity of economic phenomena and processes that guide the choices of the individuals.

During the fifties there have been important explorations along the boundaries between economics and psychology. This line of research determined the development of behavioral economics which exactly relates psychological factors to economic behavior (Rabin, 1998). Experimental results had emerged that questioned the validity of the classical model of rational choice (Simon, 1959). On the one hand, Simon's approach, developed on bounded rationality and problem solving, criticized – on the basis of analysis conducted on the field – the lack of realism of the neoclassical economic theory based on the assumption of full rationality. On the other hand, the pioneering experimental studies of Allais in 1952 set violations of utility theory and proved the systematic discrepancy between the predictions of traditional decision theory and actual behavior.

3.1. The Allais' paradox

In 1952, Maurice Allais presented in Paris his famous paradox" to an audience composed of the best economist of his generation; among others, Kenneth Arrow, Paul Samuelson, Milton Friedman, Jacob Marschak, Oskar Morgenstern and Leonard Savage.

The "paradox" consists in presenting a subject in two situations. In the first situation **(A)** the person is proposed to choose between getting for sure \$ 1,000,000 (a) and receive a lottery (b) which has 0.1 probability to win \$ 5,000,000, 0.89 probability of winning \$ 1,000,000 and 0.01 probability of not winning anything. In the second situation **(B)** the person is proposed to choose between a lottery (c) which has 0.1 probability to win \$ 5,000,000 and 0.9 probability of not winning anything and another lottery (d) which has 0.11 chance of winning \$ 1,000,000 and 0.89 probability of not winning anything. We would expect that a rational individual chooses (a) in the first situation and chooses (c) in the second situation. But this outcome, apparently evident, contradicts the utility theorem. In fact, calculating the utilities for each choice we obtain:

$u\,(a) = u\,(1M)$ $\qquad$ $u\,(b) = 0.1\,u\,(5M) + 0.89\,u\,(1M) + 0.01\,u\,(0)$

$u\,(c) = 0.1\,u\,(5M) + 0.9\,u\,(0)$ $\qquad$ $u\,(d) = 0.11\,u\,(1M) + 0.89\,u\,(0)$

From which:

u (a) - u (b) = 0.11 u (1M) – [0.1 u (5M) + 0.01 u (0)]

u (d) - u (c) = 0.11 u (1M) – [0.1 u (5M) + 0.01 u (0)]

The utility theorem tells us that if the individual prefers (a) with respect to (b): (u (a) > u(b)) in the first situation *(A),* then the individual must prefer (d) to (c) : (u (d)> u (c)) in the second situation *(B)* and vice versa, hence the "paradox".

Therefore, the results of laboratory experiments conducted by Allais[21] have shown that individuals chose inconsistently and that they preferred solutions which did not maximize the expected utility. In this way, Allais has demonstrated that the axiomatic definition of rationality did not allow the description and even prediction of economic decisions.

Another "paradox" has been identified by Ellsberg (1961), who, by means of experiments, demonstrated another type of inconsistency in preferences, showing that individuals prefer to bet on a lottery with a chance of obtaining a win already known that on a lottery with ambiguous results. This aversion to uncertainty (ambiguity) of the individual is completely ignored in the expected utility model from a descriptive point of view, while is not considered acceptable from a normative point of view.

3.2. Bounded rationality

In economics the concept of bounded rationality is associated to Herbert Simon (1955, 1956, 1957, 1972, 1979, 1991), who proposed the idea of bounded rationality as an alternative basis for the mathematical modeling of decision making.

Simon has coined the term 'bounded rationality' in *Models of Man* (1957). In his view, rationality of individuals is limited by the information they have, the cognitive limitations of their minds, and the finite amount of time they have to make decisions. Bounded rationality expresses the idea of the practical impossibility (not of the logical impossibility) of exercise of perfect (or 'global') rationality (Simon, 1955). "Theories that incorporate constraints on the information-processing capacities of the actor may be called *theories of bounded rationality*" (Simon, 1972, p.162). Simon argues that most people are only partly rational while are emotional/irrational in the remaining part of their actions. He maintains that although the classical theory with its assumptions of rationality is a powerful and useful tool, it fails to include some of the central problems of conflict and dynamics which economics has become more and more concerned with (Simon, 1959, p.255). Simon identifies a variety of ways to assume limits of rationality such as risk and uncertainty, incomplete information about alternatives, complexity (1972, pp.163-164).

Furthermore, he asserts that an individual who wants to behave rationally must consider not only the objective environment, but also the subjective environment (cognitive limitations), thus you need to know something about the perceptual and cognitive process of this rational individual. Simon, therefore, considers the psychological theory very important to enrich the analysis for a description of the process of choice in economics. This is why he adopts the notion of procedural rationality, a concept developed within psychology (Simon, 1976), which depends on the process that generated it, so rationality is synonymous of reasoning. According to Simon a search for procedural rationality is the search for computational efficiency, and a theory of procedural rationality is a theory of efficient computational procedures to find good solutions (Simon, 1976, p.133). Procedural rationality is a form of psychological rationality which constitutes the basic concept of Simon's behavioral theory (Novarese, Castellani, Di Giovinazzo, 2009; Barros, 2010; Graziano, Schilirò, 2011), in contrast to economic rationality, defined by Simon as 'substantive rationality'.[22]

Another way to look at bounded rationality is that, because individuals lack the ability and resources to arrive at the optimal solution, they instead apply their rationality only after having greatly simplified the choices available. Actually, individuals face uncertainty about the future and costs in acquiring information in the present. These two factors limit the extent to which agents can make a fully rational decision. Thus, Simon claims, agents have only bounded rationality and are forced to make decisions not by 'maximization', but rather by satisfying[23], i.e. setting an aspiration level which, if achieved, they will be happy enough with, and if they don't, try to change either their aspiration level or their decision. Satisfying is the hypothesis that allows to the conception of diverse decision procedures and which permits rationality to operate in an open, not

[21] Allais (1953).

[22] Models of procedural rationality have been devised by Rubinstein (1998).

[23] The term 'satisfying' appears in Simon, 1956. Later Simon (1957, p.205) says "*The key to the simplification of the choice process…is the replacement of the goal of maximizing with the goal of satisfying, of finding a course of action that is 'good enough*'".

predetermined, space (Barros, 2010; Schilirò, Graziano, 2011). Real-world decisions are made using fast heuristics, 'rules of thumb' that satisfy rather than maximize utility over the long run. Thus agents employ the use of heuristics to make decisions rather than a strict rigid rule of optimization. The agents do this because of the complexity of the situation, and their inability to process and compute the expected utility of every alternative action. In fact there are limits of attentional capacity, mnemonic and computational binding the computational load, hence the usefulness of automatic routines. Rationality is bounded by these internal constraints in the uncertain real world. Simon then relates the concept of bounded rationality to the complementary construct of procedural rationality, which is based on cognitive processes involving detailed empirical exploration and procedures (search processes) that are translated in algorithms, in contrast to the notion of perfect rationality, that is based on substantive rationality, which derives choices from deductive reasoning and from a tight system of axioms, an idea of rationality that has grown up strictly within economics (Simon, 1976, 1997). For Simon "as economics becomes more and more involved in the study of uncertainty, more and more concerned with the complex actuality of business-decision making, the shift in program will become inevitable. Wider and wider areas of economics will replaces the over - simplified assumptions of the situationally constrained omniscient decision-maker with a realistic (and psychological) characterization of the limits of Man's rationality and the consequences of those limits for his economic behavior" (Simon, 1976, pp.147-148).

Simon, however, does not reject the neoclassical theory *tout court* he describes a number of dimensions along which neoclassical models of perfect rationality can be made somewhat more realistic, while sticking within the vein of fairly rigorous formalization. These include: limiting what sorts of utility functions there might be, recognizing the costs of gathering and processing information, the possibility of having a "multi-valued" utility function.

Simon's work has been followed in the research on judgment and decision making, both in economics and psychology, so in his line of research psychology entered into economics. There are two major approaches which produced important insights into perception mechanisms shaping the individual's internal representation of the problem. The first is the "heuristics and biases" program (Tversky, Kahneman, 1974), which has been fundamental to the contemporary development of behavioral economics. Behavioral economics has criticized some of the tenets of mainstream economics as psychologically unrealistic (Rabin, 2002). This line of research generally begins with assumptions rooted in psychological regularity and asks what follows from those assumptions (Camerer, Loewenstein, 2003). Tversky and Kahneman offer a theoretical explanation about the observed deviations from perfect rationality, noting that people rely on "heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations" (1974, p.1124). They therefore do not abandon the assumption that individuals are intelligent and intentional in making decisions, but they assume systematic and specific biases that move away the judgment from the perfect rationality of individuals. Moreover, according to Kahneman and Thaler (2006) to accept a model where individuals maximize their utility function, which, by hypothesis, is perceived to be consistent, accurate and also stable over time, is not possible anymore, because individuals often make systematic errors in predicting their future experience and results, thus failing to maximize their utility. This occurs because individuals in acting face real difficulty in assessing their preferences and, therefore, prefer the pursuit of instant gratification, which, however, are often inconsistent with their long-term preferences (Rabin, 1998). The "failures" of perfect rationality depend also on the specific ways in which people select and process the information mentally. The attitude to risk of the individual varies depending on frame within which lies the choice (Kahneman and Tversky, 1979). Kahneman and Tversky with their 'Prospect theory' (1979, 1984, 1986) have shown experimentally the presence of inconsistent judgments and choices by an individual facing the same problem presented in different frames ('invariance of failures'). It follows that the frame, or the context of choice, *ceteris paribus*, helps to determine a different behavior. Among the violations of the expected utility paradigm, that have a psychological motivation and which are important in the financial choices, in particular there is the loss aversion, which implies the risk aversion (Kahneman and Tversky, 1984)[24]. Kahneman and Tversky have shown, for example, that many of the risks of little importance are given disproportionate weight, but also that the losses and future earnings are not treated symmetrically.

[24] For most individuals, in fact, the motivation to avoid a loss is greater than the motivation to make a profit. This general psychological principle, which is connected to a kind of survival instinct, causes that the same decision gives rise to opposite choices depending on whether the results are presented to the subject as losses rather than as loss of earnings. This type of evidence has led Kahneman and Tversky (1979, 1984) to develop the theory of prospects within the cognitive-behavioral approach.

The second approach, derived from Simon's work, is the ''fast and frugal heuristics" program (Gigenzer, Goldstein, 1996; Todd, Gigerenzer, 2003). These fast heuristics are conscious processes, accessible to introspection in humans. Following Simon's notion of satisfying, Gigenzer and Goldstein have proposed a family of algorithms based on a simple psychological mechanism: one-reason decision making. These fast and frugal algorithms violate fundamental tenets of classical rationality: they neither look up nor integrate all information (Gigenzer, Goldstein, 1996). The heuristics are determined by a trade-off between the limits of the human mind and the computing performance required by complex problems. The psychology of choice is to codify these heuristics in humans, to help apply them in situations where they work well.

4. Behavioral finance

4.1. Asset allocation

The theory of expected utility applies to financial investment decisions: it is the asset allocation theory, where the investment decision possibilities are represented by a function with the different choice of risky investments taking into account their expected value.

If A and B are two risky alternatives, the expected utility theory enables to state that

$$A < B \text{ if and only if } E[u(A)] < E[u(B)]$$

where the symbol < indicates the preference of B with respect to A and the function $u(.)$ represents the utility function.

The risky choice of B is preferred to the risky choice of A if the expected value of the utility function B is higher than the expected value of the utility function A. The utility function is increasing and concave with respect to risk aversion.

The assumptions are:

- Independence axiom: $A < B$ *implies* $\alpha A + (1-\alpha)C < \alpha B + (1-\alpha)C$*, for every C*

- Equivalent probability. Assume a lottery that gives value W_H and W_L. The probability of W_L is p. An investor is risk averse if:

$$pu(WH) + (1-p)\,u(WL) < u(pWH + (1-p)WL)$$

Consider a change of probability from p to q

$$qu(W_H) + (1-q)\,u(W_L) = u(qW_H + (1-q)W_L)$$

if changing the probability from p to q, then the final relation remains unchanged, it means that, regardless of the various probabilities, the probability that will result a lottery rather than another is equivalent.

- Certain equivalent. Let us assume a lottery. This can give as a result WH and WL. The probability of getting WH is equal to p. If

$$pu(WH) + (1-p)\,u(WL) < u\,(pWH + (1-p)WL)$$

thus investor is risk averse.

The certain equivalent is WCE if

$$pu(WH) + (1-p)u(WL) = u\,(WCE)$$

i.e. if the mathematical relation which identifies the risk averse investor is equal to WCE, then WCE is the certain equivalent, whose utility is the expected utility of the lottery.

4.2. Expected utility and risk aversion

Let's examine the relation between expected utility and risk aversion. Consider a lottery W, with mean E(W); this is a safe return, but not winning the lottery. It is a certain event. An individual is called risk-neutral if it indifferent to perceive definitely a sum equal to E(W) or the lottery W. Then:

$$E[u(W)] = u\,(E\,(W)).$$

An individual is instead risk averse if he prefers the sum E (W) to the lottery W. Then

$E[u(W)] < u(E(W))$. *i.e.* an individual is risk averse if he prefers a smaller but certain sum.

To measure the degree of risk aversion we try to determine a value π such that
$E[u(W)] = u(E(W) - \pi)$.

With the Taylor's expansion we can verify that

$\pi = \frac{1}{2}(-u''/u')\, Var(W)$ *Where* u' and u" are the first and the second derivate of the utility function.

The literature on financial behavior has set forth that the various utility functions are different for the different behavior of risk aversion, relative or absolute, to changes in wealth. There are several utility functions that differ according to the change in risk aversion with respect to wealth. That is, it has been shown that individuals, on the basis of their wealth, have a different risk aversion.

Among the utility functions we can cite the quadratic utility function, which has the unrealistic feature of risk aversion that increases with wealth. Another is the exponential function or constant absolute risk aversion (CARA). A third is the power utility or constant relative risk aversion (CRRA). A fourth is the logarithmic utility function, an extreme case of the CRRA. Last the hyperbolic absolute risk-aversion (HARA). This function represents the most general case, to measure the linear risk tolerance based on wealth. It is the function commonly used to measure risk aversion.

4.3. Behavioral finance: anomalies and biases

From the seventies onwards there has been the growth of a new branch of finance: the behavioral finance, which in itself combines aspects of cognitive psychology and financial theories in the strict sense. In practice this new approach seeks to explain the so-called financial market "anomalies" by analyzing the behavior of economic agents. The adoption of heuristics by individuals is necessary to solve the problems of everyday life, but in the financial sector it can lead to biases which have proved very expensive.

In their Prospect theory Kahneman and Tversky (1979, 1984, 1986) criticize the classical approach of the expected utility and offer a theory based on the existence of a 'Reference Point'. They argue that any individual has a deformation of the probability, which is different between gains and losses and, moreover, the individual has aversion to losses. A loss, in fact, is more weighted by a psychological point of view than a gain.

Their utility function is:

$u(r) + w^{+}(p)\,(u(Wh) - u(r)) - \lambda w^{-}(1 - p)(u(r) - u(WH))$

r = reference point
$w^{+}(p)$ = deformation of the positive probability in its functional form.
λ = risk aversion.

According to this theory the utility function of a given asset compared to a reference point is given by the utility function of the reference point itself, plus the deformation of the positive probability (which represents the probability of the pleasant event) of the utility function (Wh) less the deformation of the negative probability (the probability that the hoped event does not happen) of the utility of the reference point less the sought event.

Another issue which is relevant for the decisions of financial investment is the rejection, based on empirical tests (Thaler, 1980) of the postulate of the theory of rational choice that preferences are invariant with respect to the capital position of the individual at the time. According to Thaler (1980), your choices are influenced by the "endowment effect" if you are brought to ask for goods in your possession more than you would be willing to pay for it, if you do not already own that good.

Kahneman, Knetsch and Thaler (1990) also carried out a significant experiment based on the "endowment effect", in which they demonstrate that the individuals feel a great sorrow when they lose the objects they possess, more than the pleasure would cause them to acquire those same objects, if they do not already possess them. So the "endowment effect" is an anomaly that causes a *status quo* bias (a preference for the current state that biases the individual against both buying *and* selling his object). The "endowment effect" is also connected to a particularly pervasive phenomenon: the "loss aversion", for which the disutility of a loss is greater than the utility of a win of the same size[25].

[25] Loss aversion is a core aspect of agent's reference-based preferences. The sensation of loss relative to status quo and other reference points looms very large relative to gains (Rabin, 2002, p.9).

Thus, the literature of behavioral finance includes the lack of symmetry between decisions to acquire and maintain resources and the strong aversion to the loss of some (emotionally) valuable resources that could be completely lost. In the field of behavioral finance, the loss aversion appears to manifest itself in the investor behavior as an unwillingness to sell assets or other securities, if doing so forces the investor to achieve a nominal loss (Genesove and Mayer, 2001). This loss aversion helps to explain in particular why housing market prices do not adjust downwards during periods of low demand.

The models of behavioral finance, used in the valuation of assets, usually criticize the idea of the Efficient Markets Hypothesis, which maintains that market prices incorporate all information rationally and instantaneously (Fama, 1970, 1991). In other words, in this idea based on informational efficiency of markets, a market is informational efficient if at all times the stock prices fully and correctly reflect all the available information. However, the Efficient Markets Hypothesis has been challenged by behavioral economics and finance, since these alternative approaches argue that markets are not rational, but are driven by sentiments as fear and greed. Contributions of behavioral finance have discovered some anomalies, inspired by the hypothesis that some investors in assets have limited rationality (Camerer, Loewenstein, 2003), which, for instance, would produce excess return in financial markets. De Bondt and Thaler (1985), in particular, have shown that bonds, characterized by particularly high yields (so-called winners), record in the aftermath the worst yield and vice versa. This depends on investors' overreaction to an event. Over the time the investors realize the error and correct their assessments causing a reversal of returns.

Finally, Thaler and Shefrin (1981), who gave major contributions to behavioral finance, present their behavioral life-cycle theory arguing that economists who wish to analyze the consumption-saving decision must address the bounded rationality and impatience of consumers. The behavioral-life cycle theory models consider consumers as responding to psychological limitations by adopting rules-of-thumb, such as mental accounts, that are used to constrain the decision making of the myopic agent.

An alternative approach to behavioral economics that can be applied to financial marks is that offered by neuroeconomics, a discipline at the turn of neuroscience and economics (Camerer, Loewenstein, Prelec, 2005), which aims at studying the processes underlying the decision-making choices and that reveals what instincts are activated when you have to do with the risk, the gains and losses[26]. Neuroeconomics tries to offer a solution through an additional set of data obtained via a series of measurements of brain activity at the time of decisions. Neuroeconomics theory proposes to build brain-based models capable of predicting observed behavior (Brocas, Carrillo, 2010). The underlying idea of neuroeconomics is that the brain is a multi-system entity with restricted information and conflicting objectives characterized by bounds of rationality, so the decision-maker must be modeled as an organization. This relatively new approach can be considered another development of Simon's intuitions.

Conclusion

The financial crisis has raised many questions and created new problems for economic theory. It is not all certain that the mathematical algorithms devised by the classical theory can predict in time when the international financial crises occur, but, as this paper tried to argue, we can enrich our knowledge of the complex reality of financial markets through the fertile contribution of behavioral economics.

Thus the present contribution discussed the criticism to the classical theory of rational choice and to the expected utility coming from behavioral economics which highlights the link between psychology and economics (Rabin, 2002). Moreover, the concept of bounded rationality, as formulated by Herbert Simon, has been proposed and confronted with the classical notion of perfect rationality. It has been underlined the relation between bounded rationality and procedural rationality which is the form of psychological rationality that constitutes the basic concept of Simon's behavioral theory. Failures of classical theory of rational choice, anomalies and biases in the behavior of the economic agents in financial markets has been pointed out, although the critical part of the behavioral theory seems more convincing than the positive and proactive part of the same theory, leaving some degree of indeterminacy in defining solutions.

A possible answer to the questions posed by the financial crises could come from neuroeconomics. The latest crisis has been largely the result of a collapse of financial markets and confidence; according to neuroeconomics theory, it would be an effect of automatic processes and unconscious decisions much more than deliberate and conscious decisions. Thus, researchers in neuroeconomics argue that to prevent market crises

[26] Neuroeconomics have exploded with discoveries, because of advances in imaging techniques which permit more precise temporal and spatial location of brain activity.

may be feasible in the future, thanks to new financial models that take into account the neurocognitive constraints, i.e. mechanisms put in place by the brain in response to certain environmental stimuli, and the influence of emotions on the choices of investment.

References

[1] Allais M. 1953. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine, *Econometrica* 21: 503-546.

[2] Arrow K.J. 1971. *Essay in the Theory of Risk - Bearing*, Amsterdam, North Holland Pub. Co.

[3] Barros G. 2010. Herbert A. Simon and the concept of rationality: boundaries and procedures, *Brazilian Journal of Political Economy* 30(3): 445-472, July - September.

[4] Brocas I., and Carrillo J.D. 2010. Neuroeconomics theory, using neuroscience to understand bounds of rationality, VoxEU.org, March.

[5] Camerer C.L., and Loewenstein G. 2003. Behavioral Economics: Past, Present, Future, in Camerer C.L., Loewenstein G., Rabin M. (eds.), *Advances in Behavioral Economics*, Princeton, Princeton University Press, 2003: 3-51.

[6] Camerer C.L., and Loewenstein G., Prelec D. 2005. Neuroeconomics: how neuroscience can inform economics, *Journal of Economic Literature*, XLIII: 9-64.

[7] Constantinescu M. 2010. Neuroeconomics and decision making process, *Theoretical and Practical Research in Economic Fields* 1(2): 210-218.

[8] De Bondt, W.F., and Thaler R. 1985. Does the stock market overreact?, *The Journal of Finance* 40(3): 793-805.

[9] Ellsberg, D., 1961. Risk, ambiguity, and the Savage axioms, *Quarterly Journal of Economics* 75(4): 643–669.

[10] Fama E.F. 1970. Efficient Capital Markets: a Review of Theory and Empirical Work. *Journal of Finance* 35: 383-417, May.

[11] Fama E.F. 1991. Efficient Capital Markets: II. *Journal of Finance* 46(5): 1575-1617.

[12] Friedman M., and Savage L. 1948. Utility analysis of choices involving risk, *Journal of Political Economy* 56 (4): 279-304.

[13] Genesove D., and Mayer C. 2001. Loss aversion and seller behavior: Evidence from the housing market, *The Quarterly Journal of Economics* 116 (4): 1233-1260.

[14] Gigerenzer, G., and Goldstein D.G. 1996. Reasoning the fast and frugal way: models of bounded rationality, *Psychological Review* 103 (4): 650–669.

[15] Graziano M., and Schilirò D. 2011. Rationality and choices in economics: behavioral and evolutionary approaches, *Theoretical and Practical Research in Economic Fields* 2(2): 183-196.

[16] Kahneman D. 2003. Maps of bounded rationality: psychology for behavioral economics. *The American Economic Review* 93(5): 1449–75.

[17] Kahneman D., and Thaler R.H. 2006. Anomalies: Utility Maximization and Experienced Utility, *Journal of Economic Perspectives* 20(1): 221-234.

[18] Kahneman D., and Tversky A. 1979. Prospect theory. An analysis of decision under risk, *Econometrica* 47(2): 263-292.

[19] Kahneman D., and Tversky A. 1984. Choices, values and frames, *American Psychologist* 39: 341-350.

[20] Kahneman D., Knetsch J.L., and Thaler R.H. 1990. Experimental tests of the endowment effect and the Coase theorem, *Journal of Politica Economy* 98(6): 1325-1348.

[21] Montesano A. 2005. La nozione di razionalità in economia, *Rivista italiana degli economisti,* n.1.

[22] Novarese M., Castellani M., and Di Giovinazzo V. 2009. Procedural rationality and happiness, *MPRA Paper* 18290, October, University Library of Munich, Germany.

[23] Rabin M. 1998. Psychology and economics, *Journal of Economic Literature* 36(1): 11-46.

[24] Rabin M. 2002. A perspective on psychology and economics, *European Economic Review*, 46(4-5): 657-685.

[25] Rubinstein A. 1998. *Modeling Bounded Rationality*. Cambridge: MIT Press.

[26] Schilirò, D. 2005. Economia della conoscenza, istituzioni e sviluppo economico, *MPRA Paper* 31492, University Library of Munich, Germany.

[27] Schilirò D. 2011. Decisioni e razionalità in economia, *MPRA Paper* 29477, University Library of Munich, Germany.

[28] Schilirò D. 2012. Bounded rationality: psychology, economics and the financial crisis, *MPRA Paper* 40280, University Library of Munich, Germany.

[29] Schilirò D., and Graziano M. 2011. Scelte e razionalità nei modelli economici: un'analisi multidisciplinare, *MPRA Paper* 31910, University Library of Munich, Germany.

[30] Simon H.A. 1955. A behavioral model of rational choice, *The Quarterly Journal of Economics,* 69(1): 99-118, February.

[31] Simon H.A. 1956. Rational choice and the structure of the environment, *Psychological Review*, 63 (2): 129 – 138, March.

[32] Simon H.A., 1957. *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*, New York: John Wiley and Sons.

[33] Simon H.A., 1959. Theories of decision making in economics and behavioral science, *American Economic Review*, 49 (3): 253-283.

[34] Simon H.A., 1972. Theories of bounded rationality, in McGuire C.B., Radner R. (eds.), *Decision and Organization*, Amsterdam, North Holland Pub. Co.

[35] Simon H.A., 1976. From substantive to procedural rationality, in Latsis S.J. (ed.), *Method and Appraisal in Economics*, Cambridge: Cambridge University Press, pp. 129-148.

[36] Simon H.A. 1979. Rational decision making in business organizations [Nobel Memorial Lecture], *American Economic Review,* 69(4): 493-513, September.

[37] Simon H.A. 1991. Bounded rationality and organizational learning, *Organization Science,* 2 (1): 125–134.

[38] Simon H.A. 1997. *An Empirically Based Microeconomics*, Cambridge, Cambridge University press.

[39] Thaler R.H. 1980. Towards a positive theory of consumer choice. *Journal of Economic Behavior and Organization,* 1: 39-60.

[39] Thaler R. H., and Shefrin H. 1981. An economic theory of self-control. *Journal of Political Economy,* 89(2): 392–406.

[40] Todd P.M., Gigerenzer G. 2003. Bounding rationality to the world, *Journal of Economic Psychology,* 24: 143-165.

[41] Tversky A., and Kahneman D. 1974. Judgment under uncertainty: heuristics and biases, *Science,* 185: 1124-31, September.

[42] Tversky A., Kahneman D., 1986. Rational choice and the framing of decisions, *Journal of Business,* 59(4): S251-S278, October.

[43] Von Neumann J., and Morgenstern O. 1944. *Theory of Games and Economic Behaviour*, Princeton, Princeton University Press.

ASERS Publishing
Volume III Issue 2(6) Winter 2012
DOI: 10.2478/v10261-012-0008-z

THE CRISIS OF THE EUROPEAN MONETARY UNION[27]

Peter Alfons SCHMID
Catholic University Eichstätt-Ingolstadt, Germany
peter.a.schmid@kuei.de

Abstract:
The Euro Zone (EZ)'s economies are under great stress since last decade's financial crisis. Diverging interest rates, high debt burdens and sluggish growth in several member countries led to various rescue activities. Nevertheless, financial markets have still not calmed and the break-up of the EZ is discussed openly. Contrary to the popular belief I show that the Euro itself has been a success story so far but that the EZ suffers under a debt crisis and huge economic imbalances. An overhaul of the EZ's institutional framework is necessary.

Keywords: Financial crisis, euro crisis, current account imbalances, Monetary Union.

JEL Classification: E42, F34.

1. Introduction

The future of the common currency in the EZ is at risk and there is an open debate in politics and the academic world whether the European Monetary Union (EMU) can survive in its current form (Subacchi - Pickford 2012, 10). Even before its start in 1999 there were many critics that warned that the greatest experiment in monetary history is set to fail due to the violations of the required conditions established by Mundell 1961 for the functioning of a common currency area; it was even suggested that the EZ might carry political risk that was meant to be eliminated by the common currency (Feldstein 1997, 41).

Some political observers now speak of peace and war concerning the EMU and that its success is pivotal for the survival of the European integration. The former German chancellor Kohl even stated the issue of containing a potentially dangerous Germany within the EMU as a receipt for peace in Europe. Unfortunately, this rhetoric complicates or even prevents a serious debate on the current state of the EMU and its associated challenges as it is well known that *the truth dies first in times of war*. This article's aim is to give an unbiased summary of the *status quo* of the Euro and shed light into the roots of the current turmoil on financial markets. I conclude that the EMU is in deep economic crisis but that there is no currency crisis although the common currency might have been favorable to the economic imbalances within the EZ.

2. The performance of the Euro. Hard facts on its history

Most commentators and academics speak of a Euro crisis now (EEAG 2012, 8). Nevertheless there are voices that question the existence of a currency crisis. In fact, a crisis of the Euro cannot be assessed prima facie based on its performance since its birth. I follow and enrich Klodt 2011 in my argumentation.

First, the development of the Euro exchange rates does not look alarming. Compared to other major currencies the real effective exchange rate of the Euro even displays certain strength (Figure 1a). Whereas the US-Dollar, the Pound Sterling and the Japanese Yen lost value the Euro gained comparatively.

[27] A strongly shortened and older version of the paper in German can be read at www.oekonomenstimme.org.

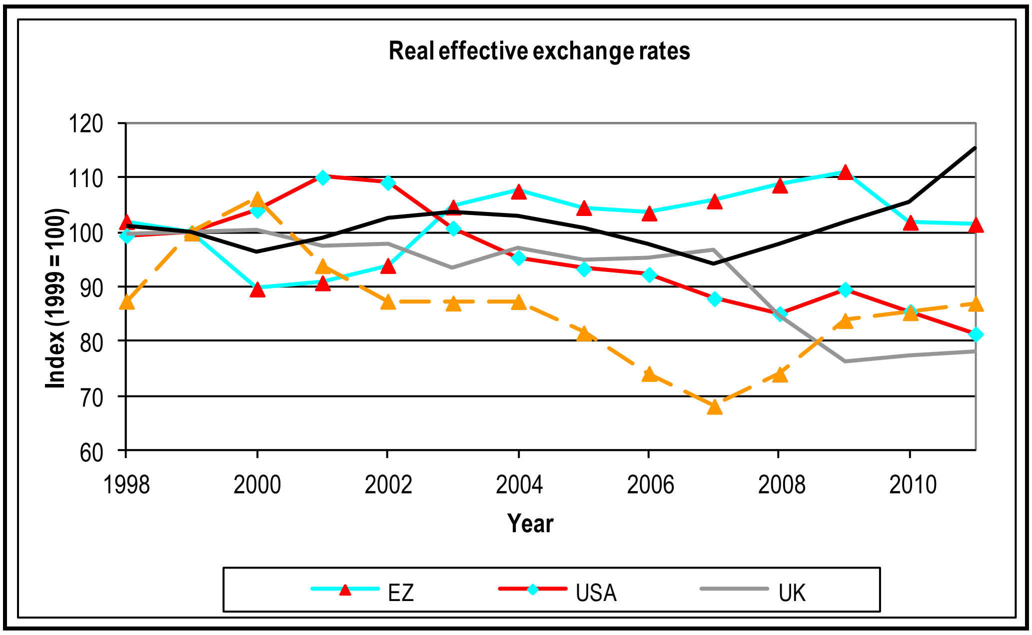


Figure 1a. Real effective exchange rates based on 41 trading partners, deflator: consumer price indices

Source: Eurostat.

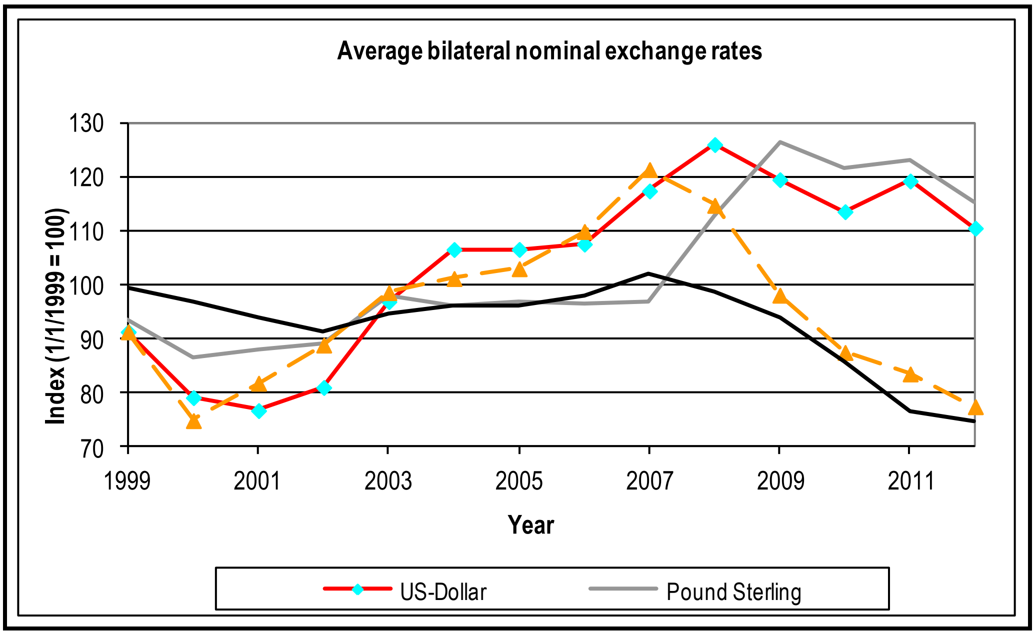


Figure 1b. Bilateral exchange rates of the Euro, forecast for 2012

Source: Eurostat.

Moreover, Figure 1b shows major fluctuations of bilateral Euro exchange rates. But these are not unusual by historical standards. The US-Dollar's exchange rate against the Deutsche Mark, for example, fluctuates much stronger around its average during the period 1971 to 1999. In addition the Euro appreciated against the US-Dollar and the Pound Sterling considerably after its decline in the first years of the currency union. The sharp loss against the Swiss Franc and the Japanese Yen only happened after the outbreak of the financial crisis and at least the Swiss Franc must be seen as a safe haven which naturally gains in turbulent times. As the Euro also gained value against major trading partners from emerging countries, e.g. +16.8% against the Chinese Yuan and

+49.4 against the Indian Rupee from 2000 to 2011, the price of the Euro in units of other currencies does not suggest a currency crisis but describes its relative strength.

Second, overall public debt in the EZ increased considerably from 72.8% in 1998 to 87.2% in 2011 (Figure 2). But it is remarkable that until the outbreak of the financial crisis a moderate decrease or stabilization of public debt was accomplished; it only increased afterwards by roughly 20 percentage points. In addition, EZ public debt is still well below the public debt in the United States and Japan that experienced debt hikes from 67.2% to 102.9% and from 183.0% to 229.8% in the time period 2007 to 2011. Thus, a global debt crisis in developed economies might describe the current situation more accurately than a currency crisis in the EZ.

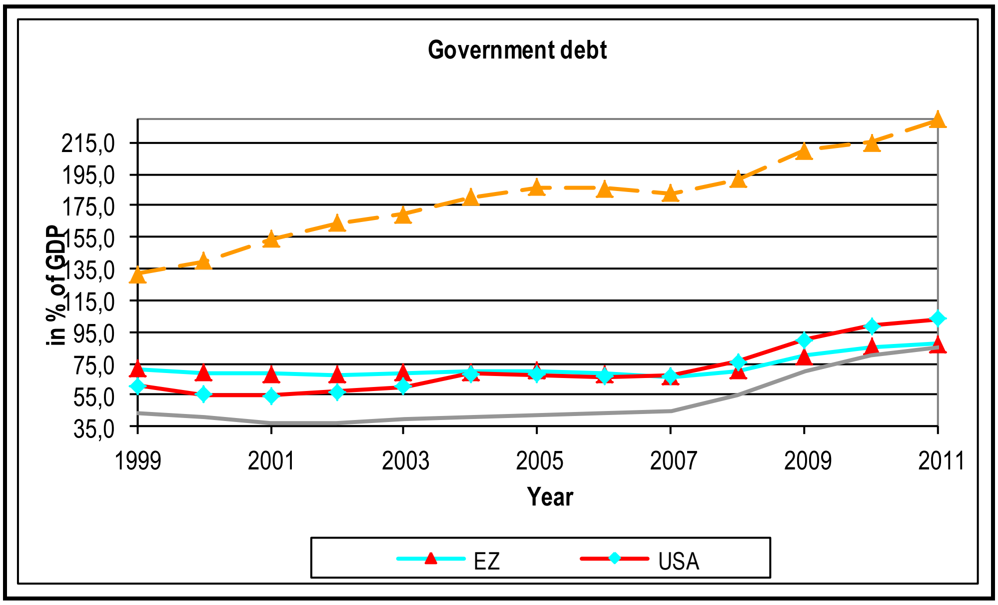


Figure 2. Public Debt in the Euro Zone, USA, UK and Japan

Source: Eurostat, IMF.

Third, not only public debt matters, of course. In addition total debt of a country or currency area against the rest of the world must be analyzed. Therefore, current account balances (Figure 3) are shown:

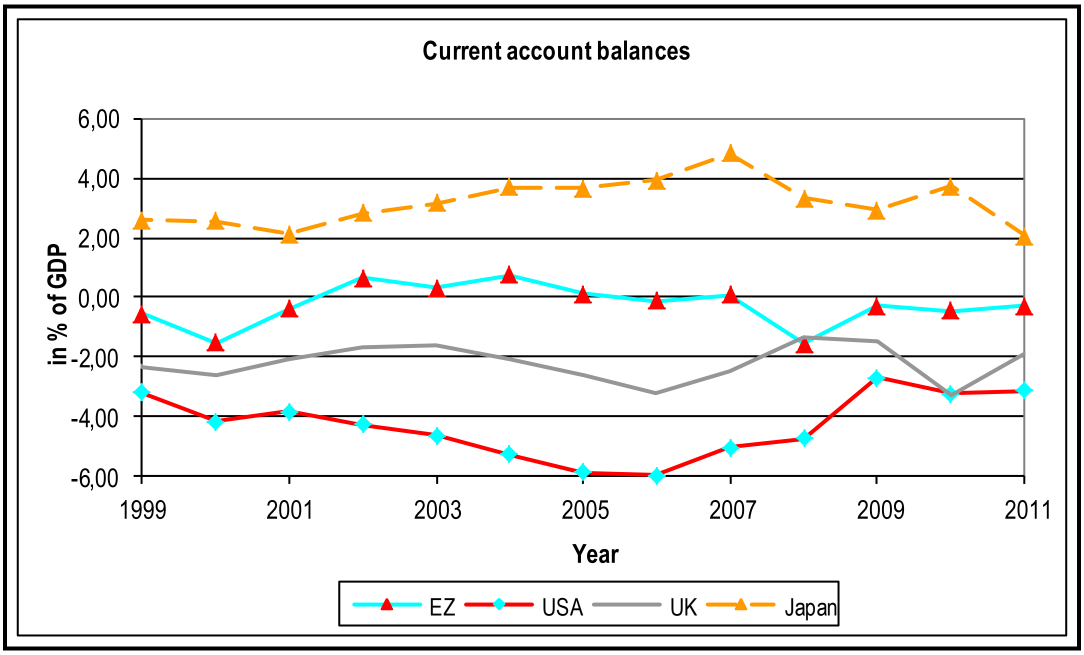


Figure 3. Current account balances in the Euro Zone, USA, UK and Japan

Source: OECD, EC European Economic Forecast Spring 2012.

The EZ looks more or less balanced – in contrast to the United States that still shows high deficits. Since 1999 the EZ's current account balance has been oscillating around zero whereas the US current account was negative for the entire period.

Finally and fourth, the official goal price stability was accomplished in the EZ in a remarkable way. The inflation rate was and is close to its official target rate of 2 % (Figure 4). Moreover, it was more stable than in the United States and still is well below the one in the UK and the USA.

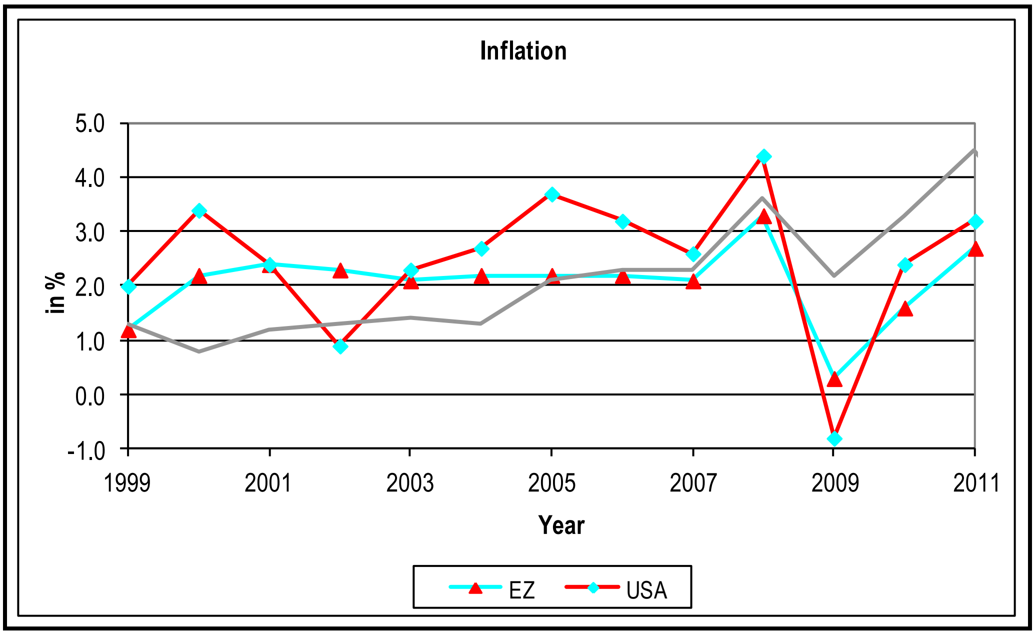


Figure 4. Inflation Rates in the Euro Zone, USA and UK

Source: Eurostat, European Economic Forecast Spring 2012.

Thus, "although there is no generally accepted formal definition of a currency crisis" (Krugman 2000, p. 1), an overall currency crisis must be denied based on the presented hard facts. A closer look at the economic situation in the EZ, however, shows deep trouble because of its imbalances and structural problems. The Euro has not been the reason therefore but created a favorable environment for the diametric development within the EZ.

3. Financial crisis, real economic development and Imbalances within the Euro Zone

The long-term interest rate levels within the EZ narrowed considerably right before and after the start of the EMU (Figure 5) and led to real economic convergence as intended (EEAG 2012, 9).

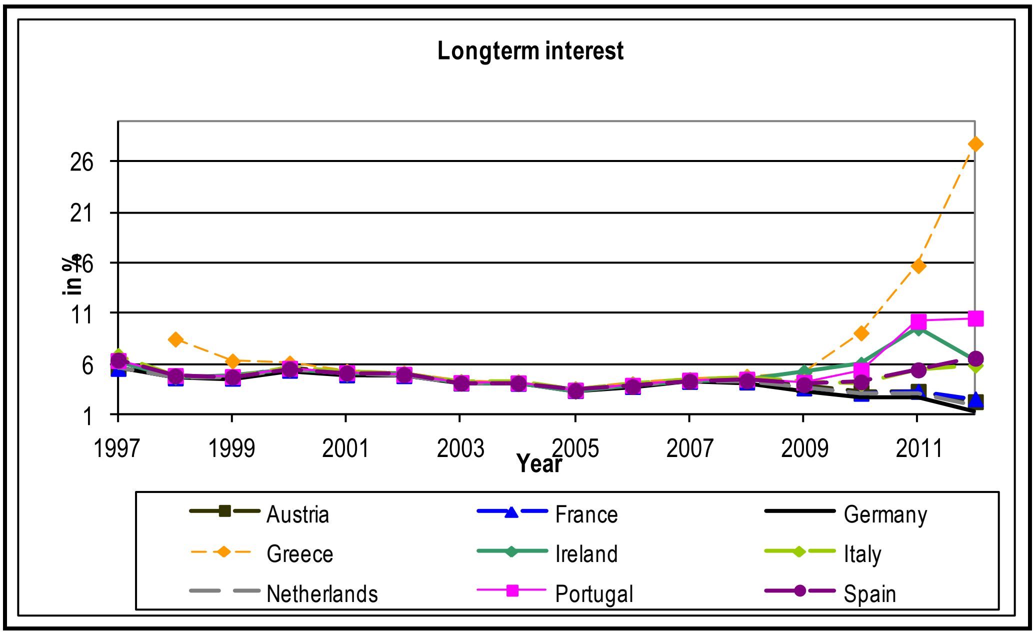


Figure 5. Long-term interest rates in the Euro Zone, latest value: June 2012

Source: OECD

These low interest rates in Greece, Ireland, Portugal and Spain (GIPS) countries allowed the financing of credit booms. Even worse, as the savings rates in these countries decreased the credit booms were financed by private capital imports (Neubäumer 2011, p. 828). Apparently, private investors must have lost faith in the creditworthiness of these countries and the sustainability of the economic catch-up process with the outbreak of the financial crisis as interest rates diverged again. At the time of this writing the interest rate spreads reached alarming levels for some countries whereas some countries like Germany can borrow at historically low rates.

3.1 Real Economic Development within the Euro Zone

Whereas Germany formed the lower bound of economic growth in the EZ and even was labeled the sick man of Europe (The Economist 2003) other countries like Spain and Ireland were seen as primes and examples for economic success (Bergheim 2007, 3 and Sweeney 2008, p. 4). Figure 6 shows how GDP per capita converged by differing growth rate patterns within the EZ until the outbreak of the financial crisis.

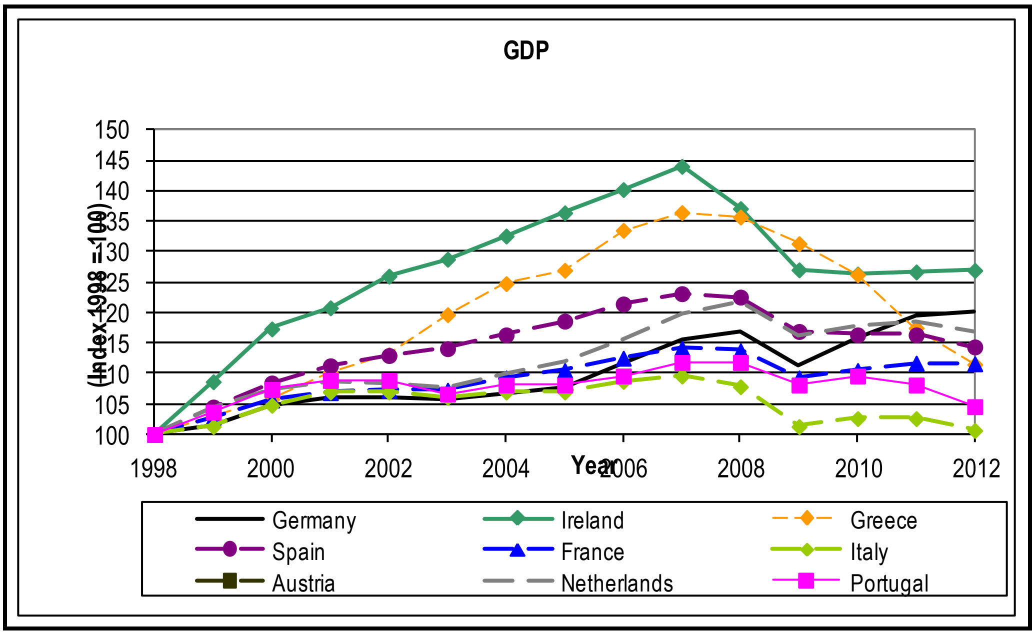


Figure 6. Development of GDP per capita in the Euro Zone, forecast for 2012

Source: Eurostat.

In its aftermath the distribution of economic growth in the EZ reversed completely. Germany, after its structural reforms in the first half of the last century and only moderate growth of unit labor costs, now leads the EZ with strong GDP growth figures. Of course, this development is also expressed in unemployment figures (Figure 7):

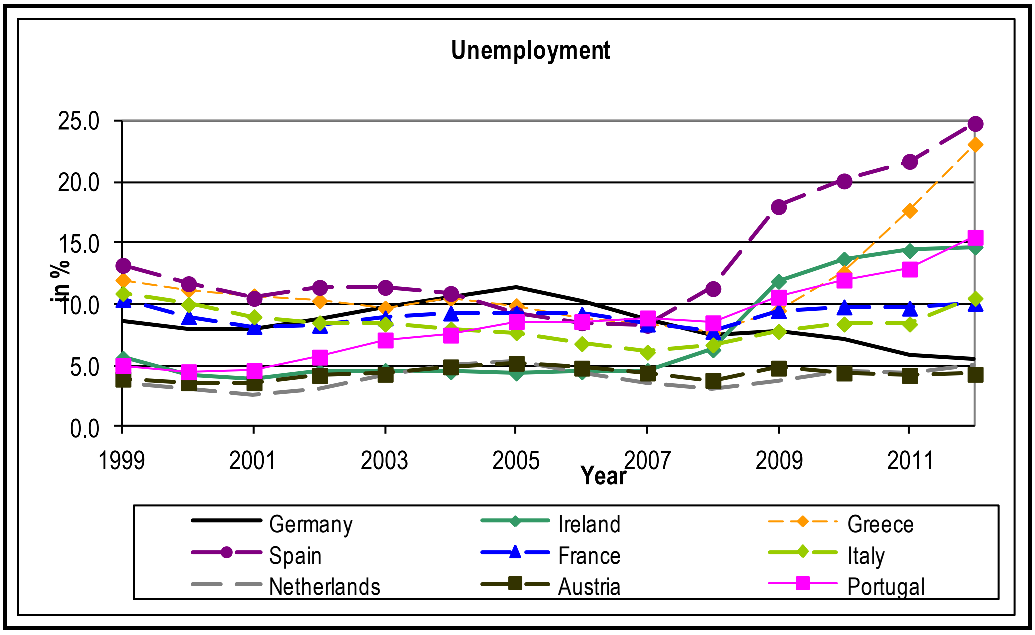


Figure 7. Unemployment rates in the Euro Zone, latest value: May 2012

Source: Eurostat.

Whereas there is full employment in some parts of Germany with labor scarcity for skilled industry workers, unemployment in Spain is unsustainable high, especially among the young (Bräuninger - Majowski 2011, p. 2). Although labor migration begins to improve within the EZ and let to a net inflow amounting to 280,000 into Germany (Deutsche Bundesbank 2012, p. 16) there is still a strong home bias and a long way to the kind of labor mobility like in the United States (Shambaugh 2012, p. 13). Thus, those imbalances do not dampen themselves sufficiently endogenously so far.

3.2 Structural Problems and Imbalances within the Euro Zone

This economic heterogeneity within the EZ is summarized in Table 1 that shows the public, private and net foreign debt positions in addition to GDP growth and unemployment figures.

Table 1. Economic situation in the Euro zone

	GDP-Growth per capita (1999-2007)	GDP-Growth per capita (2008-2012)	Unemployment (05/2012)	Public Debt (2011, % of GDP)	Private Debt (2010, % of GDP)	Net Foreign Position (2011, % of GDP)
Austria	20,5%	1,1%	4,3%	72,2%	165,7%	-3,4%
Belgium	18,0%	-1,8%	7,1%	98,0%	232,3%	57,8%
Finland	33,5%	-4,0%	7,6%	48,6%	182,8%	16,0%
Germany	15,5%	4,1%	5,5%	81,2%	127,2%	36,1%
Luxembourg	38,9%	-8,6%	5,4%	18,2%	253,9%	85,1%
Netherlands	19,9%	-2,6%	5,1%	65,2%	225,3%	41,3%
France	14,2%	-2,2%	10,1%	85,8%	159,8%	-15,9%
Italy	9,6%	-8,2%	10,5%	120,1%	126,4%	-20,6%
Spain	23,2%	-7,1%	24,8%	68,5%	226,6%	-91,6%
Greece	36,5%	-18,3%	23,1%	165,3%	125,2%	-79,5%
Ireland	44,1%	-11,9%	14,7%	108,2%	341,3%	-97,6%
Portugal	11,9%	-6,5%	15,5%	107,8%	250,2%	-102,7%
Cyprus	23,9%	-8,8%	10,7%	71,6%	278,1%	-80,7%
Estonia	90,4%	-7,6%	10,1%	6,0%	176,7%	-57,7%
Malta	7,6%[28]	3,9%	6,1%	72,0%	212,0%	5,7%
Slovakia	54,6%	10,1%	13,7%	43,3%	69,0%	-64,5%
Slovenia	46,4%	-6,5%	8,2%	47,6%	128,8%	-41,8%

Source: Eurostat.

Besides the differences in growth rate patterns the debt positions are worrying. Nine out of 17 countries show combined debt levels over 250 % and seven countries have a net foreign position below -50 % of GDP. Especially the countries that already receive international help are among both groups, a fact emphasizing the uncertainty or doubts about near economic revival. Only Italy with its low private debt, a combined relative debt level comparable to the French one and its relatively high net foreign position looks promising. The development of relative debt levels over time is displayed in Figures 8 and 9:

[28] For Malta accumulated growth for 2001-2007 is displayed.

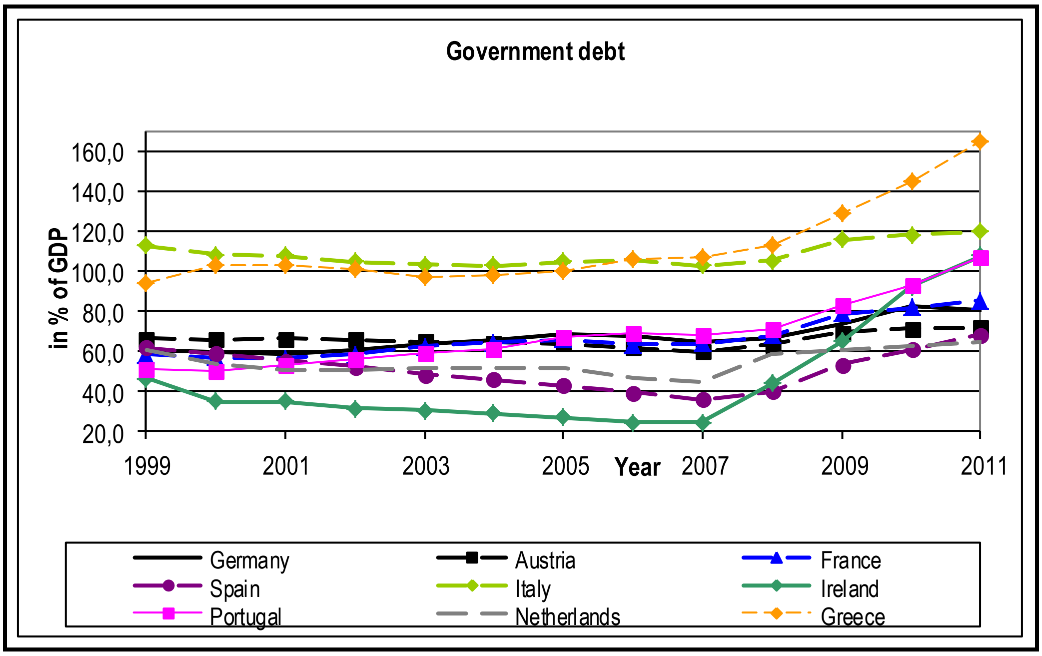


Figure 8. Public debt levels within the Euro Zone

Source: Eurostat.

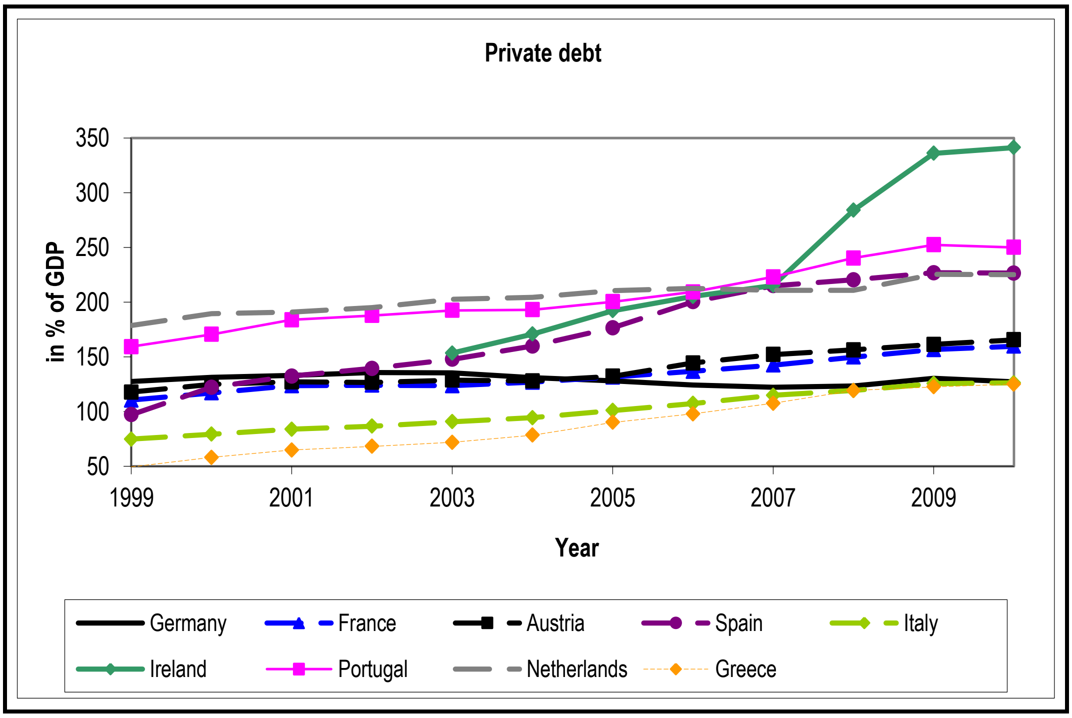


Figure 9. Private debt levels within the EZ

Source: Eurostat.

Stimulated by low long-term interest rates and accompanying fiscal incentives and investment booms Spain and Ireland could reduce their public debt levels considerably before the outbreak of the financial crisis. Even Greece, Italy and Portugal were successful in reducing or stabilizing their public debt levels. Private debt, however, already has increased steadily before and now poses severe problems on the banking sector and hinders private consumption and thus a fast economic turn-around. Moreover, a large part of this debt is foreign-held (Figure 10):

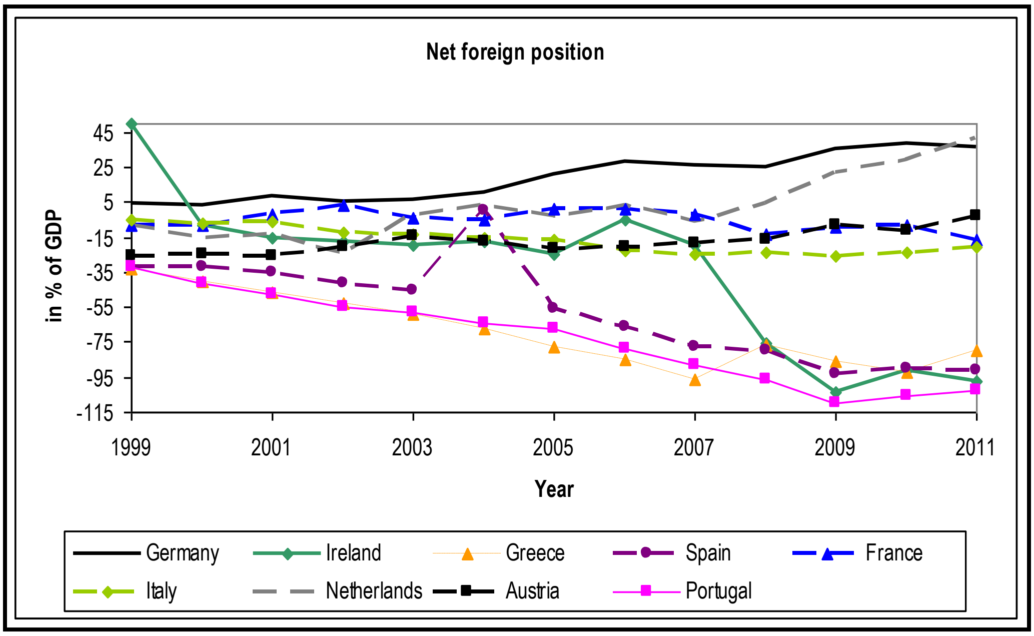


Figure 10. Net foreign position in the Euro Zone

Source: Eurostat.

Again, the GIPS countries have the highest foreign indebtedness, close or even over 100 % of GDP. Whereas the decline of Greece, Portugal and Spain was steadily, Ireland's foreign debt only surged with the outbreak of the financial crisis.

The development of the net foreign position can be interpreted by current account deficits that were financed by private capital inflows (Mayer *et al.* 2011, p. 32). It becomes apparent that the balanced current account of the EZ as a whole is bought by huge imbalances within. On the one hand there are Germany, Austria and the Netherlands with their strong and competitive export industries and on the other hand there are the GIPS countries (Thompson 2011). Obviously these imbalances cannot continue indefinitely as "no country can go on forever covering by new lending a chronic surplus on current account without eventually forcing a default from the other parties." (Keynes 1946, p.184).

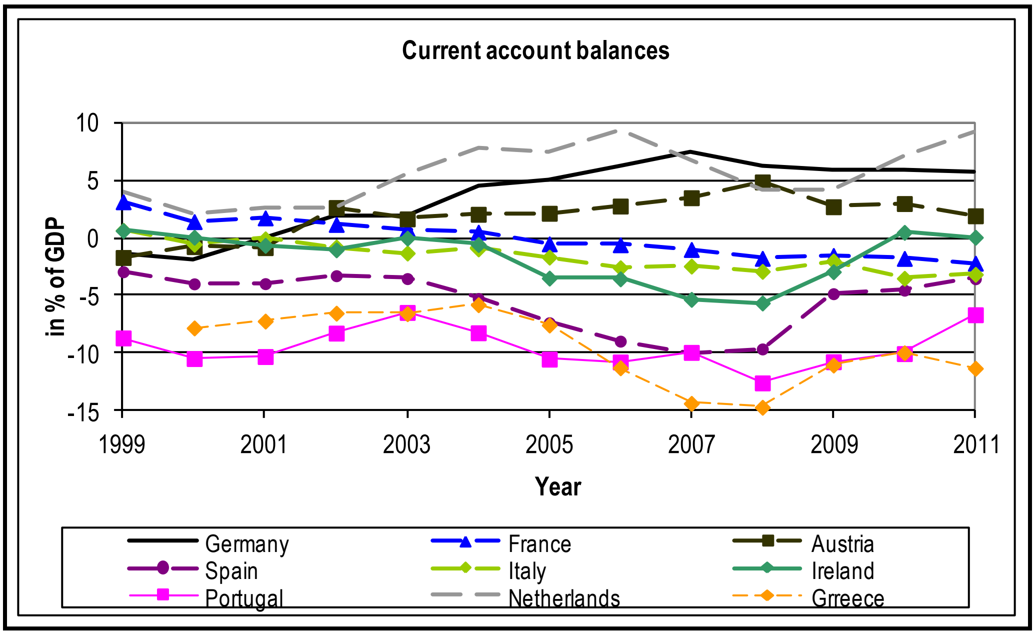


Figure 11. Current account development in the Euro Zone

Source: OECD, EC European Economic Forecast Spring 2012.

As a result the EMU economy is widely imbalanced and in deep economic crisis (Milbradt 2011, 39). Unfortunately, the economic catch-up process in the GIPS countries before the outbreak of the financial crisis was not lasting, but financed to a large part consumption and real estate booms (Neubäumer 2011, 828). Moreover, the unit labor cost rose considerably and necessary structural reforms were postponed leading to a further loss of competitiveness, sclerotic labor markets and low competition on goods markets (Barkbu *et al.* 2012, 20); although Italy did not experience this unsustainable economic boom it also suffers under the loss of its competitiveness.

4. The crisis of the European Monetary Union and stabilization activities

The common currency is not the reason for the economic imbalances and loss of competitiveness in Southern and peripheral countries. Artificially low interest rates led to the unsustainable booming years and the take-over of large combined and foreign-financed private and public debt levels. The common monetary policy, however, nourished these imbalances by too low interest rates (Neubäumer 2011, p. 828). With the outbreak of the financial crisis these imbalances became clear, the housing bubbles in Ireland, Portugal and Spain burst and private capital flows reversed their direction and even capital flight occurs leading to an amplification of economic recession.

Consequently, stabilization and rescue activities have been determined upon and implemented. As an immediate first-aid the ECB established three strands of rescue activities. First, the key interest rate for the main refinancing operations was reduced from 4.25% in July 2008 to 0.75 % since July 2012. Moreover, the minimum rate tender was changed to a fixed rate tender system with full allotment and the minimum requirements for collaterals were reduced substantially. Second, the Securities Markets Program (SMP) was established according to which the ECB currently holds government bonds of countries under financial stress for 214 billion EUR (Europäische Zentralbank 2012, S. 48). At the time of this writing a new program was decided upon according to which the ECB intervenes on the secondary market in favor of countries that applied for fiscal help under the rescue umbrellas EFSF or ESM and thus commit themselves to economic reform. Third, it allowed the build-up of huge TARGET2 balances. Until the outbreak of the financial crisis this balances were of minor magnitude, oscillated around zero and thus did not play any role. Since then they substituted private capital flows for the financing of current account deficits (Homburg 2011, 48). In June 2012 the major creditor countries Germany, Finland and the Netherlands augmented a combined surplus of 924 billion Euro whereas the GIPS countries' and Italy's combined deficit amounted to 965 billion Euro (Schubert 2012, 1). The build-up of TARGET 2 deficits is very favorable as they have no maturity and only carry the key refinancing interest and thus are much cheaper than private capital. They can be interpreted as debt financing of current account deficits and capital flight (Mayer *et al.* 2011), thus express the economic imbalances and dangerously postpone the necessary adjustment processes (Sinn 2012). Summarizing, the ECB's policies stabilized the economic framework in times when immediate reactions were necessary and thus prevented a crisis escalation (Neubäumer 2011, p. 830). But they cannot go on forever and risk perpetuating economic imbalances and the build-up of new sources of financial risk.

The fiscal policy response is established in IMF programs, bilateral credits like in the case of Ireland and Cyprus, the two rescue packages for Greece and the institutionalization of the so-called rescue umbrellas EFSM, the non-permanent EFSF and finally the permanent ESM. Under the EFSM the European Union acts as borrower and on-lends the proceeds to the beneficiaries up to 60 billion Euros. The EFSF is a legal entity on its own, is scheduled until July 2013 and has a lending capacity of 440 billion Euros. The ESM as a separate legal entity with a lending capacity of up to 500 billion Euros is meant to begin to exist at the time of this writing. It will act as a permanent mechanism to restore economic stability in single member countries hit by asymmetric shocks. The resulting fiscal transfers between member countries are favorable for the functioning of a common currency area (Kenen, 1969). Its importance must even be stressed as the EZ lacks a fiscal unity like the United States where the federal tax and welfare system works as an automatic stabilizer (The Economist, 2011). Moreover, the often stated critic that the rescue umbrellas prevent an efficient capital allocation (Sinn 2011, p. 7) must be questioned in light of spill-over effects and contagion risk (Boysen-Hogreve, 2011, p. 7).

Besides these immediate rescue activities it is necessary to prevent moral hazard strategies. Thus, closer political union and strict rules associated with fiscal aid must be implemented. The requirement of structural economic reforms is necessary and a partial sovereignty loss inevitable. Moreover, it must not be forgotten that only the financial crisis unveiled the economic imbalances in the EMU and excessive lending produced the underlying credit bubbles which are the crisis's root. The banking industry must be stabilized through higher equity ratios, stricter regulation and eventually a revival of the dual banking system. In addition, traditional monetary policy with its focus on consumer price stability and the accompanying anti-cyclical monetary policy must be rethought as asset price stability plays an important role for overall financial stability. New concepts in form of macro prudential regulation as a supplement to traditional policy might be the right remedy (Hansen *et al.* 2010).

Conclusion

Consequently, it is apparent that the EZ faces great economic challenges that might have been favored but were not caused by the common currency. Even more important the common currency system stabilized the economic situation and thus must be seen positively (Neubäumer, 2011, p. 832). Although the exit of single countries out of the EZ is discussed as an alternative it has to be doubted heavily whether this option is the one to be taken in light of the immense economic cost (Deo *et al.* 2011, p. 1).

Consequently, the various (immediate) rescue activities have been necessary and correct. Their mid- and long-term success, however, is dependent on the economic reforms implemented within the EZ, the willingness to reduce the debt levels and institutional reforms necessary for preventing moral hazard behavior of single countries. In the fifth year after the outbreak of the financial crisis it is also important to modify the rescue activities, especially the monetary ones, in order to prevent further cementation of the underlying forces of the economic imbalances.

Structural economic reform might lead to higher competitiveness and economic expansion in the coming years. In addition, an accompanying overhaul of the financial system could lead to greater financial stability and a reduction of the risk of future financial crisis and thus the EMU's current crisis must be seen as a chance for positive mid- and long-term welfare effects.

References

[1] Barkbu, B., Rahman, J., and Valdés, R. *et a*l. 2012. Fostering Growth in Europe Now. *IMF Staff Discussion Note* 06/18/2012.

[2] Bergheim, S. 2007. Spain 2020 – the success story continues. *DB Research Current Issues* 09/11/2007.

[3] Bräuninger, D., and Majowski, C. 2011. Labor mobility in the Euro area. *DB Research EU Monitor 85* 09/20/2011.

[4] Boysen-Hogrefe, Jens. 2011. Für einen Schuldenschnitt und gegen den Rettungsschirm? Argumente auf dem Prüfstand. *Kiel Policy Brief* 29.

[5] Deo, S., Donovan, P., and Hatheway, L. 2011. Euro break-up: the consequences. *UBS Investment Research* 09/06/2011.

[6] Deutsche Bundesbank. 2012. *Monatsbericht Juni 2012*. Deutsche Bundesbank.

[7] The Economist. 2003. Sick man walking: Germany needs a lot more reforms than those in this week's modest deal. *The Economist*: http://www.economist.com/node/2304703.

[8] EEAG. 2012. *The EEAG Report on the European Economy*. CESifo.

[9] Europäische Zentralbank. 2012. *Monatsbericht Juni 2012*. Europäische Zentralbank.

[10] Feldstein, M. 1997. The Political Economy of the European Economic and Monetary Union: Political Sources of an Economic Liability. *Journal of Economic Perspectives,* 11(4): 23-42.

[11] Hanson, S., Kashyap, A.K., and Stein, J. 2010. A Macroprudential Approach to Financial Regulation. *Paper prepared for the Journal of Economic Perspectives*.

[12] Homburg, S. 2012. Anmerkungen zum Target2-Streit. *ifo Schnelldienst,* 16/2011: 46-50.

[13] Kenen, P.B. 1969. *The Theory of Optimum Currency Areas: An eclectic view* in: Mundell, R. A. and Alexander K. Swoboda (Eds.), *Monetary Problems of the International Economy*, University of Chicago Press: 41–60.

[14] Keynes, J.M. 1946. The Balance of Payments of the United States. *The Economic Journal,* 56(222): 172-187.

[15] Krugman, P.R. 2000. *Currency Crisis*. The University of Chicago Press.

[16] Klodt, H. 2011. Ist die Währungsunion zu retten? Für einen anreizeffizienten Krisenmechanismus. *Kiel Working Paper* 1690.

[17] Mayer, T., Möbert, J., and Weistroffer, C. 2011. Makroökonomische Ungleichgewichte in der EWU und das Eurosystem. *ifo Schnelldienst,* 16/2011: 31-38.

[18] Milbradt, G. 2011. Die EZB auf der schiefen Bahn: Target-Salden und Eurokrise. *ifo Schnelldienst,* 16/2011: 39-45.

[19] Mundell, R.A. 1961. A Theory of Optimum Currency Areas. *The American Economic Review*, 51 (4): 657-665.

[20] Neubäumer, R. 2011. Eurokrise: Keine Staastsschuldenkrise, sondern Folge der Finanzkrise. *Wirtschaftsdienst,* 91(12): 827-833.

[21] Schubert, M. 2012. Economics aktuell: Bundesbank: Target2-Saldo sinkt leicht auf 727 Mrd. Euro. *Commerzbank Economic Research* 08/07/2012.

[22] Shambaugh, J.C. 2012. The Euro's Three Crises. *BPEA Paper* 03/12/2012.

[23] Sinn, H.-W. 2011. *Die* europäische Zahlungsbilanzkrise: Eine Einführung. *ifo Schnelldienst,* 16/2011: 3-8.

[24] Sinn, H.-W. 2012. Die Target-Kredite der Deutschen Bundesbank. *ifo Schnelldienst* 65: Special Edition March 2012.

[25] Subacchi, P., and Pickford, S. 2012. Broken Forever? Addressing Europe's Multiple Crises. *Chatham House Briefing Paper* 03/27/2012.

[26] Sweeney, P. 2008. *Ireland's Economic Success: Reasons and Lessons*. Dunedin.

[27] Thompson, D. 2011. Fiscal Union Cannot Save the Euro. *The Atlantic*: http://www.theatlantic.com/business/archive/2011/11/fiscal-union-cannot-save-the-euro/249093/.

*** The Economist. 2011. Greek Americans: Which American states enjoy the biggest fiscal transfers? *The Economist*: http://www.economist.com/node/21524887.

ASERS Publishing
Volume III Issue 2(6) Winter 2012
DOI: 10.2478/v10261-012-0009-y

A UNIVERSAL SOLUTION FOR UNITS - INVARIANCE IN DATA ENVELOPMENT ANALYSIS

Jin XU
China Center for Health Development Studies
Peking University, Beijing**, China**
xujin@bjmu.edu.cn
Panagiotis D. ZERVOPOULOS
China Center for Health Development Studies
Peking University, Beijing, **China**
Department of Business Administration of Food and Agricultural Enterprises
University of Ioannina, Agrinio, **Greece**
pzervopoulos@uoi.gr
Zhenhua QIAN
School of Social Science
University of Science and Technology Beijing, **China**
rosezhenhua@yahoo.com.cn
Gang CHENG
China Center for Health Development Studies
Peking University, Beijing, **China**
chenggang@bjmu.edu.cn

Abstract:

The directional distance function model is a generalization of the radial model in data envelopment analysis (DEA). The directional distance function model is appropriate for dealing with cases where undesirable outputs exist. However, it is not a units-invariant measure of efficiency, which limits its accuracy. In this paper, we develop a data normalization method for DEA, which is a universal solution for the problem of units-invariance in DEA. The efficiency scores remain unchanged when the original data are replaced with the normalized data in the existing units-invariant DEA models, including the radial and slack-based measure models, i.e., the data normalization method is compatible with the radial and slack-based measure models. Based on normalized data, a units-invariant efficiency measure for the directional distance function model is defined.

Keywords: Data envelopment analysis, data normalization, units-invariance, directional distance function.

JEL Classification: C02, C61, C67.

1. Introduction

Data Envelopment Analysis (DEA), originally developed by Charnes *et al.* (1978), is a nonparametric method that draws on linear programming for measuring the comparative efficiency of Decision making Units (DMUs). DEA has been applied extensively in many different areas (Cook & Seiford, 2009; Seiford, 1996). A fundamental advantage of DEA is units-invariance, which means that efficiency scores assigned to DMUs are independent of the measurement units of the inputs and outputs that are utilized in the assessment process (Lovell and Pastor, 1995; Tone, 2001). Radial DEA models, such as CCR and BBC models (Banker *et al*., 1984; Charnes, 1994), and the radial measure, such as the slack-based measure (SBM), are units-invariant (Färe, and Knox Lovell, 1978; Tone, 2001).

The directional distance function model is a generalization of radial models (Chambers *et al*., 1996; Chambers *et al*., 1998; Chung *et al*., 1997). A special feature of the directional distance function model is that the direction the DMUs under evaluation are projected to the production frontier can be customized. By assigning a direction vector in Euclidean space, one can project the evaluated DMU on a specific point on the frontier. Particularly when the direction vector points towards the origin of the coordinates, the directional distance function model is equivalent to the radial model. Two advantages of the directional distance function model are that: 1) researchers can specify the direction of decreasing inputs and increasing outputs by assigning a direction vector, and 2) researchers can easily deal with the cases where undesirable outputs exist. However, a drawback of the directional distance function is that its measurement is generally not units-invariant. Taking into account that the inputs and outputs of the evaluated DMUs serve as the direction vector, changes in the measurement units of inputs or outputs potentially can lead to significant differences in the results.

The proposed data normalization method provides a universal solution when the applied DEA model violates the units-invariance criterion. The properties of the proposed method are tested with the DEA-based directional distance function model, but the new method can be applied to all existing and future DEA models.

2. The method of data normalization of data envelopment analysis and its properties

The measurement of efficiency using radial DEA models is not affected by the measurement units of inputs and outputs because efficiency results from the comparison of the inputs and outputs of the evaluated DMU against the corresponding values of the target DMU (benchmark). For radial models, the inputs or outputs are improved in proportion. In non-radial models, such as SBM models, the "proportional improvement" restriction is loosened, but the measurement of efficiency still draws on input-output ratios. As a result, efficiency scores are not affected by the measurement units of inputs and outputs.

The concept used to develop a method for dealing with the issue of units-invariance is based on the introduction of a preparation stage prior to the application of DEA models. In this stage, the original input and output data are converted into dimensionless data. When dimensionless data are utilized, this stage ensures that the efficiency scores produced by any DEA model will meet the units-invariance criterion.

The proposed procedure is expected to satisfy the conditions below:

(1) The data conversion should not affect efficiency scores measured by any units-invariant radial or non-radial DEA model.
(2) The results produced by DEA models using converted data can be converted reversely so that to be completely consistent with the results obtained from DEA models utilizing original data. The consistency of the results should be expected regardless of the units-invariance DEA model (i.e., radial or non-radial) that is applied.

The above two conditions ensure the proposed model's compatibility with existing units-invariant DEA measures.

Taking into account the points raised above, in this paper, we develop a DEA data normalization method.

Let m represents the number of inputs and q represents the number of outputs for each of the n DMUs. Column vectors x_j and y_j express the inputs and outputs, respectively, of DMU$_j$, $\hat{x}_j$ and $\hat{y}_j$ denote the normalized value of inputs and outputs, respectively; and x_0 and y_0 stand for the original inputs and outputs, respectively, of the evaluated DMU (DMU$_o$). A conversion is applied as follows

$$\begin{aligned}&\hat{x}_{ij} = x_{ij} / x_{i0}, \quad i = 1,2,...,m\\&\hat{y}_{rj} = y_{rj} / y_{r0}, \quad r = 1,2,...,q\\&j = 1, 2, \ldots, n\end{aligned} \tag{2.1}$$

The normalization formula can be extended to inputs or outputs with negative values as follows

$$\begin{aligned}&\hat{x}_{ij} = x_{ij} / |x_{i0}|, \quad i = 1,2,...,m\\&\hat{y}_{rj} = y_{rj} / |y_{r0}|, \quad r = 1,2,...,q\\&j = 1, 2, \ldots, n\end{aligned} \tag{2.2}$$

Essentially, the inputs (outputs) of DMU$_o$ serve as measurement units for every input (output) of the sample. The data conversion presented in formulas (2.1) and (2.2) does not affect the efficiency scores measured by any DEA model that is originally units-invariant.

Unlike other data normalization methods, the proposed data normalization for DEA yields one discrete normalized dataset for each DMU$_j$, i.e., there will be n normalized data sets for the n DMUs of the sample.

Data normalization for DEA has the following properties:

(1) DEA data normalization is a dimension-free conversion. Regardless of the measurement units of the original inputs and outputs or even the changes in the measurement units used with the original inputs and outputs, the normalized data remain the same.

(2) All the inputs and outputs of DMU_o are equal to unity after normalization.

Subsequent to data normalization, the DEA models that are originally units-invariant yield efficiency scores that are identical to those obtained when non-normalized data are used. In addition, when normalized data are used, the slacks generated from DEA models can be converted reversely, as follows

$$s_i^- = \hat{s}_i^- \times x_{i0}, \ s_r^+ = \hat{s}_r^+ \times y_{r0} \quad (2.3)$$

where s stands for reversely converted slacks, $\hat{s}$ are the slacks identified by the DEA model when normalized data are utilized, x_{i0} and y_{r0} express the inputs and outputs, respectively, of DMU_o.

The input-oriented CRS model using raw data can be expressed as

$$\begin{aligned} &\min \ \theta \\ &\text{s.t. } X\lambda + s^- - \theta x_0 = 0 \\ &\quad Y\lambda - s^+ = y_0 \\ &\quad \lambda, s^-, s^+ \geq 0 \end{aligned} \quad (2.4)$$

The output-oriented CRS model using raw data can be expressed as:

$$\begin{aligned} &\max \ \varphi \\ &\text{s.t. } X\lambda + s^- = x_0 \\ &\quad Y\lambda - s^+ - \varphi y_0 = 0 \\ &\quad \lambda, s^-, s^+ \geq 0 \end{aligned} \quad (2.5)$$

In radial DEA models, radial movement and slack movement are negative for inputs, and positive for outputs. The relationship between the original inputs (outputs), radial movements, slack movements, and target inputs (outputs) are formulated as follows

Target value = original value + radial movement + slack movement

$$X\lambda = x_0 + (\theta - 1)x_0 + (-s^-) \quad (2.6)$$

$$Y\lambda = y_0 + (\varphi - 1)y_0 + s^+ \quad (2.7)$$

where (θ-1) expresses the radial movement of the input in model (2.6), and (φ-1) denotes the radial movement of the output in model (2.7).

After normalization of the data, the input-oriented CRS model becomes

$$\begin{aligned} &\min \ \theta \\ &\text{s.t. } \hat{X}\lambda + \hat{s}^- - \theta \hat{x}_0 = 0 \\ &\quad \hat{Y}\lambda - \hat{s}^+ = \hat{y}_0 \\ &\quad \lambda, \hat{s}^-, \hat{s}^+ \geq 0 \end{aligned} \quad (2.8)$$

Respectively, after normalization of the data, the output-oriented CRS model is written as

$$\begin{aligned} &\max \ \varphi \\ &\text{s.t. } \hat{X}\lambda + \hat{s}^- = \hat{x}_0 \\ &\quad \hat{Y}\lambda - \hat{s}^+ - \varphi \hat{y}_0 = 0 \\ &\quad \lambda, \hat{s}^-, \hat{s}^+ \geq 0 \end{aligned} \quad (2.9)$$

According to property (2) of the data normalization method for DEA, when normalized data are utilized in radial DEA model, all of the inputs and outputs of DMU_o are equal to unity. As a result, formulas (2.6) and (2.7) can be rewritten as

$$\hat{X}\lambda = 1 + (\theta - 1) + (-\hat{s}^-) \quad (2.10)$$

$$\hat{Y}\lambda = 1 + (\varphi - 1) + \hat{s}^+ \quad (2.11)$$

The non-oriented CRS-SBM model can be expressed as

$$\min \quad \theta = \frac{1 - \frac{1}{m}\sum_{i=1}^{m}\frac{s_i^-}{x_{i0}}}{1 + \frac{1}{q}\sum_{r=1}^{q}\frac{s_r^+}{y_{r0}}}$$

$$\text{s.t.} \quad X\lambda + s^- = x_0$$

$$Y\lambda - s^+ = y_0$$

$$\lambda, s^-, s^+ \geq 0 \quad (2.12)$$

After normalization of the data, model (2.12) becomes

$$\min \quad \theta = \frac{1 - \frac{1}{m}\sum_{i=1}^{m}\hat{s}_i^-}{1 + \frac{1}{q}\sum_{r=1}^{q}\hat{s}_r^+}$$

$$\text{s.t.} \quad \hat{X}\lambda + \hat{s}^- = \hat{x}_0$$

$$\hat{Y}\lambda - \hat{s}^+ = \hat{y}_0$$

$$\lambda, \hat{s}^-, \hat{s}^+ \geq 0 \quad (2.13)$$

In model (2.13), the inefficiency is expressed as the average of the slacks identified when normalized data are applied.

In order to prove empirically the consistency of the efficiency scores when normalized data are incorporated in units-invariant DEA models, we refer to Table 1. The testing sample consists of seven DMUs with two inputs (x_1 and x_2) and one output (y). Let DMU G be the unit under evaluation (DMU_o) and apply the input-oriented CRS model to original (raw) and normalized data. The normalized data illustrated in Table 1 are calculated using formula(2.1). The efficiency score obtained from raw data is identical with the score that resulted from the utilization of normalized data. In a radial model, radial movement (-0.31) is constituted as the degree of inefficiency.

Table 1. Illustration of DEA data normalization: efficiency measurement of DMU G using the input-oriented CRS model

DMU	Raw data				Normalized data		
	x_1	x_2	y		$\hat{x}_1$	$\hat{x}_2$	$\hat{y}$
A	10.00	40.00	10.00		0.20	0.67	0.50
B	15.00	25.00	10.00		0.30	0.42	0.50
C	32.00	24.00	16.00		0.64	0.40	0.80
D	48.00	16.00	16.00		0.96	0.27	0.80
E	24.00	48.00	16.00		0.48	0.80	0.80
F	54.00	27.00	18.00		1.08	0.45	0.90
G	50.00	60.00	20.00		1.00	1.00	1.00
Efficiency score	0.69				0.69		
Radial movement	-15.62	-18.75	-0.00		-0.31	-0.31	-0.00
Slack movement	-0.00	-0.00	-0.00		-0.00	-0.00	0.00
Projection	34.38	41.25	20.00		0.69	0.69	1.00

3. Efficiency measurement using the directional distance function model

The linear programming of the directional distance function model is defined as follows

$$\begin{aligned} &\max \quad \beta \\ &\text{s.t.} \quad X\lambda + \beta v \le x_0 \\ &\qquad Y\lambda - \beta u \ge y_0 \\ &\qquad \lambda, v, u \ge 0 \end{aligned} \tag{3.1}$$

where v and *u* denote the input and output direction vectors, respectively.

In directional distance function models, direction vectors determine the directions of movement of the inputs and outputs of the inefficient DMUs and target values (projections on the frontier), thereby determining efficiency scores. Direction vectors also reflect the relative importance of inputs and outputs in efficiency measurement. Figure 1 illustrates the impact of direction vectors on efficiency measurement drawing on an input-oriented CRS directional distance function model using normalized data.

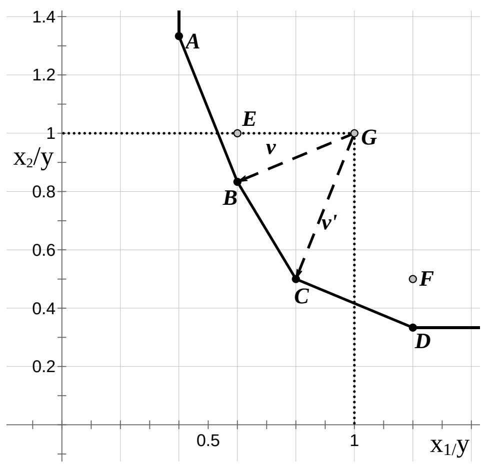


Figure 1. An input-oriented CRS directional distance function model using normalized data

In Figure 1, the horizontal coordinate represents the consumption of x_1 for each unit of output and the vertical coordinate represents the consumption of x_2 for each unit of output. When the direction vector is parallel to the horizontal axis, i.e., $v = (1, 0)$, improvement is applied solely to x_1, and the efficiency score is determined exclusively by the inefficiency of x_1. Similarly, when direction vector is parallel to the vertical axis, i.e., $v = (0, 1)$, improvement is associated only with x_2, and the efficiency score is determined exclusively by the inefficiency of x_2. Furthermore, a downward movement of the direction vector, i.e., from v to v', indicates a decrease of the impact of x_1 on the measurement of the efficiency score and an increase of the impact of x_2.

When the directional distance function models are incorporated in DEA, the inputs and outputs of DMU_o usually are utilized as direction vectors. In such situations, directional distance function models are equivalent to radial DEA models, and β, which reflects the degree of inefficiency, has the property of units-invariance. Unless the direction vectors are equal to the inputs and outputs of the DMU under evaluation, β is no longer units-invariant. Previous studies have not developed a solution for the problem of units-variance. As a result, the applicability of directional distance function models in efficiency measurement is limited.

Drawing on the definition of SBM model, we defined a units-invariant directional distance function model on the basis of DEA data normalization as follows

$$\theta = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} \beta v_i}{1 + \frac{1}{q}\sum_{r=1}^{q} \beta u_r} = \frac{1 - \beta \frac{1}{m}\sum_{i=1}^{m} v_i}{1 + \beta \frac{1}{q}\sum_{r=1}^{q} u_r}$$

$$\max \quad \beta$$

$$\text{s.t.}\quad \hat{X}\lambda + \beta v \le \hat{x}_0$$
$$\hat{Y}\lambda - \beta u \ge \hat{y}_0$$
$$\lambda, v, u \ge 0 \tag{3.2}$$

where *βv* and *βu* represent the inefficiency of the inputs and outputs, respectively. The inefficiency score of the evaluated DMU is calculated as the arithmetical mean of the inefficiency scores of inputs and outputs.

In model (3.2), when the input direction vector *v* is set equal to the input vector of DMU_o, i.e., *v* = (1, 1…, 1), and the output direction vector *u* is assigned a null vector, the model becomes equivalent to the input-oriented radial DEA model using normalized data, with efficiency score $\theta = 1-\beta$. The efficiency score obtained from the application of model (3.2) is identical with the results obtained from radial models (2.4) and(2.8).

Alternatively, in model (3.2), by assigning a null vector to the input direction vector *v*, and setting the output direction vector *u* equal to the output vector of DMU_o, i.e., *u* = (1, 1…, 1), the directional distance function model becomes equivalent to the output-oriented DEA model using normalized data. In this case, the efficiency score is defined as $\theta = 1/(1+\beta)$. The efficiency score calculated by the directional distance function model (3.2) is identical with the results provided by radial models (2.5) and(2.9).

Theorem 1: For the data set illustrated in Table 1, if the length of the direction vector changes and the direction of the same vector is unchanged, then the efficiency remains unchanged.

Proof: Assume that the direction vectors of input and output are scaled up proportionally from *v* and *u* to *bv* and *bu*, respectively, with *b* being a positive real number, and the Euclidian directions of the vectors are unchanged. Thus, model (3.2) becomes

$$\theta = \frac{1-\beta b \frac{1}{m}\sum_{i=1}^{m} v_i}{1+\beta b \frac{1}{q}\sum_{r=1}^{q} u_r}$$

$$\max \quad \beta b$$
$$\text{s.t.}\quad \hat{X}\lambda + \beta b v \le \hat{x}_0$$
$$\hat{Y}\lambda - \beta b u \ge \hat{y}_0$$
$$\lambda, v, u \ge 0 \tag{3.3}$$

If we let *α* be equal to *βb*

$$\theta = \frac{1-\alpha \frac{1}{m}\sum_{i=1}^{m} v_i}{1+\alpha \frac{1}{q}\sum_{r=1}^{q} u_r}$$

$$\max \quad \alpha$$
$$\text{s.t.}\quad \hat{X}\lambda + \alpha v \le \hat{x}_0$$
$$\hat{Y}\lambda - \alpha u \ge \hat{y}_0$$
$$\lambda, v, u \ge 0 \tag{3.4}$$

Model (3.4) is equivalent to model(3.2), so the efficiency scores they produce will be identical.

Theorem 2: For the same data set, model (3.2) is equivalent to model (3.5) shown below

$$\min \quad \theta = \frac{1-\beta \frac{1}{m}\sum_{i=1}^{m} v_i}{1+\beta \frac{1}{q}\sum_{r=1}^{q} u_r}$$
$$\text{s.t.} \quad \hat{X}\lambda + \beta v \le \hat{x}_0$$
$$\hat{Y}\lambda - \beta u \ge \hat{y}_0$$
$$\lambda, v, u \ge 0 \qquad (3.5)$$

Proof: Using normalized data, the inputs and outputs of the evaluated DMUs are all equal to unity. We know from the constraint condition of model (3.5) that

$$0 \le \beta \le 1/\max(v_1,...,v_i), \ \mathrm{i} = 1,2,...,m$$

Considering the interval of *β*, the numerator in model (3.5) is a monotonic decreasing function, while the denominator is a monotonic increasing function. As a result, within the interval of *β*, *θ* is regarded as a monotonic decreasing function. Therefore, model (3.5) is equivalent to model(3.2).

Acknowledging that model (3.5) is a nonlinear programming model, model(3.2) should be used instead for the measurement of efficiency when the directional distance function is incorporated. On the basis of model (3.2) we can introduce weights to inputs and outputs according to their relative significance in the efficiency measurement. To be more precise, model (3.6) is presented

$$\theta = \frac{1-\beta \frac{1}{m}\sum_{i=1}^{m} w_i v_i}{1+\beta \frac{1}{q}\sum_{r=1}^{q} h_r u_r}$$
$$\max \ \beta$$
$$\text{s.t.} \quad \hat{X}\lambda + \beta v \le \hat{x}_0$$
$$\hat{Y}\lambda - \beta u \ge \hat{y}_0$$
$$\lambda, v, u \ge 0$$
$$\sum_{i=1}^{m} w_i = m, \quad \sum_{r=1}^{q} w_r = q \qquad (3.6)$$

where *w* stands for the weight assigned to inputs, and *h* indicates the weight of outputs.

Efficiency measurement can be extended to cases with undesirable outputs. Namely, when undesirable outputs are present, the directional distance function model is defined as follows

$$\theta = \frac{1-\frac{1}{m}\beta\sum_{i=1}^{m} w_i v_i}{1+\omega \frac{1}{q}\beta\sum_{r=1}^{q} h_r u_r + \omega' \frac{1}{q'}\beta\sum_{r'=1}^{q'} h'_{r'} u'_{r'}}$$
$$\max \ \beta$$
$$\text{s.t.} \quad X\lambda + \beta v \le x_k$$
$$Y\lambda - \beta u \ge y_k$$
$$Y'\lambda + \beta u' \le y'_k$$
$$\sum_{i=1}^{m} w_i = m, \quad \sum_{r=1}^{q} w_r = q, \quad \sum_{t=1}^{q'} w_t = q'$$
$$\omega + \omega' = 1 \qquad (3.7)$$

where q' denotes the number of undesirable outputs incorporated in the model, h' stands for the weight of undesirable outputs, u' expresses the direction vector of undesirable outputs, and ω and ω' are the weights that determine the mix of desirable and undesirable outputs, respectively, in the measurement of efficiency.

Concluding remarks

Although units-invariance is commonly recognized as one of the most fundamental properties of DEA, some DEA models violate this property. The data normalization method we developed in this paper provides a universal solution for this problem. As inputs and outputs are rendered dimensionless, efficiency scores are independent of the measurement units of the inputs and outputs. The proposed data normalization method extends the applicability of the directional distance function model because it eliminates its units-variant problem. However, the virtues of the proposed approach are not limited to the directional distance function, since it can support any future development of DEA that may not respect the units-invariant property.

References

[1] Banker, R.D., Charnes, A. and Cooper, W.W. 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* 30: 1078-1092.

[2] Chambers, R.G., Chung, Y., and Färe, R. 1996. Benefit and distance functions, *Journal of Economic Theory* 70: 407-419.

[3] Chambers, R.G., Chung, Y. and Färe, R. 1998. Profit, directional distance functions, and nerlovian efficiency. *Journal of Optimization Theory and Applications* 98: 351-364.

[4] Charnes, A. 1994. *Data envelopment analysis: theory, methodology and application*, Kluwer Academic Publishers.

[5] Charnes, A., Cooper, W.W., and Rhodes, E. 1978. Measuring the efficiency of decision making units, *European Journal of Operational Research* 2: 429-444.

[6] Chung, Y., Färe, R., and Grosskopf, S. 1997. Productivity and undesirable outputs: a directional distance function approach, *Journal of Environmental Management* 51: 229-240.

[7] Cook, W.D., and Seiford, L.M. 2009. Data Envelopment Analysis (DEA) - Thirty Years On, *European Journal of Operational Research* 192: 1-17.

[8] Färe, R., and Knox Lovell, C. A. 1978. Measuring the Technical Efficiency of Production, *Journal of Economic Theory* 19: 150-162.

[9] Knox Lovell, C.A., and Pastor, J. T. 1995. "Units Invariant and Translation Invariant DEA Models", *Operations Research Letters* 18: 147-151.

[10] Seiford, L.M. 1996. Data Envelopment Analysis: The Evolution of the State of the Art (1978–1995), *The Journal of Productivity Analysis* 6: 99-137.

[11] Tone, K. 2001. A Slacks - Based Measure of Efficiency in Data Envelopment Analysis, European *Journal of Operational Research* 130: 498-509.

Journal of Advanced Research in Law and Economics – Biannually

Editor in Chief: PhD Mădălina **CONSTANTINESCU**
Co-Editors: PhD Russell **PITTMAN**
PhD Eric **LANGLAIS**

*J*ournal *of* **A***dvanced* **R***esearch in* **L***aw and* **E***conomics* provides readers with high quality and empirical research in law and economics. The *Journal* publishes analytical studies on the impact of legal interventions into economic processes by legislators, courts and regulatory agencies. Finally, important developments and topics in law and economics analysis will be documented and examined in special issues dedicated to that subject. The journal is edited for readability; lawyers and economists, scholars and specialized practitioners count among its readers.

Journal of Advanced Research in Law and Economics is indexed in RePEC, IndexCopernicus, CEEOL and EBSCO databases.

Web: http://www.asers.eu/journals/jarle.html Email: jarle@asers.eu

Journal of Advanced Research in Management Biannually

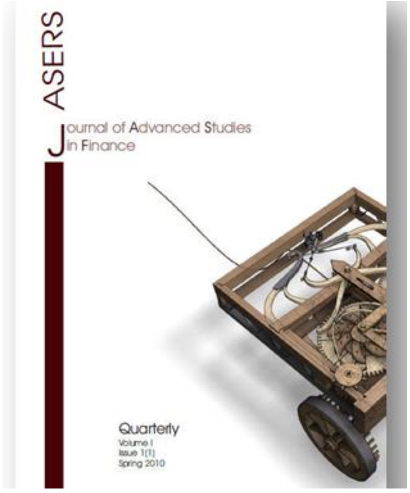
Editor in Chief: PhD Andy **ŞTEFĂNESCU**
Co-Editor: PhD Rajesh K. **PILLANIA**

The Journal aims to serve researchers, scholars through prompt publications of significant advances in any branch of management science, and to provide a forum for the reporting and discussion of news and issues concerning management science.

Journal of Advanced Research in Management is indexed in RePEC, IndexCopernicus, CEEOL and EBSCO databases.

Web: http://www.asers.eu/journals/jarm.html
Email: jarm@asers.eu

Journal of Advanced Studies in Finance – Biannually



Editor in Chief: PhD. Laura **ŞTEFĂNESCU**
Co-Editor: PhD Rajmund **MIRDALA**

The Journal aims to publish empirical or theoretical articles which make significant contributions in all areas of finance, such as: asset pricing, corporate finance, banking and market microstructure, but also newly developing fields such as law and finance, behavioral finance, and experimental finance. The Journal will serves as a focal point of communication and debates for its contributors for better dissemination of information and knowledge on a global scale.

Journal of Advanced Studies in Finance is indexed in RePEC, IndexCopernicus, CEEOL and EBSCO databases.

Web: http://www.asers.eu/journals/jasf.html
Email: jasf@asers.eu

Journal of Environmental Management and Tourism – Biannually

Editor in Chief: PhD Cristina **BARBU**

Journal of Environmental Management and Tourism will publish original research and seeks to cover a wide range of topics regarding environmental management and engineering, environmental management and health, environmental chemistry, environmental protection technologies (water, air, soil), pollution reduction at source and waste minimization, energy and environment, modelling, simulation and optimization for environmental protection; environmental biotechnology, environmental education and sustainable development, environmental strategies and policies, etc.

Journal of Environmental Management and Tourism is indexed in RePEC, IndexCopernicus, CEEOL and EBSCO databases.

Web: http://www.asers.eu/journals/jemt.html Email: jemt@asers.eu

Journal of Research in Educational Sciences – Biannually

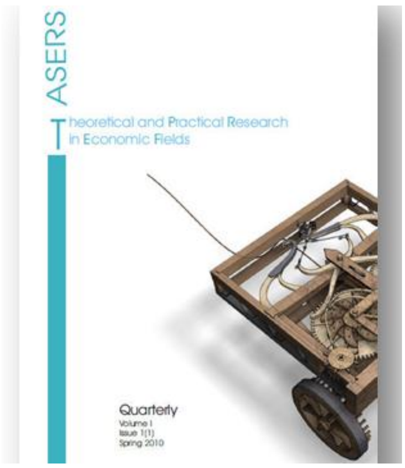
Editor in Chief: PhD Laura **UNGUREANU**

The Journal is design to promote scholars thought in the field of education with the clary mission to provide an interdisciplinary forum for discussion and debate about education's most vital issues. We intend to publish papers that contribute to the expanding boundaries' of knowledge in education and are focusing on research, theory, current issues and applied practice in this area.

Journal of Research in Educational Sciences is indexed in RePEC, IndexCopernicus, CEEOL and EBSCO databases.

Web: http://www.asers.eu/journals/jres.html Email: jres@asers.eu

Theoretical and Practical Research in Economic Fields – Biannually



Editor in Chief: PhD Laura **UNGUREANU**
Co-Editor: PhD Ivan **Kitov**

Theoretical and Practical Research in Economic Fields publishes original articles in all branches of economics - theoretical and empirical, abstract and applied, providing wide-ranging coverage across the subject area. Journal promotes research that aim at the unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking.

The Journal is indexed in RePEC, IndexCopernicus, CEEOL and EBSCO databases.
Web: http://www.asers.eu/journals/tpref.html Email: tpref@asers.eu

ASERS Publishing
ASERS

ASERS Publishing
Web: www.asers.eu
URL: http://www.asers.eu/asers-publishing
ISSN 2068 – 7710