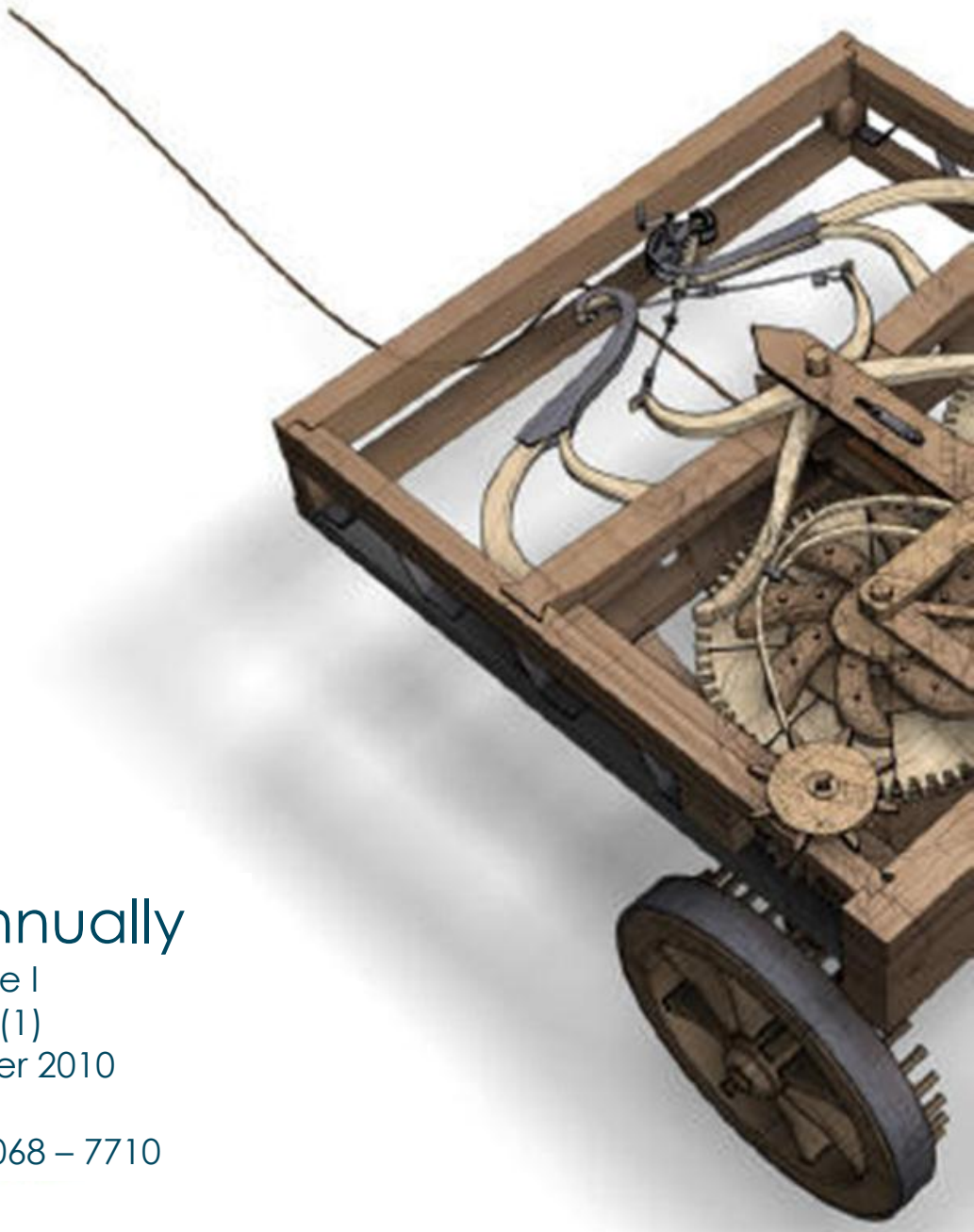
TASERS

Theoretical and Practical Research in Economic Fields



Biannually
Volume I
Issue 1(1)
Summer 2010

ISSN 2068 – 7710

Summer 2010
Volume I, Issue 1(1)

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

Contents:

Call for Papers
Winter_Issue 2010

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, will be indexed in CEEOL and very soon in IndexCopernicus and EBSCO 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 September 25st, 2010, and shall go through the usual, albeit somewhat expedited, refereeing process.

Deadline for submission of proposals:	25st September 2010
Expected Publication Date:	15th December 2010
Web:	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.

THE NEXUS BETWEEN REGIONAL GROWTH AND TECHNOLOGY ADOPTION: A CASE FOR CLUB-CONVERGENCE?

Stilianos ALEXIADIS
University of Piraeus*, **Greece**
Department of Economics
salexiadis7@aim.com

Abstract

Although the importance of technology adoption has been acknowledged, nevertheless, at a more general level, a critical question arises: what is the implication of a 'low' or a 'high' adoptive ability for regional convergence? A model is developed in which the pattern of convergence is attributed to the rate of technological adoption across regions. According this model convergence towards leading regions is feasible only for regions with a sufficient ability to adopt technology. A scheme of measurement is developed to calibrate this argument and data for the EU27 NUTS-2 regions for period 1995-2006 are used to develop an empirical analysis of the processes and conditions that have been hypothesised as generating differential regional economic change. The results suggest that adoption of technology has a significant effect on regional growth patterns in Europe, and hence the analysis has important implications for the direction of regional policy in Europe.

Keywords: convergence-club, technology adoption, European regions

JEL Classifications: O18, R11

1. Introduction

The debate on regional convergence has bred, and continues to do so, dozens of empirical studies (e.g. Button, and Pentecost 1995; Neven, and Gouyette 1995; Martin 2001). In this fast growing literature technological progress has been acknowledged to be of critical importance in promoting regional convergence. Nevertheless, the impact of the adoption of technology has received less attention. Indeed, Bernard and Jones (1996) claim that empirical studies on convergence have over-emphasised the role of capital accumulation in generating convergence at the expense of the diffusion of technology. It is the intention of this paper to develop and apply a model that incorporates technology adoption in an extensive regional context, namely that of the NUTS-2 regions of the EU. Divided into five sections, the theoretical framework upon which the empirical analysis will be conducted is articulated in Section 2. Data related issues are overviewed in Section 3, and the models are submitted to the usual econometric test yielding the main findings in Section 4. In the concluding section we offer a possible explanation for the results we obtain and suggest that might afford an interesting policy conclusion.

2. Regional Convergence and Technology Adoption

In the neoclassical model, a factor that promotes, and accelerates, regional convergence is the process of technology diffusion; a process that occurs instantly across regions. However, several criticisms have been raised against the conclusions, which this model yields, because of various simplifying assumptions underlying the results. To be more concrete, when the neoclassical model is applied to a system of regional economies, (exogenous) technology is assumed to be a public good characterised by two features, namely non-rivalry in consumption and non-excludability. Under the assumption of perfect competition it may be argued that technology has such characteristics and is, as Borts and Stein (1964) argue 'available to all' (p.8). A process of technology adoption, however, is not a simple and automatic process. Instead, it requires[1] that lagging economies should have the appropriate infrastructure or conditions to absorb technological innovations. If infrastructure conditions are not 'sufficiently developed' then it cannot be presumed that there is an 'advantage of backwardness' associated with a high technological gap[2].

Based on this argument a model, alternative to the neoclassical, is developed in this paper. The model considers a system of a large number n of *interrelated* regions, indexed by $i = 1, \ldots, n$. In such a framework, a distinction must be made between two sources of technological change. The first is a process of intentional creation of technology; an autonomous process that takes place exclusively within the 'borders' of a region. As

* The author is grateful to Mrs Judith Tomkins for her valuable comments, which substantially helped to sharpen the arguments and the focus of this paper.

[2] An argument commonly attributed to (Abramovitz 1986).

regions are, by definition, open economies technology is also affected by technological improvements that take place in other regions. This process is usually termed as *technology adoption* and constitutes the second source of technological change. Alternatively, this source refers to the part of technology that is generated from *interaction* between spatial units. An essential assumption for the purpose of this paper is that technology adoption is related to the size of the 'technological gap'[3]. This can be defined as the difference between an exogenously determined best-practice frontier (x), and the prevailing level of technology in a region (a_i), i.e. $b_i = a_i - x_i$; a measure which can be conceived as an approximation of 'technological proximity'. Thus, the growth of technology in a region ($\dot{a}_i$) can be described as follows:

$$\dot{a}_i = \tilde{\theta}_i + \xi b_i \tag{1}$$

In Equation (1) $\tilde{\theta}_i$ denotes the autonomous part of technology growth, i.e. technology created within a region. The ability of a region to implement technological innovations is represented by the parameter ξ , which reflects the opportunities for technological catch-up. Given that $b_i = a_i - x_i$, then the technological distances between a leading and a follower region, are given by: $b_l = a_l - x$ and $b_f = a_f - x$, respectively or $\dot{a}_l = \tilde{\theta}_l + \xi b_l$ and $\dot{a}_f = \tilde{\theta}_f + \xi b_f$. The growth rate for the technology gap between the two regions ($\dot{b}_{lf}$) is therefore:

$$\dot{b}_{lf} = \dot{a}_l - \dot{a}_f = \left(\tilde{\theta}_l - \tilde{\theta}_f\right) + \xi\left(b_l - b_f\right) \tag{2}$$

Defining $b_{lf} = b_f - b_l$ and $\tilde{\theta}_{lf} = \left(\tilde{\theta}_l - \tilde{\theta}_f\right)$, Equation (2) can be written as follows:

$$\dot{b}_{lf} = \tilde{\theta}_{lf} - \xi b_{lf} \tag{3}$$

Assuming that ξ is a decreasing function of the *initial* technological gap in a region, i.e. $\xi_i = f\left(b_{lf_{i,0}}\right)$ with $f' < 0$, then a relatively high initial level of technological gap implies that the prevailing conditions are not favourable for technology adoption and, consequently, the distance from the technological leader increases through time. Conversely, a relatively low level of the initial technological gap can be taken as an indication that conditions allow adoption of technological innovations, reflected in a relatively high value of ξ . Obviously, regional disparities in the absorptive parameters generate a strong tendency for regional per-capita output to diverge.

It becomes of crucial importance, therefore, to determine the dynamic path of convergence that this model implies. This can be shown using an example in which the economy is divided into three regions, one 'leader' (l) , which is at the technological frontier $(b_l = a_l - x = 0)$, and two followers, i.e. $i = 1, 2$. Assume that $\tilde{\theta}_{lf_1} - \tilde{\theta}_{lf_2} = 0$ and $b_{lf_{1,0}} - b_{lf_{2,0}} > 0$, which implies that $\xi_1 - \xi_2 < 0$. If this difference remains unchanged over a given period of time, then a catch-up, in terms of technology, between region 1 and 2 is not feasible. Stated in alternative terms, if $\left(\Delta\xi_{1,2}\right)_t \to \infty$, then $\left(\Delta b_{lf_{1,2}}\right)_t \to \infty$, as $t \to \infty$ and the two regions move towards different directions (Figure 1). Only regions with low technology gaps are likely to converge towards a steady-state equilibrium growth path, as represented by the growth rate of the leading region. Regions with relatively large technology gaps may fall progressively behind. It seems thus legitimate to ask, if there is a way for the 'technologically poor' regions to catch-up with the 'technologically rich' regions? In this example a catch-up is feasible only if region 1, viz. the 'technologically poor' region, improves its adoptive ability, i.e. if the value of ξ increases through time, from ξ_1 to ξ_1' , as shown in Figure 2. Provided that $\left(\Delta\xi_{1,2}\right)_t \to 0$, then gradually $\left(\Delta b_{lf_{1,2}}\right)_t \to 0$, allowing region 1 to catch-up with the 'technologically rich' region 2. The conclusion to drawn is that a pattern of club-convergence is the most probable outcome, if the adoptive parameters differ across regions. Movements towards overall convergence occur only as regions become similar in terms of their adoptive abilities.

[3] It is possible to extend this model by assuming that the process of technology creation is also a function of the 'technological' gap. *See* Alexiadis (2010) for an elaboration of this argument.

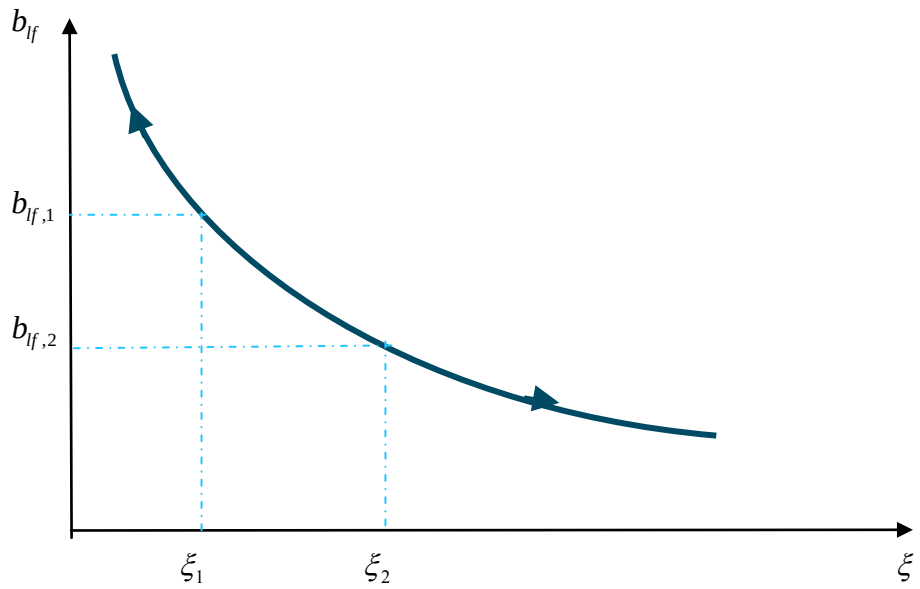


Figure 1. A pattern of regional divergence

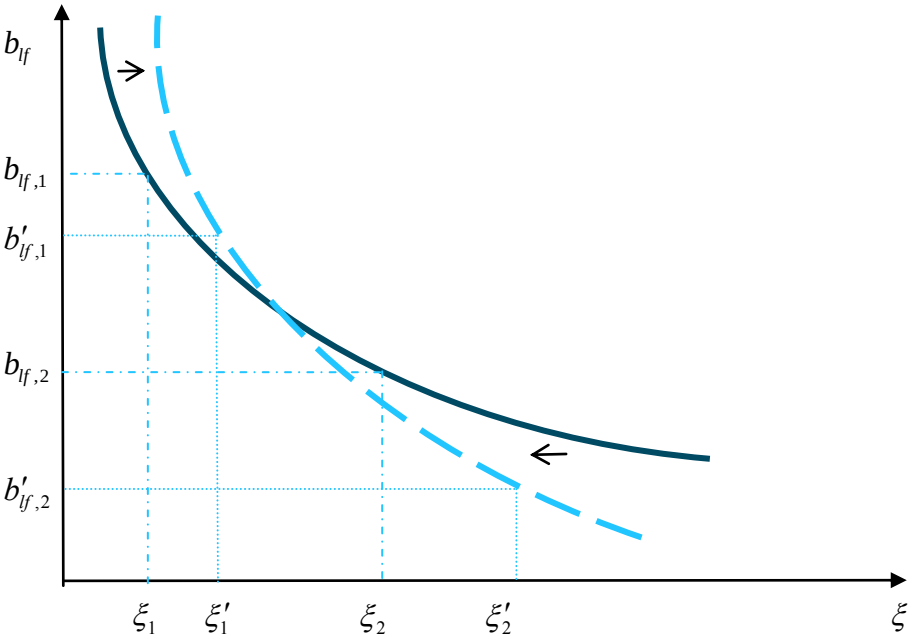


Figure 2. Club-convergence

Overall, this model suggests that convergence towards the leading region(s) is feasible only for regions with sufficient absorptive capacity. There is the distinct possibility that only regions with low technology gaps are able to converge towards a steady-state equilibrium growth path, relative to the growth rate of the leading region. Regions with large technology gaps may fall progressively behind.

3. Econometric Specification

The empirical literature on regional convergence (e.g. Barro and Sala-i-Martin, 1992) makes extensive use of two alternative tests for convergence, namely absolute and conditional convergence:

$$g_i = a + b_1 y_{i,0} + \varepsilon_i \quad (4)$$

$$g_i = a + b_1 y_{i,0} + b_{\mathbf{X}_i} \mathbf{X}_\mathbf{i} + \varepsilon_i \quad (5)$$

where y_i typically represents per-capita output, or output per worker, of the i[th] region (in logarithm form), $g_i = (y_{i,T} - y_{i,0})$ is the growth rate over the time interval $(0,T)$, and ε_i is the error-term, which follows a normal distribution.

Absolute (unconditional) convergence is signalised by $b_1 < 0$. On the other hand, conditional convergence is based upon the argument that different regional characteristics will lead to different steady-states. Conditional convergence requires that $b_1 < 0$ and $b_{\mathbf{X}_i} \neq 0$. Consider two groups of regions, let $i = k, l$, that differ not only in terms of initial labour productivity, i.e. $\mathbf{\Delta y_{kl,0}} \equiv y_{k,0} - y_{l,0} \neq 0$, but also in terms of their structural characteristics, i.e. $\mathbf{\Delta X_{kl}} \equiv \mathbf{X}_k - \mathbf{X}_l \neq 0$. Assume that $\mathbf{\Delta y_{kl,0}} > 0$ and $\mathbf{\Delta X_{kl}} > 0$. An implication of this assumption is that a superior (inferior) regional infrastructure, approximated in terms of a high (low) $\mathbf{X_i}$, is associated with a high (low) level of initial level of labour productivity. Absolute convergence amongst these groups is possible if $g_{k,T} - g_{l,T} < 0$. However, given that $\mathbf{\Delta X_{kl}} > 0$, a relatively slow process of convergence is expected. It follows, therefore, that a test for conditional convergence is more suitable for the empirical application of the model developed in Section 2, with variable(s) representing technology the principal focus, which is what the remaining paragraphs of this section will be dealing with.

Technical change, leading to regional productivity growth, originates either from within the region, namely indigenous innovation (PI_i), or from other regions, i.e. technological spillovers from adopting innovations created elsewhere (ADP_i). In the former case, technical change may be approximated by the 'propensity to innovate' ($PI_{i,t}$), as proposed by Pigliaru (2003), and can be measured in terms of the number of patents per-capita in each region.

The second source of technical growth, namely the ability of a region to adopt technological innovations, is approximated as the percentage of total employment in technologically dynamic sectors:

$$ADP_{i,t} = \frac{\sum_{\rho=1}^{k} \eta_{i,t}^{\rho}}{\sum_{j=1}^{m} L_{i,t}^{j}} \tag{6}$$

where $\eta_{i,t}^{\rho}$ refers to personnel employed in high-tech manufacturing and knowledge-intensive high-technology services ($\rho = 1, \ldots, k$), while $L_{i,t}^{j}$ is the employment in all the sectors ($j = 1, \ldots, m$).

Equation (6), represents the level of technological development, but also, indicates a capacity for technology adoption, since these are taken to apply high technology. Therefore, it is possible to express a model of 'technologically-conditioned' convergence as follows:

$$g_i = a + b_1 y_{i,0} + b_2 PI_{i,0} + b_3 ADP_{i,0} + \varepsilon_i \tag{7}$$

In Equation (7) the variables related to technology are expressed in initial values. There are two primary reasons for such an approach. The first is related to the fact that creation and adoption of innovations, normally, have future or long-run effects on regional growth (Funke and Niebuhr 2005). In other words, future growth is affected by current efforts to enhance technology. Therefore, including the two technological elements at the initial time captures these long-run effects of technology on regional growth over a specific time period. A second reason for using initial values is that it tests the hypothesis that initial conditions 'lock' regions into a high or low position, for example, how high or low levels of technology affect the pattern of regional growth and convergence.

Equation (7), thus, incorporates the potential impact of both internally generated technological change and technology adoption upon a region's growth. Broadly speaking, it is anticipated that $b_2 > 0$, since high levels of innovation are normally associated with high levels of growth and vice versa. However, it is not automatically the case that this condition promotes convergence. In other words, if low productivity regions have a high initial level of intentional technology creation, then this will have positive impacts on convergence, by enhancing their growth rates. On the other hand, if such regions have a low propensity to innovate, then no significant impacts on growth are anticipated and, hence, it may be difficult to converge with technologically advanced regions. The latter case is the more likely.

The $ADP_{i,0}$ variable reflects two distinct features, namely the *initial* level of 'technological adoption' as well as the degree to which existing conditions in a region allow further adoption of technology. A low initial level

of $ADP_{i,0}$ combined with fast growth may indicate, ceteris paribus, that less advanced regions are able to adopt technology, which is transformed into high growth rates and, subsequently, convergence with the technologically advanced regions. It may be argued, therefore, that the condition $b_3 < 0$ promotes convergence. On the other hand, a low initial value for $ADP_{i,0}$ may indicate that although there is significant potential for technology adoption, initial infrastructure conditions are not appropriate to technology adoption and, therefore, there are no significant impacts on growth. If the latter effect dominates then $b_3 > 0$, and convergence between technologically lagging and technologically advanced regions is severely constrained.

4. Econometric Estimation and Discussion

In this paper we exploit data on Gross Value Added (GVA) per-worker since this measure is a major component of differences in the economic performance of regions and a direct outcome of the various factors that determine regional 'competitiveness' (Martin 2001). The regional groupings used in this paper are those delineated by EUROSTAT and refer to 267 NUTS-2 regions. The EU uses NUTS-2 regions as 'targets' for convergence and are defined as the 'geographical level at which the persistence or disappearance of unacceptable inequalities should be measured' (Boldrin, and Canova 2001, 212). Despite considerable objections for the use of NUTS-2 regions as the appropriate level at which convergence should be measured, the NUTS-2 regions are sufficient small to capture sub-national variations (Fischer, and Stirböck 2006). The time period for the analysis extends from 1995 to 2006, which might be considered as rather short. However, Durlauf and Quah (1999) point out that 'convergence-regressions' are valid for shorter time periods, since they are based on an approximation around the 'steady-state' and are supposed to capture the dynamics toward the 'steady-state'.

The potential or otherwise for β-convergence is indicated in Figure 3, which shows a scatterplot of the average annual growth rate against the initial level of GVA per-worker. Even a cursory analysis of the EU27 data suggests that the inverse relationship between growth rate and initial level of labour productivity is not so obvious. Figure 4 indicates that this relationship is more probable to occur among regions that exceed a certain level of initial labour productivity. The presence or absence of β-convergence, however, cannot be confirmed by visual inspection alone. Therefore, the cross-section test, based on estimation of Equation (4) for the 267 NUTS2 regions, is applied to the period 1995-2006. Furthermore, the conventional test of regional absolute convergence is modified to include the hypothesis of 'technologically-conditioned' convergence. The relevant results are set out in Table 1.

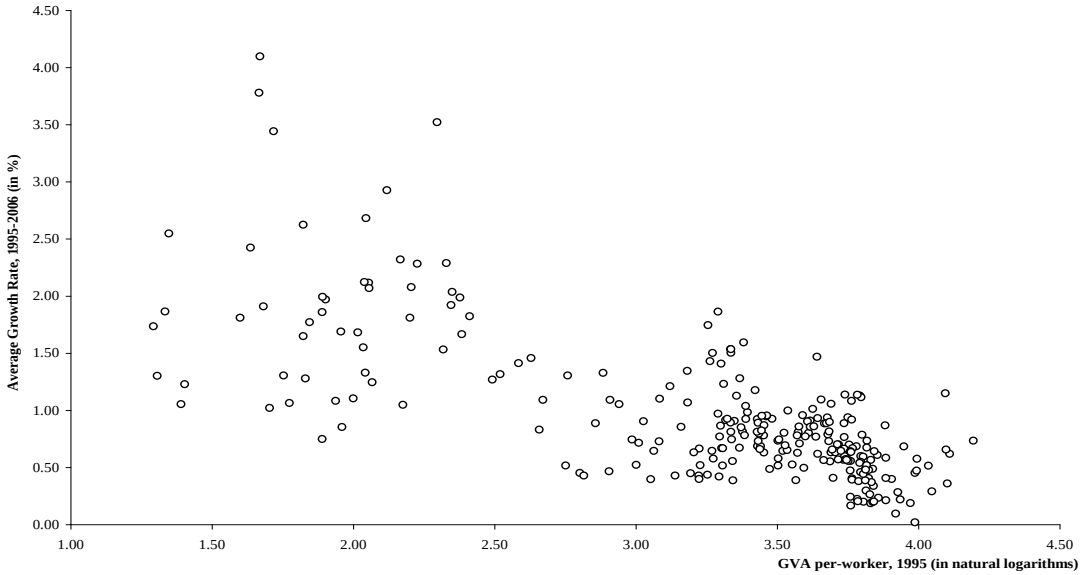


Figure 3. Absolute Convergence, GVA per-worker, EU NUTS-2 Regions, 1995-2006

Table 1. Regional Convergence, GVA per-worker, EU regions: 1995-2006

	Equation (4)	Equation (7)
Depended Variable: g_i , n = 267 NUTS-2 Regions, Ordinary Least squares		
a	0.5714**	0.4734**
b_1	-0.0747**	-0.0275*
b_2		-0.0406**
b_3		0.0650**
Implied β	0.0065**	0.0023**
LIK	147.552	164.440
AIC	-291.104	-320.879
SBC	-283.929	-306.530

Notes: ** indicates statistical significance at 95% level of confidence, * 90% level. AIC, SBC and LIK denote the *Akaike*, the *Schwartz-Bayesian* information criteria and Log-Likelihood, respectively.

As shown in Table 1, there is a statistically significant inverse relationship between growth over the time period, and the level of GVA per-worker at the start of the period. Nevertheless, the rate of convergence of labour productivity is a slow one, estimated to be 0.65% per annum. A negative and statistically significant coefficient is estimated for the variable describing technology creation. As argued in Section 3, a positive value of b_2 does not necessarily promote convergence as such, since regions with relatively high initial level of innovation exhibit relatively higher rates of growth. In this light, a negative coefficient implies that any attempts to improve technology creation, especially in lagging regions, lead to faster growth; a condition that indubitably promotes regional convergence. On the other hand, the $ADP_{i,0}$ variable is positive in sign. This suggests that, on average, regions with low values of $ADP_{i,0}$ at the start of the period, i.e. the technologically backward regions, grow slower than regions with high values, ceteris paribus. If technologically backward regions of the EU were successful in adopting technology, which subsequently is transformed into faster growth, then the estimated coefficient b_3 would be negative. Since $b_3 > 0$, and bearing in mind that a low value of $ADP_{i,0}$ implies a relatively high initial technological gap, this indicates that infrastructure conditions in regions with high technological gaps are inhibiting this process of technology adoption. A large initial technological gap may constitute an obstacle to convergence. This proposition is supported by the empirical analysis which suggests that, the rate of convergence implied by the 'technologically-conditional' model is considerably slower to that obtained from the absolute-convergence specification, 0.23% and 0.65%, respectively.

The superiority of the model described by Equation (7) is supported by both the criteria for model selection applied here, namely the *Akaike* (AIC) and the *Schwartz-Bayesian* (SBC) information criteria.[4] Further support is also provided by the value of the Log-likelihood (LIK), which increases, as anticipated, with the introduction of the technological variables. Overall, these results suggest a significant technological dimension in the process of European regional convergence. However, the relatively low rates of convergence imply that technological differences are considerable across the EU-27 regions. This brings the possibility of club-convergence into consideration.

It is possible to examine this hypothesis empirically by estimating Equation (4) and (7) for two different regional groupings. The first (leading) group includes regions with initial level of labour productivity more than 75% of the EU-27 average, while labour productivity in the members of the second (lagging) group is less than 75% of the EU-27 average. The relevant results are set out in Tables 2 and 3.

[4] As a rule of thumb, the best fitting model is the one that yields the minimum values for the AIC or the SBC criterion.

Table 2. Regional Convergence in Europe, Leading group: > 75% of the EU27 Average

	Equation (4)	Equation (7)
Depended Variable: g_i , n = 183 NUTS-2 Regions, Ordinary Least squares		
a	1.2668**	1.0515**
b_1	-0.2635**	-0.1855**
b_2		-0.0382**
b_3		0.0626**
Implied β	0.0255**	0.0171**
LIK	132.4982	141.8320
AIC	-260.9951	-275.6630
SBC	-254.5760	-262.8250

Notes: ** indicates statistical significance at 95% level of confidence, * 90% level. AIC, SBC and LIK denote the *Akaike*, the *Schwartz-Bayesian* information criteria and Log-Likelihood, respectively.

Table 3. Regional Convergence in Europe, Lagging group: < 75% of the EU27 Average

	Equation (4)	Equation (7)
Depended Variable: g_i , n = 84 NUTS-2 Regions, Ordinary Least squares		
a	0.5689**	0.5026*
b_1	-0.0802**	-0.0515
b_2		-0.0358**
b_3		0.0682**
Implied β	0.0069**	0.0044
LIK	31.6536	37.2461
AIC	-59.3069	-66.4919
SBC	-54.4453	-56.7687

Notes: ** indicates statistical significance at 95% level of confidence, * 90% level. AIC, SBC and LIK denote the *Akaike*, the *Schwartz-Bayesian* information criteria and Log-Likelihood, respectively.

The results clearly indicate that the two groups converge at different rates. The estimated rate of absolute convergence for the leading group is 2.5% while the lagging regions move towards steady-state equilibrium at a considerably smaller rate (about 0.7% per-annum). Similarly, the rate of convergence, after conditioning for technological differences, in the group of lagging regions hardly exceed 0.4% per-annum, as opposed to rate more than four times higher rate in the leading group. It is important to note, however, that the two technological variables are statistically significant and positive for both groups. A necessary condition for lagging regions to converge is a negative value of the adoptive variable. It follows, therefore, that adoption of technology, although it might be the best 'vehicle' for lagging regions to converge with leading regions, is nevertheless a process which might be difficult for lagging regions, especially during the early stages of development when conditions are least supportive. In order, therefore, for the adoption of technology to set the lagging regions of the EU in a process of convergence with the leading regions requires an improvement in infrastructure conditions. The message, therefore, from the empirical application of the model developed in this paper is clear. The adoption of technology to set the lagging regions of the EU in a process of convergence with the leading regions requires an improvement in infrastructure conditions.

5. Conclusion

It is beyond argument that, although an increasing number of empirical studies have paid attention to issues of economic convergence in the EU, the impact of technology adoption in regional convergence has so far received more limited attention. We have attempted in this paper to address this question by developing a model of regional convergence that puts primary focus upon the infrastructure conditions in a region. This model is flexible enough to be applied to several regional contexts, especially to the new member-states of the European

Union. Estimating this model using data for the 267 NUTS-2 regions of the EU-27 over the period 1995-2006 yields some interesting results. To be more specific, it is established that the NUTS-2 regions of EU-27 exhibit a very slow tendency towards convergence in terms of labour productivity. An important conclusion to emerge from the empirical application is that the EU-27 regions exhibit even slower tendencies to converge *after* conditioning for technological differences across regions. This can attributed, possibly to inappropriate infrastructure conditions prevailing in lagging regions, which prevent or constrain convergence with the more technologically advanced regions. In this context, club-convergence seems to be a more probable outcome across the regions of an enlarged Europe. The evidence reported in this paper seems to confirm this hypothesis.

In terms of implications for regional policy, this paper raises a number of pertinent issues. A primary aim of regional economic policy in the context of an enlarged Europe should be the promotion of high-technology activities, and R&D, including universities, scientific and research institutions. Moreover, in order to enhance regional growth and convergence, policy should seek to reorient these activities. High-technological and knowledge-creating activities should be directed, if possible, at regions with unfavorable infrastructure conditions, the purpose being to stimulate the production structures of those regions to shift to activities that implement high technology.

References

[1] Abramovitz, M. 1986. Catching Up, Forging Ahead and Falling Behind. *Journal of Economic History* 46: 385-406.

[2] Alexiadis, S. 2010. Interregional Differences in Adoptive Abilities: An Alternative Framework. *Forthcoming in Regional Science Inquire*

[3] Bernard, A., and Jones, C. 1996. Technology and Convergence. *Economic Journal* 106: 1037-44.

[4] Barro, R., and Sala-i-Martin, X. 1992. Convergence. *Journal of Political Economy* 100: 223-51.

[5] Boldrin, M., and Canova, F. 2001. Inequality and Convergence in Europe's Regions: Reconsidering European Regional Policies. *Economic Policy*. 16: 207-53.

[6] Borts, G., and Stein J. 1964. Economic Growth in a Free Market. *Columbia University Press*.

[7] Button, K., and Pentecost, E. 1995. Testing for Convergence of the EU Regional Economies. *Economic Inquiry* 33: 664-71.

[8] Durlauf, S., and Quah, D. 1999. The New Empirics of Economic Growth. In J. Taylor and M. Woodford (eds) *Handbook of Macroeconomics.* Amsterdam: Elsevier, 235-308.

[9] Fischer, M., and Stirböck, C. 2006. Pan-European Regional Income Growth and Club-Convergence. *Annals of Regional Science*, 40: 693-721.

[10] Funke, M,. and Niebuhr, A. 2005. *Regional* Geographic Research and Development Spillovers and Economic Growth: Evidence from West Germany. *Regional Studies* 39: 143-53.

[11] Martin, R. 2001. EMU versus the Regions? Regional Convergence and Divergence in Euroland. *Journal of Economic Geography* 1: 51-80.

[12] Neven, D., and Gouyette, C. 1995. Regional Convergence in the European Community. *Journal of Common Market Studies* 33: 47-65.

[13] Pigliaru, F. 2003. Detecting Technological Catch-up in Economic Convergence. *Metroeconomica* 54: 161-78.

CAN SHIFT TO A FUNDED PENSION SYSTEM AFFECT NATIONAL SAVING? THE CASE OF ICELAND

Mariangela BONASIA
University of Naples, Parthenope, **Italy**
mariangela.bonasia@uniparthenope.it
Oreste NAPOLITANO
University of Naples, Parthenope, **Italy**
napolitano@uniparthenope.it

Abstract

Across industrialised and developing countries public pension systems have been heavily reformed during the last two decades. The major concern relates the financial sustainability of pay-as-you-go pension schemes. The proposals to privatize social security lead to the creation of a multipillar system. This study assesses the validity of the effect of pension reforms on domestic savings in two steps: first, to test for the existence of a long-run relationship, we use an ARDL model; second, we employ the Kalman filter algorithm, in order to recover the parameter dynamics overtime. We analyse the case of Iceland because its pension system is characterized by a large mandatory private pillar. The empirical evidence supports the widely held view that growing mandatory pension funds financial assets has significantly positive impact on national saving. Moreover, we show that the pattern of the pension funds' coefficients seems to capture well the economic dynamic of the period. The coefficients of the funded pillars illustrate a shift upward soon after the launch of the reforms in 1998. Later on, these coefficients show a negative trend till the middle of 2004 and they increase sharply until the beginning of 2006. Afterwards, following the Icelandic and international financial crisis, they strongly declines.

Keywords: national saving, pension funds, multipillar system, ARDL model, Kalman filter

JEL Classification: E21, G23, H55

1. Introduction

Within the past two decades many countries around the globe have implemented intense pension reforms, which have often involved an increased use of funded pension programmes managed by the private sector.

Most of the world's countries are characterized by public pension provision financed by tax contributions on a pay-as-you-go basis, even if there could be important differences in size and degree of targeting (universal and/or means-tested). Currently, the western countries are facing lower growth rates while declining fertility rates and lower mortality rates are leading to an increase in the dependency ratio. As a consequence, the most unfunded systems are faced with financial sustainability problems. Many countries have enacted reforms which cut benefits and increase contributions and consequently most reformed pension systems are increasingly based on the rejection of the *dominance* of the public pension pillar. Hence, the aim is to move to a multipillar[5] system with a greater role given to the private sector, so that it can provide a minimum floor income for the elderly.

There are several reasons to increase funding within public or private pension systems. Among others[6], the one that is strictly related to the object of this paper is that such a policy could produce an equivalent increase in national saving, reducing the burden placed on the future workers who must support the retired elderly. Since Feldstein's seminal work (1974), this topic has been on the research agenda for over two decades. A heated debate has developed over the merits of a social security reform based on the building of a multipillar system and, in particular, over its impact on saving. This paper intends to investigate whether this explanation for advanced funding is valid. In doing so we try to estimate the effect on national saving of an increase in mandatory private pension funding in Iceland, a country that is increasingly meeting the criteria of the prototype three pillar system as suggested by the World Bank (1994). Its social security system could be a typical example of a great

[5] Each pension pillar provides the three functions in different ways. The first pillar, public pay-as-you-go, usually defined-benefit and redistributive, should provide social safety net support to everyone. The second one is a privately managed funded pillar that handles peoples' mandatory retirement savings; unlike the public pillar – which is redistributive, centrally controlled, and tax-financed - the second mandatory pillar should highlight savings. Finally, the third pillar encourages discretionary retirement savings and capital development.

[6] Another justification for proposals to shift from pay go to private pension schemes is that such a policy will remedy to labour market distortions (Arrau, and Schmidt-Hebbel 1995; Feldstein 1996; Feldstein, and Samwick 1997; Kotlikoff 1996a, and 1996b). However, the aim of this paper is not to analyse the relationship between funded pension schemes and labour market.

role given to the second mandatory private pension pillar since it exhibits the largest pension funds in relation to the size of its economy.

Hence, the approach taken in this study is that of long-run equilibrium via cointegration. Instead of relying on the more common Johansen multivariate cointegration procedure, autoregressive distributed lag (ARDL) cointegration modelling of Pesaran, Shin, and Smith (1999) will be carried out to determine if there is a long-term equilibrium relationship between national savings and pension funds in Icelandic economy. In this regard, we test the hypothesis that the contemporaneous convergence of the variables in the system exhibits minimum systematic error due to the presence of an error correction mechanism (short-term relationship).

Moreover, this study implements an additional analysis based on the Kalman filter methodology in order to analyse the paths of the pension funds with respect to national savings. The time varying-parameter technique uses the basic idea that the parameter vector in an econometric relationship may be subjected to sequential variation over time because of the problems of structural change, misspecification and aggregation.

The results of this study reveal that there exists a long-run equilibrium relationship between national savings and pension funds. Additionally, the combined impact of private pension provision and life insurance on national savings is positive. Moreover, further evidence comes from the Kalman filter methodology where the behaviour of the coefficients of pension funds' variable shows a positive trend.

The innovative feature of our study is that, analysing Iceland, we consider a typical example of multipillar system characterized by a small public pension pillar, a large mandatory private pension pillar and a considerable voluntary pension saving. In order to be efficient a pension system should take into account lifetime risks facing individuals and society as a whole. Furthermore a certain degree of income equalisation should be built into the system. Then the pension system should have at least some degree of flexibility and scope for choice for the individuals. Finally, it should be designed in such a way as to promote economic performance that is, saving, growth and financial sector development.

Our findings indicate that increases in pension funds financial assets increase national saving when pension funds are the result of a mandatory pension program. An interesting feature of the results about Icelandic pension system, which are compatible with results from other developed countries, is that the shift to a multipillar system implemented by Iceland could give useful suggestions to policy makers that are facing problems with social security reforms. However despite the strong impact of the 2007 financial crisis on the Icelandic economy, the reformed pension system seems to work as well as the co integration relationship with the national savings. Moreover, a novelty of this work is that the contribution rate to pension funds, increasing national saving in Iceland, could be an important policy instrument whose effect can be twofold: **1.** to secure a minimum replacement rate at retirement; **2.** to raise the future national income through a faster capital formation.

The paper is organized as follows: Section 2 presents the relevant theoretical and empirical literature. Section 3 reviews the evolution of the Icelandic economy and the pension system. The data, the ARDL cointegration methodology, the Kalman filter and the empirical results are explained in Section 4. Policy implications and the main conclusions are summarised in the fifth section.

2. A Brief Literature Review

The impact of the pension system on individual saving has been a concern since Feldstein's (1974) pioneering paper, which centres around the funding status of social security and, in particular, on the degree to which an unfunded pension system reduces private saving. The latter is the central question of the current debate[7]. However, the question analysed in this study is the impact on national saving of the implementation of a multipillar system, consisting of public pay-as-you-go scheme and of compulsory saving in funded pension plans, as in the Icelandic pension system.

According to Feldstein (1996) one of the benefits claimed for the private funded pension schemes is that they rise national saving. The reduction of the size of public unfunded scheme could rise saving and wealth directly by reducing government debt and indirectly increasing saving through two other paths. On one side, the elimination of payroll tax may lead to an increase in labour supply thus boosting GDP and increasing saving. On the other side, the shift to funded pension scheme may stimulate the capital market, leading to an increase in the efficiency of investment and thereby to an increase in economic growth and saving, depending on the extent of

[7] This literature is reviewed and summarized in CBO (1998) and Atkinson (1987).

capital market development prior the reform[8]. While theoretical arguments tend to be consistent with the view that higher private savings are associated with funding than with PAYG, convincing empirical support is missing.

Several studies have analysed transition from unfunded to fully funded system in an overlapping generation framework (Arrau, and Scmidt-Hebbel 1999; Cifuentes, and Valdes Prieto 1997, and Kotlikoff, Smetters, and Walliser 1998). The most significant burden due to this shift has to be supported by the transitional generations since they have to pay double contributions, their own and that of current retirement individuals under the unfounded system.

As Samwick (2000) remarks, the intergenerational allocation of the transition cost is the most important consideration to determine the effect of the regime change on saving.

How the government manages this additional financing requirement is crucial to the effect of the transition on saving, as discussed in Holzmann (1997). If the government tries to finance the implicit pension debts by issuing explicit debt, then public savings would decrease, so the overall national saving rate might be unchanged or even fall (Cesaratto 2006). For example, a simulation study by Hviding and Merette (1998) gives evidence that debt financed transitions may not have considerable effects on national saving and output; all that may happen is that the government has altered the form of the debt.

Another possibility is to finance the transition increasing taxation. As a consequence, the disposable income of current workers will decline and if their consume will fall by the same amount, the national saving will raise by an equal amount of tax increase. According to the life-cycle model, individuals will try to smooth the tax raise over their entire life time. Hence part of tax income will be financed by a lower private saving (Samwick 2000).

Auerbach and Kotlikoff (1987), Kotlikoff (1996a), Mitchell and Zeldes (1996) underline that if individuals are completely rational life-cycle savers, the introduction of a mandatory saving program would not have any net effect on national saving because rational individuals will reduce their previous saving by an equal amount. An investment based program should raise aggregate saving only if individuals are myopic or do not save for other reasons (for instance, bequest).

Two well-known cross sectional studies by Cagan (1965) and Katona (1965) have shown that employees covered by private pensions do not save less and may even save more than employees not covered by private pensions.

From an empirical point of view, different findings were obtained by Munnel (1976). She used a sample of 5,000 men with ages included between 45 and 59 in 1966. The author presented evidence showing that the saving of employees covered by private pensions is largely less than the saving of employees not covered by private pensions. From cross-country empirical evidence, Baillu and Reisen (1997) studied the effect of funded pension wealth as a determinant of private saving working with a sample of 10 countries and a panel of more than 100 observations. The authors showed that funded pension wealth increases private saving rates in developing countries with mandatory funded pension programs.

More recently, Murphy and Musalem (2004) conducted an empirical study on the effect of the accumulation of pension fund financial assets on national saving, using a panel of 43 industrial and developing countries. The authors divided the countries into two groups: the first one includes countries whose data on pension asset are mainly due to compulsory funded pension program; the second group of countries is instead characterized by data that are the result of voluntary funded pension program. They focused on pension funds financial assets and their impact on national saving rate. Their findings suggested that increases in pension funds financial assets increase national saving when pension funds are the result of a mandatory pension program.

Despite the goals underlined by Murphy and Musalem (2004), we are interested into studying the effect of the reforms implemented by Iceland on national saving. The aim of the present paper is to verify if, reforms consisting in a shift from a pay-as-you-go to a multipillar system with a greater role give to the second pillar (mandatory pension funds), foster national saving.

The largest pension funds in relation to the size of their respective economies are the ones that have had mandatory or quasi mandatory pension funds for many years, like Australia, Iceland, the Netherlands and

[8] Orszag and Stiglitz (1999) explore – in a deliberately provocative manner – the multipillars system of the "World Bank model" with a particular attention to the private mandatory defined contribution component. They debunk ten myths to stress that the general opinions most frequently used to support investment-based programs are often not confirmed in theory and practice. Among them, the first myth is about the common idea that private defined contribution plans raise national saving.

Switzerland[9]. The largest voluntary pension fund systems are those in the United States, United Kingdom and Canada (OECD 2005). In this paper we focus our analysis on Iceland that is increasingly meeting the criteria of the prototype three pillar systems.

3. Icelandic Economic and Pension System

The economic effects of reformed pension systems are important. It is always those who are economically active today who support today's pensioners, irrespective of the specific economic system. However to better analyse the impact of the reformed pension provision on the economic performance it is useful to summarize the evolution of the Icelandic economy.

3.1. Economic performance of Icelandic economy within the past twelve years

The first attempt of transforming the structure of the Iceland economy started in the sixties when a primitive process of liberalization was not accompanied by the necessary process of de-politicization of the economic life. Nor was the second, begun in the late eighties. The latter involved the deregulation of domestic interest rates and the flow of foreign capital, the indexation of financial obligations, the partnership agreement with the European Economic Community in 1994 and, from 1998 to 2003, the privatization of banks, investment funds and the reform of the pension system. The two main banks were sold simultaneously at a price considered low by the National Audit Office. Moreover, they were not sold to foreign banks, as was the case of banking privatization in Eastern Europe, but to people very close to the ruling political parties.

In 2001, banks were deregulated in Iceland. This set the stage for banks to upload debts when, at the same time, the share of the Icelandic assets hold by foreign institutions increased. According to the official statistics about the key indicator of the Icelandic economy, the middle of the 2004 is considered a sort of turning point. In fact, we have found at least four important changes in the main macroeconomic variables: **1.** the growth rate of GDP per capita changes its trend becoming negative after a decade of continuous positive movement; **2.** the wages started increasing; **3.** the monetary policy changed implying a turn towards a tight or restrictive monetary policy; **4.** the current account balance reached its negative peak in 2005.

The 2007 crisis stretched out when banks became unable to refinance their debts. It was estimated that, in 2008, the three major banks hold foreign debt in excess of €50 billion that is, almost 6 times the Iceland's GDP. The króna, was under pressure because the international markets considered it an overvalued currency. Due to the fact that Iceland has a small domestic market, Iceland's banks financed their expansion with loans on the interbank lending market and abroad, increasing the external debt of about 213 per cent of disposable income. Finally, the inflation rate rose in all the period but from 2004 the house component increased considerably.

3.2. Icelandic social security system

The Icelandic pension system is based on three pillars. The first pillar, according to the accepted terminology in this field, is a tax-financed public pension. The occupational pension funds are the second mandatory pillar. The third pillar is voluntary pension saving with tax incentives[10]. A comprehensive pension reform took place in 1997 and 1998 that affected the second and third pillar.

Iceland is not facing the problems of the most developed countries due to the aging population. First of all, Icelandic people are younger and, according to the predictions for 2030, its dependency ratio will remain lower than the other European countries. Moreover, labour participation rates of the elderly are also higher and the effective retirement age is higher than the most industrialized nations. This happens because of the particular Icelandic social security system in which individuals are entitled to receive public pension only from the age of 67 and pension funds are regulated so that no incentives are given for early retirement.

The first pension pillar provides an old age, disability and survivors pension. It is divided into a basic and supplementary pension, both are means-tested.

The second pillar is based on occupational pension funds. It is mandatory by law to pay at least 10 per cent of all wages and salaries into fully funded pension plans that provide lifelong retirement and disability benefits. Many of the funds were established through a collective labour agreement in 1969. Most of them are

[9] In Iceland, however, the participation in pension funds, which should be left to individual choice, is instead mandatory in agreements between unions and employers; that private pensions receive not only the effective regulation and supervision of financial markets but also of a public guarantee reintroducing elements of PAYG financing (Castellino, and Fornero 2007).

[10] For more details see Gudmundsson (2001, 2004) and Gudmundsson, and Baldursdóttir (2005).

managed jointly by representatives from the trade unions and employers. Only in 1998, a more comprehensive legislation covered the operation of pension funds into the financial markets.

In the 1980s and 1990s pension fund assets grew from 14% to 80% per annum in real terms of GDP. This placed Iceland fourth among EU and EFTA countries in terms of the size of second pillar pension fund assets as a percentage of GDP, after the Netherlands, Switzerland and the UK. At the beginning of 2005 there were forty-eight pension funds in Iceland. Ten of them were no longer receiving contributions.

The *third* pillar is characterized by voluntary private pension savings for which in 1998 tax incentive legislation was adopted as part of the general pension reform.

In 2003 and 2004 the *assets* of Icelandic pension funds and life insurance was respectively 138.4 and 146.2 as a percent of GDP (OECD 2005). The reason is that contribution to pension funds exceed benefits paid from them because of pensioners are few in proportion to working fund members. Furthermore, most of them have contributed to the fund only for short period of their working life ad so they are entitled only to relatively small benefits. As a consequence, pension funds fed on their own investments. In 2006 and 2007 the importance of pension funds relative to the size of the economy were 132,7 and 134 per cent of GDP, respectively. The situation changed dramatically in the 2008 when the assets of pension funds fell to 99.5 per cent of GDP. Due to the sharp decline of GDP to -7.2 per cent in 2009, the ratio of assets of pension funds on GDP does not give any useful information.

4. Data, methodology and empirical results

The principal task in this paper is to search for a long-run equilibrium relationship and short-term dynamics between national savings and a set of explanatory variables (here after, "Z") like public pension provision and private pension funds that could affect savings patterns. In doing so we follow three steps. Firstly, we consider the orthogonal problem that usually arises when we work with financial variables. To overcome this peculiar aspect we implement the SURE methodology. Secondly we move our analysis to the study of the long-run relationship between national saving and a vector of exogenous variables "Z". Finally we estimate, using a Kalman filter approach, a 'backward-looking' process for national savings with parameters varying with fundamental component of the private pension plans. The time varying methodology allows us to recover an unobservable factor that could affect national savings. For each endogenous variables of the model it is therefore possible to observe how the respective coefficients have changed over time.

4.1. Data analysis

The choice of the sample 1997-2009 using monthly observation was essentially based on the need of analysing the behaviour of savings after substantial reforms of the pension systems took place. For the estimation of the Equations used in this work, the variables considered are[11]: SAV is the national aggregate saving of Iceland based on the current account definition of savings as the residual difference between total income, consumption and all the retirement saving generated by the three pillars.

PAYG is defined as the logarithm of total expenditure on public pension benefits. The variable is from the Central Bank of Iceland. We expect a negative relationship between PAYG and national saving; PENS is defined as the logarithm of total expenditure on private pension funds and is expected with a positive sign. The Iceland PENS is mandatory by law to pay at least 10% of all wages and salaries into fully funded pension schemes that provide lifelong retirement and disability pensions; INS is the logarithm of insurance providing payment of a sum of money to a beneficiary on the death of the insured person or if the insured person reaches a certain age also known as life insurance and should have a positive impact on national savings. The monthly statistics for assets of pension funds including life insurance are compiled from monthly reports from 23 pension funds published by Central Bank of Iceland; INT is the nominal interest rate. We use money market rate (average rate on money market) as a measure of interest. The impact of interest rates on household savings is ambiguous because income and substitution effects work in opposite directions. Finally, the variable DR is the dependency ratio. It is calculated as the ratio of people over 64 to people aged 20-64 years. According to the life-cycle model, the higher the ratio of the elderly population to the working-age population, that is DR, the lower could be the aggregate saving because the old are retired and do not save while the young work and do save. Hence we expect a negative relationship between dependency ratio and national saving. The logarithms transformation of pension variables (PENS, PAYG and INS) reflects the general interest emphasized in the social security literature in relative versus absolute effect of this variables on national saving.

[11] Data source: IMF – Financial Statistics, Iceland Central bank and OECD statistics.

4.2. Methodology and empirical results

4.2.1. A preliminary analysis

One of the central conditions to achieve identification when we deal with financial variables is that the structural form shocks are orthogonal to one another. That is, we assume that the error term is orthogonal to the variables on the right side of the Equation (1) below. In reality, this condition may not be satisfied. Common shocks for such variables within a country may be news about economic fundamentals or announcements of releases of relevant macroeconomic data. Moreover, there may be common shocks among countries, such as oil price shocks.

Following the approach commonly used in the related literature we address the issue that the three series $PENS_t$, $PAYG_t$, and INS_t are nearly orthogonal or uncorrelated. The correlation between them measures the extent to which each series provide "orthogonal information". The former problem is related to the possibility that these variables are simultaneously determined. This can occur either because they cause each other or because they have some common omitted determinants. For instance, we assume that pension funds assets are nearly orthogonal to the life insurance assets. One reason for a violation of this condition would be a contemporaneous response of interest rates to pension and life insurance funds.

Generally speaking, ignoring this potential correlation might reduce the efficiency of the estimates, or even produce biased estimates if these variables are correlated with other included explanatory variables. To examine the impact of controlling for this correlation, we estimate a Seemingly Unrelated Regression Estimation (SURE) system of four Equations: one for private pension (PENS), one for the public pension (PAYG), another for the additional private pension (INS) and another for the interest rate (i_t), while allowing for their error terms to be correlated. The below set of Equations has contemporaneous cross-Equation error correlation so that, the Equations seem unrelated which states that the idiosyncratic shocks of the four markets are independent.

$$\begin{cases} PENS_t = \alpha_{1t} + \sum_{n=1}^{k_1} \beta_{1,j} PAYG_{t-n} + \sum_{m=1}^{k_2} \beta_{2,j} INS_{t-m} + \sum_{h=1}^{k_3} \beta_{3,j} i_{t-h} + \mu_{pens_t} \\ PAYG_t = \alpha_{2t} + \sum_{s=1}^{k_4} \beta_{4,j} PENS_{t-s} + \sum_{f=1}^{k_5} \beta_{5,j} INS_{t-f} + \sum_{g=1}^{k_6} \beta_{6,j} i_{t-g} + \mu_{payg_t} \\ INS_t = \alpha_{3t} + \sum_{v=1}^{k_7} \beta_{7,j} PAYG_{t-v} + \sum_{l=1}^{k_8} \beta_{8,j} PENS_{t-l} + \sum_{d=1}^{k_9} \beta_{9,j} i_{t-d} + \mu_{ins_t} \\ i_t = \alpha_{4t} + \sum_{e=1}^{k_{10}} \beta_{10,j} PAYG_{t-e} + \sum_{r=1}^{k_{11}} \beta_{11,j} PENS_{t-r} + \sum_{u=1}^{k_{12}} \beta_{12,j} INS_{t-u} + \mu_{i_t} \end{cases} \quad (1)$$

In Equation (1) we use impulses in a separate system so we can investigate the relationship among the impulses (PENS, PAYG, INS and *i*). The estimated results are presented in Table 1.

Table 1. SURE estimation for Iceland

	Coefficient	S.E.
α_1	-1.165762**	0.425066
β_1	0.805752**	0.057489
β_2	-0.569065**	0.057638
β_3	0.076296**	0.007398
α_2	2.035606**	0.296579
β_4	0.765641**	0.054627
β_5	0.798689**	0.027706
β_6	-0.036351**	0.008338
α_3	-2.238607**	0.448642
β_7	1.174182**	0.040732
β_8	-0.794957**	0.080517
β_9	0.044685**	0.010199
α_4	-3.078200	4.244247
β_{10}	-3.116786**	0.714934
β_{11}	6.216143**	0.602711

β_{12}	2.606142**	0.594818	
Determinant residual covariance		2.44E-07	
Sample 1997:01 2009:12	obs. 156		
Equation PENS = $\alpha_1 + \beta_1$ PAYG + β_2 INS + β_3 i			
R^2 0.4695		Adj. R^2	0.4589
Equation PAYG = $\alpha_2 + \beta_4$ PENS + β_5 INS + β_6 i			
R^2 0.7925		Adj. R^2	0.7884
Equation INS = $\alpha_3 + \beta_7$ PAYG + β_8 PENS + β_9 i			
R^2 0.7237		Adj. R^2	0.7183
Equation i = $\alpha_4 + \beta_{10}$ PAYG + β_{11} PENS + β_{12} INS			
R^2 0.2238		Adj. R^2	0.2084

The residuals from this system of Equations are then the new variables ($\mu_{pens_{it}} =$ RPENS, $\mu_{payg_{it}} =$ RPAYG, $\mu_{ins_{it}} =$ RINS and $\mu_{i_{it}} =$ R*i*, henceforth).

4.2.2. The Cointegration analysis

The second step of the empirical analysis is related to the concept of cointegration. It is developed from the belief that certain economic variables should not diverge from each other by too far a distance or diverge without bound. Such variables may drift apart in the short-run but if they continue to be too far apart in the long-run, then economic forces will bring them together again (Granger 1986). As such, an error correction model can be established to capture the short-run equilibrium. Hence, cointegration techniques seem to be ideally matched with the many economic theories that are contained in long-run equilibrium terms (McKenzie 1997).

In light of this, the autoregressive distributed lag (ARDL) approach to cointegration (Pesaran, Shin, and Smith 2001) will be used as it has good small sample properties in comparison with these techniques, as well as circumventing the problem of the order of integration of the individual variables.

The main advantage of the ARDL approach is that it can be applied regardless of whether the regressors are *I(0)* or *I(1)*, and this avoids pre-testing problems associated with the standard cointegration analysis such as Full Information Maximum Likelihood (FIML) which requires the classification of the variables into *I(1)* and *I(0)*.

The ARDL approach to cointegration involves estimating the following model:

$$SAV_t = \alpha_0 + \alpha_1 \sum_{i=1}^{n-1} SAV_{t-i} + \alpha_2 \sum_{i=1}^{n-1} Z_{t-i} + u_t \qquad (2)$$

where n is the sample size, SAV_t is the log of national savings and Z_t is a vector of exogenous variables described in Section 4.1.

It may be of interest to test the stability of regression coefficients in Equation (2) because a change in parameters between two periods can be seen as an indicator of structural changes. We run a basic regression of Equation (2) in order to investigate the stability of the coefficients and, consequently, to set also the right number of dummies. We use the recursive residual test to identify a likely break in Equation (2) and then test for its statistical significance using a Chow test. In Figure 1 the plot of the recursive residuals about the zero line ± two standard errors are shown at each point. Since residuals outside the standard error bands suggest instability in the parameters of the Equation, we need to construct several binary variables.

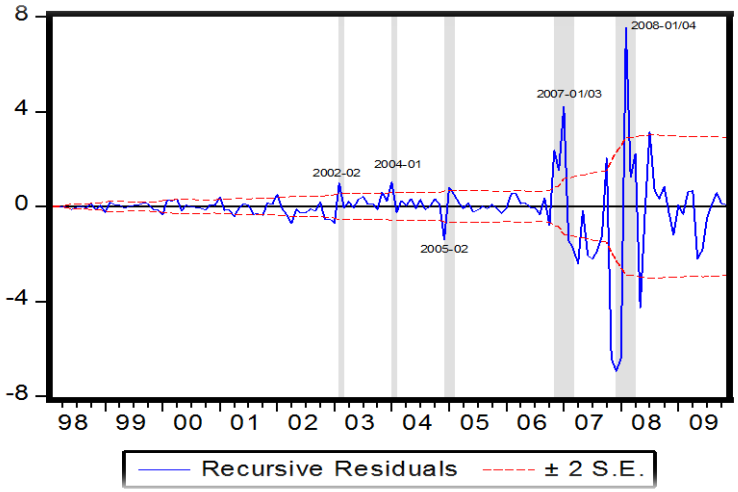


Figure 1. Recursive residual test for Iceland

Moreover, we also checked the Chow test for structural breakpoints in the sample of Equation (2). Table 2 shows the results of this test.

Finally, based on the results obtained from the above two tests, the following dummies were generated[12]: 2004:01-02 (dummy1), 2004:12-02 (dummy2), 2006:11-03 (dummy3) and 2007:12-04 (dummy4).

Table 2. Stability test results

Chow tests results* of Equation (2): $SAV_t = \alpha_0 + \alpha_1 RPENS_{t-1} + \alpha_2 RPAYG_{t-1} + \alpha_3 RINS_{t-1} + \alpha_4 GRW_t + \alpha_5 SAV_{t-1} + \varepsilon_t$		
2002:02-03	1.49312	(no break)
	(0.174)	
2004:01-02	3.6028	
	(0.0013)	
2004:12-02	9.84656	
	(0.000)	
2006:11-03	19.2155	
	(0.000)	
2007:12-04	200.252	
	(0.000)	

* F-statistics with levels of significance in parentheses. Null Hypothesis: No breaks at specified breakpoints

4.2.3. Testing for cointegration

For investigating the presence of a long-run relationship between the variables we will use two separate statistics. The first involves an F-test on the joint null hypothesis that the coefficients on the level variables are jointly equal to zero (*see* Pesaran, Shin, and Smith 1999, 2001). The second is a t-test on the lagged level dependent variable. The statistics have a non-standard distribution and depend on whether the variables are individually *I(0)* or *I*(1).

To implement the "*bound testing procedure*" it is essential to model eq. (2) as a conditional autoregressive distributed lad model (ARDL). Basically, the lagged level terms are added to an error correction form of the underlying ARDL model and the F-statistic is computed[13]. Instead of the conventional critical values, this test involves two asymptotic critical value bounds, depending on whether the variables are *I*(0) or *I*(1) or a mixture of both.

[12] These dummies are defined as one in the specified period and zero elsewhere.

[13] Pesaran, Shin and Smith (2001) tabulate two sets of asymptotic critical values to provide *critical value bounds* for all classifications of the regressors into pure *I(1)*, purely *I(0)* or mutually cointegrated.

Pesaran, Shin and Smith (2001) provide critical values for this bounds test from an extensive set of stochastic simulations under differing assumptions regarding the appropriate inclusion of deterministic variables in the error-correction model (ECM)[14]. The investigation involves testing the existence of a long-run relation among national savings, private pension funds, and public pension provision and life insurance funds.

The model in Equation (2) is more appropriately regarded as representing the equilibrium relationship in the long-run, but is unlikely to hold exactly in every single period. Hence a dynamic specification is required in order to allow the model to capture the short-run adjustment process without losing important information about the long-run equilibrium behaviour of the variables. This study uses an ECM with unrestricted intercept and no trend[15]. The empirical relationship of the error-correction model has the following general dynamic representation:

$$\Delta SAV_t = \alpha_0 + \alpha_1 T + \alpha_2 ecm_{t-1} + \sum_{i=1}^{p} \beta_{1,i} \Delta Z_{t-i} + \sum_{i=1}^{p} \beta_{2,i} \Delta SAV_{t-i} + \sum_{i=1}^{p} \gamma_{1,i} Z_{t-i} + \sum_{i=1}^{p} \gamma_{2,i} SAV_{t-i} + \sum_{j=1}^{7} \gamma_{3,j} D_j + \varepsilon_t \quad (3)$$

where α_0 is the drift component, T is the deterministic trend, ΔSAV_t is the variation of national savings, Z_t is a vector of pension variables (private pension funds, pay-as-you-go, life insurance funds), ecm_{t-1} is the equilibrium correction term (μ_t) given in Equation (2), D_j are dummies (dummy1 to dummy4) as described in Section 4.2.2, with γ's as the long-run multipliers, β's as short-run dynamic coefficients, (p) as the order of the underlying ARDL-model (p), and ε_t are white noise errors. This dynamic model shows that the movement of the variables in any period is related to the previous period's gap from long-run equilibrium. In other words, from Equation (3) whenever ΔSAV_t turns out to differ from Z_{t-1}, some sort of adjustment must occur to restore the equilibrium in the subsequent period.

Following Pesaran, Shin and Smith (2001) we try to determine the proper lag length p in Equation (3) with and without a deterministic linear trend. Table 3, presents these results using the Akaike's, Schwarz's Bayesian Information Criteria (AIC and SBC) and the Lagrange Multiplier (LM) statistics for testing the hypothesis of residual correlation of order 1 and 4. The results of the AIC and SBC suggest the use of four lags or more, while $\chi^2_{SC}(1)$ and $\chi^2_{SC}(4)$ recommend the use of a lag order of four. Since it is of a particular importance for the validity of the bound test the assumption of serially uncorrelated errors, for the sake of parsimony, we use the lag selection criterion with p=4.

Table 3. Lag-length selection criteria

	With deterministic trend				Without deterministic trend			
Lags	AIC	SBC	$\chi^2_{SC}(1)$	$\chi^2_{SC}(4)$	AIC	SBC	$\chi^2_{SC}(1)$	$\chi^2_{SC}(4)$
p=1	130.0160	117.8423	59.70	70.69	130.0170	119.3650	62.61	75.27
p=2	232.5858	221.9565	5.288	16.44	232.8846	223.7738	6.433	34.64
p=3	237.5847	222.4325	0.004	14.20	237.0085	223.3715	0.360	33.22
p=4	243.3617	222.1945	4.195♦	9.582♦	235.0044	221.3969	3.413♦	11.48♦
p=5	255.8904	225.7176	0.319	14.77	249.6885	224.0416	2.317	15.82
p=6	259.3507	227.7391	2.236	7.326	248.3567	224.2717	6.814	10.61

Notes: the lag order is selected on the basis of AIC and SBC and ♦ indicates the lag length choice according to the two criteria respectively.

Based on the previous discussion, we test for the absence of a long-run relationship between the national savings and the pension funds variables employing an F-test for the joint null hypothesis $\gamma_{1,i} = \gamma_{2,i} = 0$, under the alternative hypotheses that there is a stable long-run level relationship between the aforementioned variables.

[14] If the calculated test statistic (which is a standard F test for testing the null hypothesis that the coefficients on the lagged levels terms are jointly equal to zero) lies above the upper bound, the result is conclusive and implies that a long-run relationship does exist between the variables. If the test statistic lies within the bounds, no conclusion can be drawn without knowledge of the time series properties of the variables. In this case, standard methods of testing would have to be applied. If the test statistic lies below the lower bound, no long-run relationship exists.

[15] Pesaran, Shin and Smith, 2001, p. 296.

The bounds test results for the complete sample period are presented in Table 4 where Equation (3) is estimated and then the F- statistics are computed.

Table 4. Bounds test

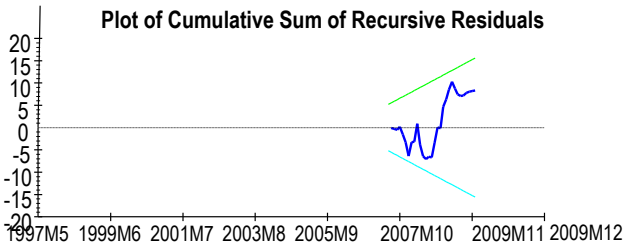
Unrestricted intercept and no trend			
	F-stat	Upper critical value	Upper critical value
K=4	F(3, 147)= 6.1213**	2.86	4.01

Notes: *t*he *F*-statistic is used to test for the joint significance of the coefficients of the lagged levels in the ARDL-ECM. Asymptotic critical values are obtained from Table CI(iii) Case III: unrestricted intercept and no trend for k=1 and K=2 and from Table CI(iv) Case IV: unrestricted intercept and restricted trend for k=1 and K=2 (Pesaran, Shin and Smith, 2001, pp. 300-301).

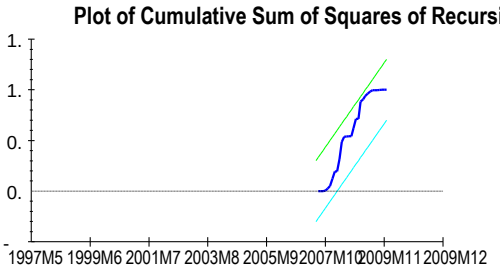
** indicates that the statistic lies above the 0.05 upper bound; * that it falls within the 0.10 bounds; • that it lies below the 0.10 lower bound.

Using the asymptotic critical value bounds computed by Pesaran, Shin and Smith (2001), the F-statistic lies above the 0.10 upper bound. Hence the null hypothesis of no long-run relationship is rejected. At this stage of the ARDL method, it is also possible to perform a parameter stability test for the appropriately selected ARDL representation of the model. The stability of coefficients are, by and large, tested by means of Hansen (1992), and Hansen and Johansen (1993). The Chow stability test requires *a priori* knowledge of structural breaks in the estimation period and its shortcomings are well documented.

In Hansen (2002) and Hansen and Johansen (1993) procedures, stability tests require I(1) variables and they check the long-run parameter constancy without incorporating the short-run dynamics of a model into the testing – as discussed in Bahmani-Oskooee and Chomsisengphet (2002). However, it is possible to overcome these shortcomings by employing the Brown *et al.* (1975)[16] procedure if we follow Pesaran and Pesaran (1997). These tests are usually implemented by means of graphical representation. It can be seen from figure 2 the plots of CUSUM and CUSUMSQ statistics are within the critical bounds implying that the coefficients in the model are stable.



The straight lines represent critical bounds at 5% significance level



The straight lines represent critical bounds at 5% significance level

Figure 2. CUSUM and CUSUMSQ plots for Stability Tests

[16] The Brown *et al*. stability testing technique, also known as cumulative sum (CUSUM) and cumulative sum of squares (CUSUMSQ) tests, is based on the recursive regression residuals. The CUSUM and CUSUMSQ statistics are updated recursively and plotted against the break points of the model. Providing that the plot of these statistics fall inside the critical bounds of 5% significance then we assume that the coefficients of a given regression are stable.

Once established that a long run cointegration exists, Equation (3) is estimated. A maximum lag order of one is allowed in the ARDL model using the Schwarz Bayesian Criteria since it presents the smaller residual sum of squared. The following ARDL (4,3,2,3) specification is used.

Table 5 reports the results of the ARDL where the regressions fit reasonably well and pass the main diagnostic tests. All levels estimates are highly significant and have also the expected signs. The results suggest that there is a strong evidence between what the theory predict (about the relationship between savings and pension funds) and the estimated variables. The constant term is positive and highly significant at 0.01 level.

The estimated retirement savings variables (RPENS, RPAYG and RINS) show the coefficients with correct signs and are statistically significance at 0.01 and 0.05. The variable DR is with the right signs and statistically significant at 0.01 level.

Table 5. ARDL (4,3,2,3) Equation (3)

Section 1, Short-run coefficient estimates					
Lag order		**1**	**2**	**3**	**4**
ΔSAV			0.6874**	0.3213**	-0.1703*
			(8.971)	(3.466)	(-1.972)
ΔRINS				7.9250**	
				(4.0295)	
ΔRPENS			0.25532♦		
			(1.6884)		
ΔRPAYG				9.9141**	
				(4.0232)	
ecm		-0.028164♦			
		(-1.8313)			
Section 2, Long-run coefficient estimates					
C	**RINS**	**RPAYG**	**RPENS**	**DR**	
1.679842**	0.62959*	-0.93436*	0.09025*	-0.128499**	
(3.156399)	(2.352)	(-2.3852)	(1.689)	(-4.0493)	
Dummy4					-0.04248*
					(-2.2338)
Section 3 Diagnostics					
Adjusted R-squared: 0.84591; Durbin-Watson stat: 1.8272;					
Serial Correlation $\chi^2_{SC}(2) = 1.412[0.235]$;					
Functional Form $\chi^2_{FF}(1) = 0.52528[0.469]$;					
Normality $\chi^2_N(2) = 6.847[0.6321]$;					
Heteroscedasticity $\chi^2_H(1) = 0.33334[0.564]$.					
♦significant at the 0.10 level; *significant at the 0.05 level; **significant at the 0.01 level ; *Obs. 152 (* monthly)					

The results of Equation(3) for the short run estimations show the complex dynamics that seem to exist between changes in saving and changes in the pension funds. The coefficients are all significant. Among them, ΔRPENS is the only one statistically significant at 0.10 level.

The equilibrium correction coefficient (*ecm)*, estimated (*-0.028*) is statistically significant at 0.10 per cent and has the correct sign, implying that a deviation from the long-run equilibrium, following a short run shock, is corrected by about 2.8 per cent after one month. Finally, Table 5 (Section 2) shows the results of the dummy4 used[17]. It is statistically significant at 0.01 level and the coefficients' value is -0.0242.

Overall, our results are consistent with the ones by Granville and Mallick (2004) that found evidence suggesting that increases in mandatory pension funds financial assets increase national saving.

As regards the relation between dependency ratio and changes in national saving, it should be noted that population ageing will itself generate changes in saving which may have a major macroeconomic impact. According to the life cycle theory, savings rates tend to decline in countries where there are larger numbers of

[17] We made estimations including all the dummies selected in 4.2.2. and dropped those not significant.

retired people (Disney 2000). The changes in savings lead to changes in demand for financial assets. Econometric studies find a strong effect of aging population on private saving as in Masson *et al*. (1998) who found the total dependency ratio to have a significant negative effect on private saving in a panel of both advanced and developing countries.

4.2.4. The Kalman filter

In this final section we estimate, using a Kalman filter approach, a 'backward-looking' process for national savings with parameters varying with fundamental component of the private pension plans. The choice of this methodology is stringently related to the economic evolution of Iceland within the past twelve years. As stressed in Section 3.1, Iceland has faced at least with three main different economic periods (1997-2004, 2004-2006, and 2006-2009). Therefore, it is important to analyse the timing path of each single coefficient of the explanatory variables. The Kalman filter is a recursive procedure of computing the optimal estimator of the state vector at time *t*, based on the information available at time *t*. One of the reasons for the central role of the Kalman filter is that "[...] when the disturbances and the initial state vector are normally distributed, it enables the likelihood function to be calculated via what is known as the prediction error decomposition. This opens the way for estimation of any unknown parameters in the model"[18]. The time varying methodology allows us to recover an unobservable factor that could affect national savings. For each endogenous variables of the model it is therefore possible to observe how the respective coefficients have changed over time.

Assuming that saving, SAV_t, is driven by a AR(n) process, we apply the following time varying parameters model:

$$SAV_t = \beta_{0,t} + \beta_{1,t} RPENS_{t-n} + \beta_{2,t} RPAYG_{t-n} + \beta_{3,t} RINS_{t-n} + \mu_t \qquad (5)$$

where μ_t is an independent white noise, the coefficients are assumed to be random walks. This can be written in state space form where the observation Equation is given by (5) above and the state Equations are given by:

$$\begin{bmatrix} \beta_{0,t} \\ \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{0,t-1},\ldots\beta_{0,t-n} \\ \beta_{1,t-1},\ldots\beta_{1,t-n} \\ \beta_{2,t-1},\ldots\beta_{2,t-n} \\ \beta_{3,t-1},\ldots\beta_{3,t-n} \end{bmatrix} + \begin{bmatrix} \mu_t \\ \mu_{1t} \\ \mu_{2t} \\ \mu_{3t} \end{bmatrix} \qquad (6)$$

Where Equation (6) is the measurement Equation with S_t and ε_t are *[nx1]* vectors[19]. The relevant results and estimates are reported in Figure 3 and Table 6.

The pattern of the coefficients "*β*" seems to capture well the economic dynamic of the period. The most important result is that the model "works" as shown in figure3 in which the behaviour of the pension funds for Iceland vary over time. In particular, the coefficients β_1 (RPENS) illustrates a shift upward soon after the launch of the reforms in 1998. Later on, this coefficient shows a negative trend till the middle of 2004 and it increases sharply until the beginning of 2006. Afterwards, following the Icelandic and international financial crisis, β_1 strongly declines.

Table 6. The Kalman estimations

	$\beta_{1,t}$	$\beta_{2,t}$	$\beta_{3,t}$	$\sigma^2_{\mu,t}$
AIC=13.41	0.4602**	- 0.7565**	0.0686*	-0.485**
Schwarz=12.42	(93.932)	(-230.68)	(148.212)	(-217.92)
Obs. 155 (M)	[0.000]	[0.000]	[0.000]	[0.000]

*significant at the 0.05 level; **significant at the 0.01 level ; z-statistics in brackets; p-value in squared brackets; (M)=monthly

[18] Harvey (1989, p. 10).

[19] For a more complete explanation of the Kalman filter approach, the state space form and the measurement and transition equations, see Harvey (1989).

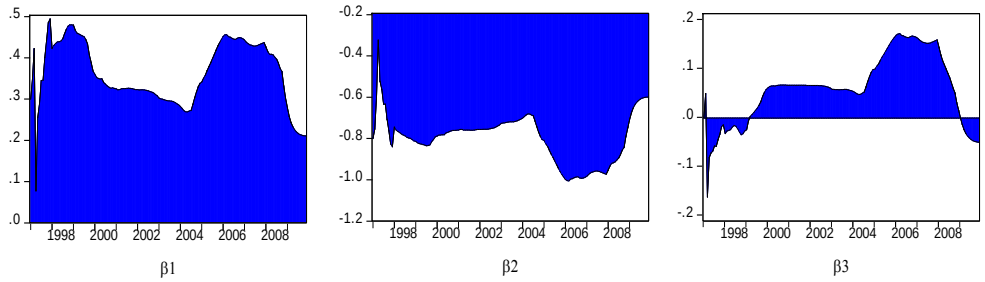


Figure 3. Time varying coefficients for Iceland

As for β_1, the coefficient β_3 (RINS) shows a similar trends for the same periods. It rises sharply reaching the peak at the beginning of 1999 and it remained stable till the end of 2003. Afterwards, it continues to rise till 2006 when it starts to decline dramatically inverting the sign in 2009. A possible explanation of this positive trend (1998-2006) of the value of β_3 could be related to tax incentive legislation adopted in 1998 as part of the general social security reform.

Finally the coefficient β_2 (RPAYG) shows a quite stable trend from 1998 till 2004 in which it sharply decline reaching its minimum value of -1.03 and inverting its path soon after.

It is worth noting that β_2 path is specular to the paths of β_1 and β_3. This could be due to the opposite impact on national saving of an unfunded and a funded pension benefits as asserted by the economic theory (Feldstein 1974, 1996, and Samwick 2000).

5. Policy Implications and Conclusion

According to Munnell (1982) a perfect world should be characterized by perfect capital and labour markets, no uncertainty, no taxes and, in this framework, pension saving should be a perfect substitute for other kinds of savings. Unfortunately we live in a not perfect world in which there are imperfect markets, taxation, uncertainties and several different opportunities for savings. Because of that, the complete substitution of pension saving for other forms of savings may not occur. Hence, should be reasonable to expect that countries which have funded occupational pension scheme would have higher national savings rate than countries which do not.

In this work we have tried to examine the pension systems of Iceland in a more general perspective with the purpose to analyse how its retirement income provisions have evolved into a multi-pillar arrangement and its impact on national saving. Moreover, we would like to ask – without completely answering the question – if there are important lessons for others in Iceland's pension experience.

Icelandic occupational fully funded pension plans have become just as important as the public pay-as-you-go system, while the latter is the dominant pillar in many other OECD countries. It is also worth noting the differences in aging population structure and in social security history of this country as describe in section 3.2.

The novelty aspects of our results can be summarized as follow.

Firstly, in this work a combinations of different econometric methodologies have been implemented. The innovative use of these methodologies allowed us to overcome the critics about the ambiguous impact of social security variables on national saving and to get time varying coefficients in order to show the reaction of national saving to the implementation of private second pension pillar.

Secondly, we have found substantial evidence that a mandatory pension fund has a positive impact on national saving in Iceland. Moreover the results showed that unfunded pension pillar had a negative influence on national saving as predicted by the theory. These findings are consistent with the current empirical literature. Moreover, using the Kalman filter methodology we were able to show that, the pattern of the coefficients of private pension, pay-as-you-go and life insurance, capture the economic dynamic of the period.

Nevertheless, the following two elements can be considered as the main policy implications for others.

As discussed in section one, the results exhibit that the social security reform that takes the form of three-pillar system is, in particular, beneficial for small open economies. In fact, in small open economy, the effects on financial market development are much clearer due to the contribute of inflow of capital on pension funds. This effect can be different in magnitude depending on the degree of openness and the financial integration of the

economy. For instance, in Iceland the percentage of foreign assets of pension funds moved from 2 percent in 1990 to 16 percent in 2002. As it is confirmed by our empirical analysis, the change of magnitude of these assets has a significant effect on national saving.

In addition, our empirical findings showed that a mandatory retirement saving policy could have improved the financial sustainability of social security systems and, through the positive impact on national saving, could be seen as an additional engine of economic growth.

Noticeably, general principles will always have to be applied with a great degree of attention to the specific economic situation in each country. In fact the implementation of the funded pension system started in a period characterized by a demographic crisis and high returns of financial assets. This made necessary to move, completely or partially, to a reformed pension system.

However we have found that, despite the strong impact of the 2007 financial crisis on the Icelandic economy, the reformed system seems to remain valid as well as the cointegration relationship with the national savings. Nevertheless we do not know how the pension system will evolve in the future. That is why, in light of the current financial crisis, it should be interesting to investigate the reaction of the new designed pension systems to changes into the economic environment. In other words, what will be the future of the second funded pension pillar in a more risky and volatile financial markets?

References

[1] Arrau, P., and Schmidt-Hebbel, K. 1999. Pension System and Reforms – Countries Experiences and Research Issues, *Policy Research Working Paper*, n. 1470, The World Bank, June.

[2] Auerbach, A., and Kotlikoff, L. 1987. *Dynamic Fiscal Policy*. Cambridge: Cambridge University Press.

[3] Bahmani-Oskooee, M., and Chomsisengphet, S. 2002. Stability of M2 Money Demand Function in Industrial Countries. *Applied Economics* vol. 34(16): 2075-83.

[4] Bailliu, J., and Reisen, H. Do funded pensions contribute to higher aggregate saving? A cross country analysis. *Working Paper* no. 130, OECD Development Center.

[5] Brown, R.L., Durbin, J., and Evans, J.M., 1975. Techniques for testing the constancy of regression relations over time, Journal of the Royal Statistical Society, vol.37:149-163.

[6] Cagan, P. The effect of Pension Plans on Aggregate Saving: Evidence from a Sample Survey, *NBER*, New York.

[7] Castellino, O., and Fornero, E. 1997. Privatizzare la previdenza sociale? Condizioni, modalità e limiti. *Politica Economica* vol. 13, iss. 1:3-25.

[8] Cesaratto, S. 2006. Transition to Fully Funded Pension Schemes: A Non-orthodox Criticism. *Cambridge Journal of Economics* vol. 30, n.1: 33-48.

[9] Cifuentes, R., and Valdes-Prieto, S. 1997. *Transitions in the Presence of Credit Constraints. In: The economics of pensions: Principles, policies, and international experience*, Cambridge University Press, 160-89; Cambridge; New York and Melbourne.

[10] Disney, R. 2000. Declining public pensions in an era of demographic ageing: will private provision fill the gap. *European Economic Review* 44: 957-73.

[11] Feldstein, M. 1974. Social Security, Induced Retirement and Aggregate Accumulation. *Journal of Political Economy* 82, n. 5; 905-926.

[12] Feldstein, M. 1996. The Missing Piece in Policy Analysis: Social Security Reform. *American Economic Review*, Papers and Proceedings 1996; 86; 1-14.

[13] Granger, C.W.J. 1986. Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics* 48: 213-228.

[14] Granville, B., and Mallick, S. 2004. Fisher hypothesis: UK evidence over a century. *Applied Economics Letters* 11, 87-90.

[15] Gudmundsson, M., and Baldursdóttir, K. 2005. Pension funds – Future prospects and uncertainties. *Monetary Bulletin*, Central Bank of Iceland, 4.

[16] Gudmundsson, M. 2001. *The Icelandic Pension System*. National Association of Icelandic Pension Funds.

[17] Gudmundsson, M. 2004. *The Icelandic pension system: Design and lessons for others*. Monetary Bulletin, Central Bank of Iceland, 1.

[18] Hansen, B.E. 2002. Tests for Parameter Instability in Regressions with I(1) Processes. *Journal of Business & Economic Statistics, American Statistical Association* 20(1): 45-59.

[19] Hansen, H., and Johansen, S. 1993. Recursive estimation in cointegrated VAR models. *Working Paper, Institute of Mathematical Statistics*, University of Copenhagen 1993.

[20] Harvey, A. 1989. *Forecasting structural time series models and Kalman filter*. Cambridge University Press.

[21] Holzmann, R. 1997. Pension Reform, Financial Market Development, and Economic Growth: Preliminary Evidence from Chile. *International Monetary Fund Staff Papers* vol. 44, iss.2:149-78.

[22] Hviding, K., and Merette, M. 1998. Macroeconomic effects of pension reforms in the context of ageing populations: overlapping generations model simulation for seven OECD countries. *OECD* Working paper, 14.

[23] Katona, G. 1965. *Private Pensions and Individual Saving. In Ann Arbor: Survey Research Center,* Institute for Social Research, The University of Michigan.

[24] Kotlikoff, L.J., Smetters, K.A., and Walliser, J. 1998. Social Security: Privatization and Progressivity. *American Economic Review* 88, iss. 2:137-41.

[25] Kotlikoff, L.J. 1996b. Simulating the Privatisation of Social Security in General Equilibrium. *NBER Working Paper* n.5776.

[26] Kotlikoff, L.J. 1996a. Privatising Social Security at home and abroad. *American Economic Review*, Papers and Proceedings 86: 1-14.

[27] Masson, P.R., Bayoumi, T., and Samiei, H. 1998. International Evidence on the Determinants of Private Saving. *World Bank Economic Review*, Oxford University Press 12(3): 483-501.

[28] McKenzie, C. 1997. Unit Roots and Cointegration Analysis: The Impact on Empirical Analysis in Economics. *Japanese Economic Review* March: 18-28.

[29] Mitchell O.S, Zeldes S.P. Social Security Privatization: A structure for Analysis. American Economic Review 1996; vol.86, n.2; 363-367.

[30] Munnell, A. 1982. *The Economics of Private Pensions*. Washington, D.C., The Brookings Institution.

[31] Murphy, L.P., and Musalem, A.R. 2004. Pension funds and national saving. *Policy Research Working Paper Series* 3410, The World Bank.

[32] OECD. 2005. Pension Markets in Focus. *OECD Newsletter*, December 2005, Iss.1-2.

[33] Orszag, P.R., and Stiglitz, J.E. 2000. Rethinking Pension Reform: Ten Myths about social Security System. In: Holman R, Stiglitz J. *New Ideas About Old Age Security*, Washington, World Bank: 17-56.

[34] Pesaran, M.H., and Pesaran, B. 1997. *Working with Microfit 4.0: Interactive Econometric Analysis*. Oxford University Press.

[35] Pesaran, M.H., Shin, Y., and Smith, R.J. 2001. Bounds Testing Approaches to the Analysis of Levels Relationships, *Journal of Applied Econometrics*, Special Issue, in honor of JD Sargan on the theme "Studies in Empirical Macroeconomics", (eds) D.F. Hedry and M.H. Pesaran 16: 289-326.

[36] Pesaran, M.H., Shin, Y., and Smith, R.J. 1999. An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis. In (ed) S. Strom. *Econometrics and Economic Theory in the 20th Century*. The Ragnar Frisch Centennial Symposium, Chapter 11, Cambridge University Press.

[37] Samwick, A. 2000. Is Pension Reform Conducive to Higher Saving. *The Review of Economics and Statistics*; 82(2): 264-272.

[38] World Bank. 1994. *Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth*. Washington DC: World Bank and Oxford University Press.

GLOBAL SUPPLY CHAINS AND THE GREAT TRADE COLLAPSE: GUILTY OR CASUALTY?

Hubert ESCAITH[20]
World Trade Organization, **Switzerland**
hubert.escaith@wto.org

Abstract

With globalization, trade and production have been increasingly interlinked, thanks to the vertical integration of industrial production processes through outsourcing and off-shoring. The expansion of international supply chains determined the apparent increase in trade elasticity observed since the late 1980s, and may explain also the overshooting of trade elasticity during the 2008-2009 trade collapse. After reviewing the available evidences, the article analyses the future of globalized production networks in a post-crisis scenario. In the short term, global rebalancing might prove easier than expected, because trade in intermediate goods inflated artificially some bilateral trade deficits, albeit bilateral exchange rate adjustments have reduced impacts. But supply chains may become smaller and more regional as a result of this rebalancing. This scenario creates a challenge for labour abundant less advanced developing countries in the periphery of the large regional networks, which will find more difficult to attract productive investments. Yet deglobalization remains a distant threat as long as the technical and institutional factors that made possible the internationalization of production are preserved.

Keywords: global supply chains, trade, outsourcing, off-shoring, trade elasticity, internationalization of production

JEL Classification: F42, E32, F23, O24, O19, G01

1. Global supply chains and the great trade collapse: guilty or casualty?

International trade and the nature of globalization have changed dramatically in recent years, with the emergence of new global players and a radically different competitive landscape. This new pattern emerged during the late 1980s and early 1990s, when the Berlin Wall fall brought down the barriers that had split the post-WWII world, and the Brady Bonds put an end to the decade-long debt crisis that plagued many developing countries. The 1990s saw the conclusion of the Uruguay Round and the birth of the WTO, which brought down many trade barriers and led to further liberalization in areas like telecommunications, financial services and information technologies.

This transformation, which was both geopolitical and economic, was accompanied by the emergence of new business models that built on new opportunities to develop comparative advantages (Krugman 1995; Baldwin 2006). With the opening of new markets, the technical revolution in IT and communications, and the closer harmonization of economic models worldwide, trade became much more than just a simple exchange of merchandise across borders. It developed into a constant flow of investment, of technologies and technicians, of goods for processing and business services, in what has been called the "Global Supply Chain".

Those changes have led the American author Tom Freidman to proclaim that now "The World is flat". This "Copernican revolution" where countries do not trade wine for clothes anymore (Grossman, and Rossi-Hansberg 2006) obliged trade analysts and international economists to revise their old beliefs and models, while trade statisticians and national account specialists were struggling to adapts their instruments to the new reality (Escaith 2008). Despite important advances, the analysis is still lacking appropriate models and good data to understand and measure appropriately this new dimension of globalization. Yet, less than twenty years after its emergence, this new business model is now challenged on two grounds: firstly, because global supply chains may have been a causing factor of the Great Recession which followed the financial collapse of September 2008; secondly because a return to business as usual after the international crisis is not possible on objective economic ground, nor it is desirable for normative social or environmental reasons.

Indeed, if the cause of the global crisis is clearly financial, its particular mode of dissemination through real channels has been related to the interdependency created by global productive networks, which served both as transmission and amplification channels. While the crisis spread initially to all developed countries through toxic

[20] Economic Research and Statistics Division-WTO, Geneva; DEFI, Aix-en-Provence. The essay builds on a series of related works which benefited from the contribution of Ch. Degain, F. Gonguet, N. Lindenberg, A. Maurer and S. Miroudot. The views expressed in this document are those of the author and do not represent a position, official or unofficial, of the WTO Secretariat or WTO Members.

assets and exposure to the US financial market and seemed to spare the developing economies, the premise of a new North-South "decoupling" vanished rapidly when the contagion spread to the real economy through the trade collapse.

The amplitude and simultaneity of the transmission of shocks came as a surprise to many analysts. International supply chains, one of the most salient features of the "new globalization" were rapidly identified as one of the main factors for such a synchronization of shocks. With unemployment increasing as recession spread in developed countries, the debate was also put on the public place as the delocalization of investment and jobs that rests behind these new productive networks, together with the lack of governance of the international finance, became the focus of much public scrutiny.

Since the every-day reality of international business models is running faster than the elaboration of new analytical paradigms, the crisis hit the global economy in largely uncharted waters. Guided by "old" economic models, the major developed and developing economies, known as G-20, met head-on the risk of a world-wide depression and coordinated global responses in the early months of 2009. The package, reminiscent of traditional Keynesian recipes, addressed principally the macro-economic transmission channels through massive fiscal stimulus. With the risk of global recession diminishing after the second quarter of 2009, the debate has now shifted to the exit strategies.

The huge fiscal deficits have sustained public consumption in industrialized countries, but private consumption and investment remains depressed. If the danger of inflation seems under control, rebalancing the current account imbalances which characterized the pre-crisis period would create an additional negative shock, as high spending countries would have to increase their national savings. The alternatives medium-term forecasts range from "back-to-business-as-usual" to "deglobalization" scenarios, producing an alphabetical string of V, U, L or W profiles.

Against this background, the present article explores in a first section the particular role of supply chains in transmitting and amplifying external shocks. A second part analyses their potential responsibility in the 2008-2009 trade collapse, and a third one is dedicated to exit scenarios and the perspective of global supply chains after the crisis, including some of the consequences of global rebalancing for less advanced developing countries.

2. Supply chains as transmission channels

Like in previous global financial crisis, the international banking system came to a "sudden stop" after September 2008. Two aspects were nevertheless original: the shock emanated from the largest world financial centre instead of initiating in developing countries, and the shocks spread very quickly and almost simultaneously to many industrial and emerging countries. In particular, trade reacted very strongly to the first signals of recession, and sectors were differently affected. The sectors most affected by the recession were fuels and minerals (due to a strong price effect), and machinery and transport equipment (strong demand effect).

With the financial crisis, the sectors producing consumer durable and capital goods were on the front line, as demand for these products relies on credit. In turn, the lower industrial activity reversed brutally the trend in the prices of key primary commodities, which had been rising substantively since 2003. Between the third and the fourth quarter of 2008, the difference in growth rate is, respectively, 41 and 37 percentage points for iron and steel, and for minerals (Table 1). The collapse in trade mostly affected merchandises; except the financial transactions, the commercial services, other than those related to trade in goods, were more resilient.

Indeed, world trade dropped five times more rapidly than global GDP, supply chains playing their part in explaining the magnifying effect of the crisis on international trade. Some of the mechanisms are purely of accounting nature: while GDP is computed on a net basis, exports and imports are registered on their gross value. In addition, because supply chains cover various countries, a lot of double counting takes place while goods for processing cross the borders at each step of the production process. But the core of the explanation is to be found in the nature itself of the 2008-2009 crisis.

International transmission of shocks takes usually two forms, commercial and financial. In previous instances of global crisis, most of the systemic commercial and financial shocks where of macroeconomic nature. A recession in a foreign economy reduced demand for exports, which in turn depressed the activity in the home country. The propagation of such demand-driven shocks through the productive sectors of the home economy can be traced using an input-output model, through traditional Input-Output modelling. In addition, both financial and real channels are interlinked at the macro level, because credit crunch affects household consumption and firms' investment.

Table 1. Quarterly growth of world manufactures exports by product, Q1/08-Q3/09
(percentage change over previous quarter, current dollar values)

Quarter/Sectors	Q1/08	Q2/08	Q3/08	Q4/08	Q1/09	Q2/09	Q3/09	Q4/09
Manufactures	-1	9	-2	-15	-21	7	9	9
Office and telecom equipment	-12	5	5	-10	-27	13	14	16
Automotive products	1	6	-14	-18	-33	14	13	21
Iron and steel	11	23	7	-34	-31	-8	10	10
Ores and other minerals	10	20	4	-33	-35	13	24	7

Source: WTO.

The gradual substitution of trade in goods by trade in tasks that took place during the 1990s have changed this traditional mode of transmission, and added another layer of transmitters which are operating at micro economic level but have also an international dimension because of the geographical segmentation of the productive chains. When industrial production is spread across various countries, and that all segments of the chain are critical (supplied constrained), a shock affecting one segment of the chain will reverberate through all the chain. At the difference of the macro-economic case, shocks are moving forward, from supplier to clients, and not backward as in the demand-driven Leontief model (from client to suppliers).

These effects are still largely unknown as they affect firms that are eminently "heterogeneous", as recognized by the new "new trade theory". It is a cliché to say that time is accelerating, but it carries a lot of truth in the present situation. In the race between the practitioners –engineering and business schools – on the one hand, and the university, on the other hand, praxis is well ahead despite significant advances in the theoretical aspects (Box 1).

On the empirical side, Escaith and Gonguet (2009) (E-G09 thereafter) jointly models the real supply-side and financial effects from a complementary viewpoint of monetary circuit and international Input-Output (I-O) matrices. The rest of this section builds on their results.

The monetary circuit: The financial aspects are introduced through endogenous money. In order to produce, individual firms need to obtain a loan from a bank. The bank grants the loan in relation to three parameters: the macroeconomic context, the specific behaviour of the sector of activity in the business cycle, and the specific situation of the firm (credit rating, soundness of the management). Credit money is created by the bank when according the loan, and is spent by the firm on wages and other production costs. The money remains in the circuit as long as the firm does not sell the products and reimburse the loan. In particular, a traditional result of the endogenous money theory is that any increase in stock of credit money corresponds to an increase in inventories in the national account circuit.

E-G09 adds to this classical building block [21] a late XXth century feature: the capital-asset adequacy ratio, a prudential mechanism –such as Basel II– set by the authorities and designed to guarantee liquidity and solvability of the banking sector. At the difference of monetary circuit and I-O tables, which track flows, the adequacy ratio is a stock variable reflecting the accumulation of loans and assets. Under normal conditions, the ratio is not binding and the circuit is almost a pure flow model. Money is endogenous to investment, as banks can modulate their assets to accommodate new credits.

Box 1 The Microeconomics of Supply Chains and Trade in Tasks

The question of what steers imports and exports has a very long history in the economic literature. Traditional normative trade theory is built on the premises of gains from international trade, based on the Ricardo's arguments of comparative advantages which allow to increase output by reallocating resources within countries. The Hecksher-Ohlin model extends the results when there are many productive factors and different factor endowments which create gains from international specialization. Economic theory points also to other sources of gains which are not liked to comparative advantages, for example access to a wider variety of goods and economies of scale on the consumption and production sides, or contesting monopoly powers on the institutional one. The critic and testing of these traditional hypotheses has led to a vast literature (see WTO 2008 for a review).

[21] Even if endogenous money and sectoral modelling seems quite heterodox now-a-days, both monetary circuit and supply-use tables can be traced to the Physiocrats.

Trade in tasks and the fragmentation of production along global supply chains has challenged the validity of the traditional models, based on the exchange of final goods. As for trade in intermediate goods, two main approaches have been explored. The first one rests on the hypothesis that factors (capital and labour) are fixed. Offshoring is similar to technical progress in the production of the final good: by shifting production processes to the countries with comparative advantages, more of the final goods can be produced with the same (fixed) amount of factors. A firm with better technology at home wills offshore some tasks if the initial wage gap is larger than the offshoring costs. Offshoring releases domestic workers who, under the traditional neo-classical market assumptions, can focus on the tasks where they have a trade-cost-adjusted comparative advantage.

When some factors are mobile sectorally and internationally, the law of comparative advantage can be generalized, albeit analyzing the outcome of even simple normative models becomes a complex matter, particularly regarding the distributional effects. Comparative advantages are no more a robust predictor of a country's trade pattern. For example, reversal of comparative advantages is possible under certain assumptions. When factors are mobile, movement of factors between sectors and between countries that tend to equalize endowments reduce the incentives to trade. The return of the mobile factor rises, while those of to sector-specific factors decline. But if factor prices are not equalized, factor mobility will at contrary lead to an increase in the volume of trade.

The new trade theory, by introducing imperfect competition, consumer preference for variety and economies of scale, look at explaining why countries that are similar in factor endowment and technology, have a significant part of their trade in the same industries. When two such identical countries open up to trade, firms with differentiated products gain access to larger markets and offer more choice to consumers. While consumers face a greater choice of products, product differentiation imparts firms with a degree of market power. But the entry of new firms modifies the way consumers substitute between products as prices and product offerings change, and reduce the market power of the firm. In addition, the total size of the market does not increase, and some firms will go out of business.

A similar pattern can be used to model trade in intermediate goods, provided the cost of production is lower the larger is the number and scale of production of intermediates. If trade is restricted, the domestic firm needs to use outsourcing to realize these economies of scale and choice. For example, firms (and governments) may solicit bids to supply goods or services in context of information asymmetry and imperfect competition. If international trade is free, it does not matter anymore where the production of intermediaries is located (besides the additional transaction costs), and the firm producing the final good uses a mix of outsourcing and offshoring. Even in presence of free trade, it is expected that monopolistic behaviour tend to predominate, as organized supply chains (business groups) exercise market power in their sales of intermediate inputs. Feenstra and Hamiston (2006) relate that firms belonging to such groups treated their member firms preferentially, buying and selling at prices different than those used for non-member firms.

The real circuit: The industrial supply side aspects and the transmission of real shocks across sectors and countries is modelled by E-G09 using an international input-output matrix (a set of interlinked national I-O matrices), rearranged to track forward linkages.[22] E-G09 computes an indicator, called **"***imported real supply-driven impact coefficient***"** (IRSIC), defined in the price space to simulate cost-push due to intermediate inputs:

$$IRSIC = \Delta Q(I\text{-}B)^{-1} \bullet 1/Q \qquad (1)$$

where, for "*n*" countries and "*s*" sectors; *Q*: line vector of initial sectoral output (1 x *n.s*), B : matrix of [b_{ij}] allocation coefficients (*n.s* x *n.s*) [23], $(I\text{-}B)^{-1}$: Ghosh-inverse matrix, i.e. the sum of direct and indirect forward effects, Δ*Q*: line vector of supply shocks (initial increases in sectoral production costs emanating from the shock-exporter country) (1 x *n.s*), and ● 1/Q : Hadomard (entry wise) product.

[22] In an international *IO* matrix (*I-IO*), cross-national transactions of intermediate goods are identified: exports of intermediate consumption are separated from final demand and traced to the importing country and sector. This provides a better definition of vertical network participation of the countries included in the *I-IO*, allowing tracking shock transmission caused by vertical integration.

[23] In a supply-use table, inter-sectoral relationships are represented by the coefficients Q_{ij}, with the use of goods and services to supply other firms, final consumers and rest of the world (exports) on horizontal lines and requirements (purchases from domestic and foreign suppliers; primary inputs or value added) in columns. The *technical* coefficients used by the Leontief matrix represents interindustry sales by sector *i* and are derived by normalizing the intermediate coefficients Q_{ij} by the value of total production ($a_{ij}= Q_{ij}/Q_j$). The Ghosh matrix *B* also is build using supply-use table, but the *allocation* coefficients represent now the purchase by *j* of interindustry inputs from *i* ($b_{ij}= Q_{ij}/Q_i$).

Albeit the supply oriented Ghosh matrices are very similar to the Leontief demand-driven model, their theoretical robustness for modelling real shocks is much weaker, in particular because it cannot track substitution effects in presence of bottlenecks. For this reason, *IRSIC* is used only **1.** in the price space, and **2.** as a tracking mechanism. [24] Results based on an international *IDE* – Jetro *I-O* matrix updated for 2006 by the authors, and covering the USA, Japan, Korea and selected emerging Asian countries, indicate that:

1. In 2000 and 2006, Japan is the largest potential exporter of cost-push supply shocks, because it is a large supplier of intermediate goods to the other economies;

2. Malaysia and Thailand are the largest importers of such shocks, because of the high degree of integration of their manufacturing sectors in international supply chains and their reliance on imported inputs rather than domestic ones;

3. Between 2000 and 2006, China increased notably its forward international linkages and its domestic backward linkages. It became a large exporter of "shocks" in 2006, at par with Japan, but its vulnerability to an imported shock remained relatively stable because Chinese manufacturers are increasingly relying on domestic suppliers;

4. Repatriating the production of manufactured parts in Japan and the USA would lead to an average increase in sectoral production costs of 2%. Albeit this seems a small impact, it should be remembered that the input-output matrix are based on nation-wide samples of enterprises. In developed economies, most intermediate consumption is sourced domestically and only a minority of firms actively engage in outsourcing: this average impact would fall disproportionately on a few outward oriented firms, causing serious disruptions at microeconomic level.[25] Because these firms are also the most dynamic and innovative ones in a given sector, these microeconomic disruptions would have significant negative systemic effects.

Real-Financial resonance: The real and monetary circuit can interfere negatively when an unexpected financial shock, such as a credit crunch affecting production or trade finance, impede a foreign supplier to fulfil its contractual engagement. Confronted to such a disruption of their supply chain, client firms can shift to alternative partners. But shifting to an alternative supplier, when decision results from an unexpected event (a shock) takes time and always carries a cost. Higher production costs are transmitted along the global supply according to the *IRSIC* mechanism, reducing firms' profitability, while the longer production circuit increase demand for loans.

The accumulation of micro-disruptions in the productive chain, typical of a serious credit crunch in a large country, disturbs the international monetary circuit: production plans take longer to be completed, leading to an accumulation of outstanding loans, a decrease in profit margins. These microeconomic effects, combined with the overall negative macroeconomic outlook caused by the initial crisis, cause a reduction of the credit worthiness of firms. Confronted to this situation, banks have to adjust their asset holding in order to compensate for the higher risk of their loan portfolio and respect the Basel II capital adequacy ratio.

This is not an issue when financial markets are functioning normally, but in times of global crisis and flight to liquidity, not only the risk profile of borrowers deteriorates, but also the market value of assets goes down. Because assets are priced to market when evaluating the capital adequacy ratio, banks can rapidly been squeezed between the rising risk-rating of their debtors on the one hand, and the shrinking value of their asset portfolio on the other hand. When such situation arises, as happened after September 2008, the circuit unravels and the credit rivers run dry: banks look for safety, stop extending new credit and even do not renew existing credit lines. The very same pro-cyclical mechanisms that led to the apparition of financial bubbles, with the concomitant asset price inflation and lower perception to risk (meaning lower interest rates and larger volume of credit), can have a catastrophic outcome when the trend is reversed and a resonance effect between real and financial circuits amplifies the initial supply shocks.

Moreover, the accumulation of supply shocks leads to secondary demand-driven adjustments, either through a price effect (increasing production costs translate into higher retail prices and lower demand) or income effect (lower activity leading to unemployment). The succession of micro waves followed by secondary macro shocks leads to a "*W*" or "*L*" crisis patterns.

Crisis, exits and (de)globalization

Trade in tasks and the greater interconnections of the global economy have created, as we saw, newer and faster channels for the propagation of adverse external shocks. Because production is internationally

[24] The dual form of the Leontief can also be used to simulate cost-push due to primary inputs (Oosterhaven, 1996).

[25] In the United States, for example, in 2002 the top 1 percent of firms by size generated 81% of manufacturing exports, while the top 10 percent controlled 96% (WTO 2008).

diversified, adverse external shocks affect firms not only through final demand, but also through a rupture in the flow of inputs received from their suppliers.

3. Trade Collapse: the Role of Supply Chains

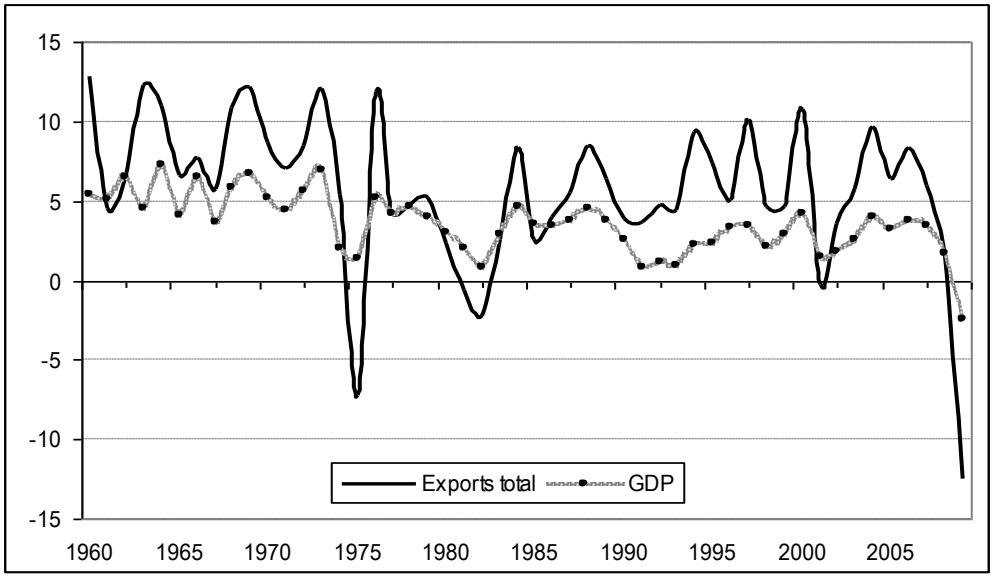


Figure 1. World merchandise exports and GDP, 1960-2009 (Real annual percentage change)

Source: WTO, International Trade Statistics and 2009 forecasts

▪ Trade Elasticity

Various authors attribute the large drop in trade registered since the end of 2008 (Figure 1), with an apparent trade-GDP elasticity larger than 5, to the leverage effect induced by the geographical fragmentation of production (Tanaka 2009; Yi 2009). Others contest the hypothesis of higher demand elasticity due to vertical integration (Benassy – Queré *et al.* 2009) because it affects only the relative volume of trade in relation to GDP, while elasticity should remain constant in a general equilibrium context. It is probable that the *observed* reality lays somewhere in-between the variable trade elasticity hypothesis and the constant one.

As seen in Figure 2, the world trade elasticity is shaped like an inversed "U", increasing at the end of the 1980s and decreasing in the most recent years. Because elasticity should indeed remain constant in an equilibrium context, this humped shape probably signals a long-term transition from one steady state to a new one.

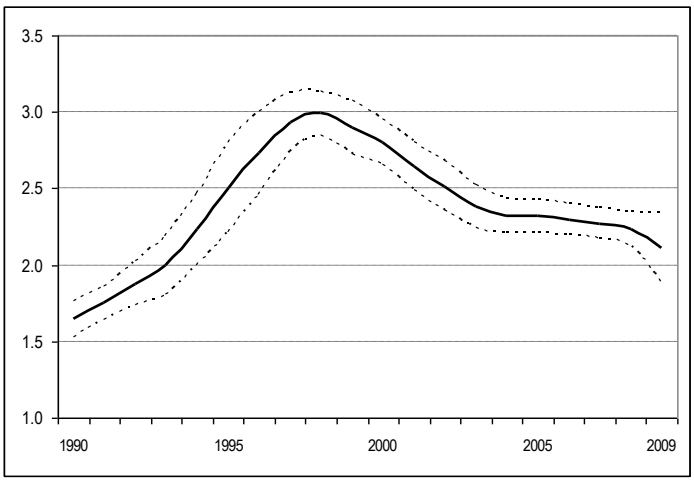


Figure 2. World: GDP Elasticity of Imports, 10 year average

Note: Rolling windows of 10 years; the date indicates the last year. World GDP is the sum of countries' data using market exchange rates; 2009 based on forecasts.

Source: Escaith, Lindenberg, and Miroudot (2010)

Short-term shocks can also affect *apparent* elasticity, even in presence of stable structural relationships. These short-term variations are to be expected when external shocks do not apply uniformly to all industries. A negative shock suffered by a single sector with high reliance on imported input will initially translate into a higher change in trade than in total GDP, leading to higher elasticity.

In absence of structural changes affecting production function (i.e., when technical coefficients, as described by an input-output matrix, are constant), the relationship linking demand for intermediate inputs with can be described by the following linear relationship:

$$\Delta M^{IC} = u \,.\, M^{\circ} .\, (I-A)^{-1} \,.\, \Delta D \qquad (2)$$

Where, in the case of a single country with "*s*" sectors): [26]

ΔM^{IC} : variation in total imported inputs (scalar)

u: summation vector (1 x *s*)

M°: diagonal matrix of intermediate import coefficients (*s* x *s*)

$(I-A)^{-1}$: Leontief inverse, where *A* is the matrix of fixed technical coefficients (*s* x *s*)

ΔD: initial final demand shock (*s* x 1) [27]

Similarly, changes in total production caused by the demand shock (including the intermediate inputs required to produce the final goods) is obtained from:

$$\Delta Q = u.\, A.\, \Delta Q + \Delta D \qquad (3)$$

Solving for ΔQ yields the traditional result:

$$\Delta Q = u.\, (I-A)^{-1} \,.\, \Delta D \qquad (4)$$

The comparison between Equations 2 and 4 is illustrative. Since $[M^{\circ} .\, (I-A)^{-1}]$ is a linear combination of fixed coefficients, the ratio $(\Delta M^{IC} / \Delta Q)$ is a constant.

Nevertheless, this tells only part of the story, because the initial shock ΔD is not a scalar, but a vector (*s* x 1), and the individual shocks affecting each element – a sector – do not need to be always in the same proportion from one year to another one. As the sectoral import requirements $[M^{\circ}_{s}]$ differ from industry to industry, and then the apparent import elasticity will change according to the sectoral distribution of the shock.[28]

It was in particular the case after the financial crisis of September 2008, as the demand of consumer durable and investment goods (consumer electronics, automobile and transport equipment, office equipment and computers, etc.) was particularly affected by the sudden stop in bank credits. Because these sectors are also vertically integrated, the impact on international trade in intermediate and final goods was high. And because services sectors, which are the main contributors to GDP in developed countries, were also more resilient to the financial crisis contribute, the drop in imports was much higher than the drop in GDP. Thus, in the initial phase of the financial crisis, the apparent Trade-GDP elasticity soared to 5.

When the initial shock reverberates through the rest of the economy, transforming the global financial crisis into a great recession, GDP will continue to slow down; meanwhile, the rate of decrease in trade will tend to stabilise as the import content of services sectors is much lower than those of manufacturing sectors. As can be seen in, there is a negative correlation between the variation in the volume of production of goods and the growth in total GDP that can be attributed, at least partially, to this time lag effect between goods and services. It is thus normal to expect a regression to normality of the trade elasticity for 2010.

[26] The model can be extended easily to the case of "*n*" countries, as in E-G09 by modifying accordingly the summation vector "*u*".

[27] In this traditional *IO* framework considering one country and the rest of the world, exports of intermediate goods are considered as being part of the final demand. The situation differs when extending the *IO* relationship to include international transactions of intermediate consumptions, as in Equation 1.

[28] The more complex the production process, the more potential for gaining in outsourcing part of it (Box 1); thus it is natural to expect much more vertical integration in the manufacturing sector. Miroudot and Ragoussis (2009) show that manufacturing sectors in OECD countries generally use more imported inputs than other industrial and services sectors. It is specially the case for final consumer goods like 'motor vehicles' and 'radio, TV and communication equipments', or computers. Services are, as expected, less vertically integrated into the world economy. But even these activities show an upward trend in the use of imported services inputs (e.g. business services).

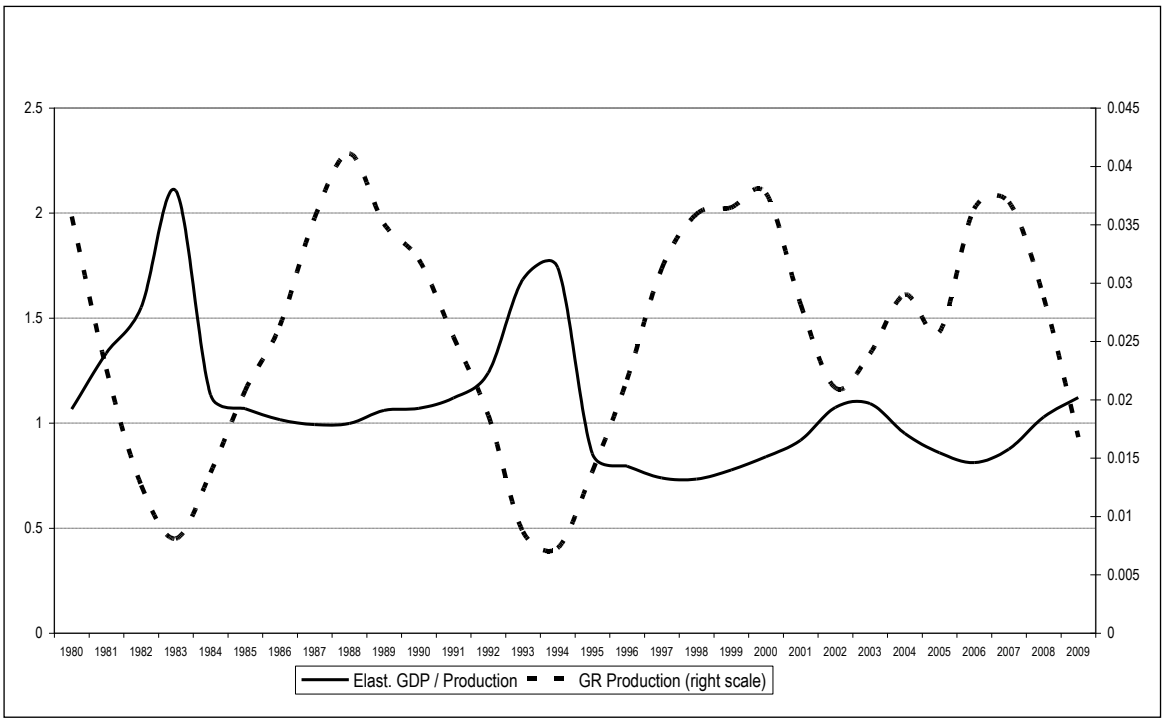


Figure 3. Delays in world production and GDP response, 1980-2009 (percentage growth and elasticity)

Notes: Five year rolling periods. Production includes agriculture, mining and manufactures.

Source: Based on WTO International Trade Statistics data base.

▪ **Inventory effects**

But recent changes in the *apparent* trade elasticity are also probably linked to global supply chain management practices. Even under "just-in-time" management (production-to-order), geographically fragmented networks need to maintain a minimum level of inventories (buffer stocks) in order to face the usual risks attached to international transportation. While large players try to keep their inventories at the lowest possible level considering their sales plans and the acceptable level of risk, they tend in the same time to force their suppliers to maintain large stocks (production-to-stock) in order to be able to supply them quickly upon request. In addition, some up-stream suppliers, engaged in highly capitalistic processes such as foundries, need to process large batches in order to benefit from economies of scale and lower their unit costs.

As a result, there is always a significant level of inventories in a global supply chain, translating into a higher demand for banking loans (Escaith, and Gonguet 2009). When a drop in final demand reduces the activity of down-stream firms, or/and when they face a credit crunch, their first reaction is to run down their inventories. Thus, a slow-down in activity transforms itself into a complete stand-still for the supplying firms that are located up-stream. These amplified fluctuations in ordering and inventory levels result in what is known as "bullwhip effect" in the management of production-distribution systems (Stadtler 2008). As long as the down-stream inventories have not been reduced to their new optimum level, suppliers are facing a sudden stop in their activity and must reduce their labour force or keep them idle.

The timing and intensity of the international transmission of supply shock may differ from traditional demand shocks applying on final goods. For example, the transmission index proposed by E-G09 implicitly assumes that all secondary effects captured by the Ghosh inverse matrix occur simultaneously, while these effects may propagate more slowly than traditional final demand shocks depending on the length of the production chain. Also, there might be contractual pre-commitments for the order of parts and material that manufacturers have to place well in advance in order to secure just-in-time delivery in accordance to their production plans (Uchida, and Inomata 2009). Indeed, since the 1990s, in high-tech manufacturing, suppliers are no more minor partners in global supply chains. These suppliers have consolidated, taking a more prominent role in the industry through a process of mergers and acquisitions (Sturgeon, and Van Biesebroeck 2009). Lynn (2009) provides an example from the US automobile industry where key suppliers of parts had a quasi-

monopolistic position and were able to impose their terms to the up-stream firms (automobile constructors).

Nevertheless, in closely integrated networks, these mitigating effects are probably reduced, especially when the initial shock is large. A sudden stop in final demand is expected to reverberate quickly thorough the supply chain, as firms run-down their inventories in order to adjust to persistent changes in their market. This inventory effect magnifies demand shocks and is principally to blame for the initial collapse of trade in manufacture that characterised the world economy from September 2008 to June 2009.

Dvorak (2009), reporting on the exposition of the electronic equipment sector during the crisis, mentions that a fall in consumer purchase of 8% reverberated into a 10% drop in shipments of the final good and a 20% reduction in shipments of the chips that go into the product. The velocity of the cuts, according to operators, was much faster than in previous slumps, as reordering is now done on a weekly basis, instead of the monthly or quarterly schedules that prevailed up to the early 2000s. In addition, previously, supply chains were simpler, involving fewer links; the complexity of today's productive networks makes their management much more complex. When faced with an unprecedented crisis such as the September 2008 one, "everybody under-bet to a certain extent", with forwards and backwards ricocheting effects through the supply chain, affecting in turn investment plans and capital goods providers.

Thus, the apparent change in trade elasticity results probably from a mixture of short-term and long term effects. Escaith, Lindenberg and Miroudot (2010) analyses the issue from a the dual angle of comparative static and dynamic models. They conclude that, in the short term, inventory and composition effects did predominate to explain the trade collapse of 2008-2009 and were certainly accentuated by global supply chains. But long term effects on trade elasticity are unclear. As mentioned (Figure 2) aggregate data point to a rise of trade elasticity during the 1990s due to globalization, with a return to normal in the late 2000s once the new global trade pattern is in place. Yet, this long term variation is not robust to alternative specification, when long term and short term effects are modelled simultaneously and data is disaggregated at country level.

4. Exit scenarios and global supply chains

The debate on the role of vertical specialization in shaping globalization and international trade is central for understanding the present crisis, but even more crucial for analysing alternative exit scenarios. In the second half of 2009, there were signs that the crisis was reaching a bottom. But analysts remained divided on the medium-run prospects, offering a menu of alphabetical potages made of *L*, *U*, *V* and *W*s. The last three scenarios are roughly based on a return to normal, after a period of recession that could be short (*V*), long (*U*) or bumpy (*W*). The *L* scenario is more pessimistic for international trade, as it involves a lasting deterioration from the high levels of globalization registered during the 1990s and the 2000s.

After the collapse of world trade in 2008-2009, and with the rise of murky protectionism as well as a higher risk aversion after the crisis, the risk is that manufacturers abandon global strategies to repatriate their operations domestically, or maintain them within a closer regional perspective. The globalization process may effectively be expected to slow down in the years to come. A September 2009 report by OECD, UNCTAD and WTO prepared before G-20 leaders meet in Pittsburgh states that as most leading economies have invoked "trade defence mechanisms" to weather the downturn, and the growing unemployment due to the crisis will also continue to fuel protectionist pressures for the years to come.[29] Analysts are also concerned about longer term developments, fearing that the West-to-East repositioning of the world trade gravity centre may spur a series of "buy domestic" and "produce where you sell" pressures, leading to renewed trade and investment confrontations and an increase in protectionism.

This outcome would have dire consequences on global value chains, and the related international production networks. [30] Indeed, supply chains are very sensitive to even small increase in transaction costs, be they caused by higher tariffs or oil prices. A series of not-so-anecdotic evidences tend to support this hypothesis. In August 2009, the head of Ernst & Young's supply chains department declared that regulatory changes and also the downturn are forcing many organizations to consider restructuring their supply chains, leading to smaller and more regional supply chains (Financial Times 9 August 2009).

This deglobalization is not only linked to the present crisis situation, but may be more structurally caused by the difficulties of decentralizing increasingly complex industrial procedures. For example, after an

[29] "Report on G20: Trade and Investment Measures", 14 September 2009.

[30] Global value chains include the conception and marketing aspects of the products, and are sensitive to risk of breaches in intellectual property and patents; their production component (the supply chains) are also highly sensitive to international and cross-border transaction costs as the goods for processing typically cross several border during production.

accumulation of delays, and confronted with a series of difficulties in the production of its latest model, Boeing decided to abandon the original fragmented chain and repatriate key production processes in its main establishments. Differentiated regional markets, as well as political pressure to voluntarily restrict exports and "build where they sell" have also encourage automobile constructers to relocate their final assembly closer to their final markets, keeping only the heavy engineering work truly global (Sturgeon, and van Biesebroeck 2009).

Other structural factors are also at work, which may increase transaction costs and push global firms to reconsider their production networks. Since 2003, oil price has been increasing, reflecting, *inter alia*, the additional demand coming from large emerging countries. This trend is probably installed for some time now, and production mangers will have to adapt to a future where energy is more expensive and less plentifully available. The same structural factors that led to an increase in the consumption of fossil fuels are also causing a change in the perception of businesses, consumers, and policy makers. Indeed, carbon footprints (accumulated CO_2 emissions) calculated through the life-cycle of a product has become increasingly associated with transportation of goods within the supply chain. Using input-output matrices, Hertwich and Glenp (2009) calculates that trade margins "representing the accumulated emissions from distribution between the producer and final consume" account for 5.5% of greenhouse gases emissions. As concerns about global warming increases, purchasing managers are concerned about the carbon footprint of their supply chains, while consumers are increasingly offered carbon-labelled products and economists talk about internalizing environment costs through ad hoc consumption taxes. As these trends accentuate, there will be a natural tendency to shorten global supply rely more on regional or domestic networks.

If this trend is confirmed, these underlying deglobalization forces would hinder the medium-term possibilities of recovery for international trade at its pre-crisis level. Because the most dynamic markets are in emerging countries, a regionalization or a repatriation of global supply chain would negatively affect developed countries' exports of intermediate goods, slowing down their recovery and augmenting the risk for an L shaped exit pattern. But because global supply chains are a source of efficiency gains and technical progress diffusion (*see* Box 1 again), even emerging countries would see their potential growth reduced in the process.

More importantly, it may also deprive the poorest developing countries, located far from established or emerging markets, of the opportunities of following the industrialization path taken by China or Mexico, a powerful strategy for frog-leaping through the Rostovian take-off model by attracting foreign direct investment, creating large volume of manufacturing jobs and transferring technologies (Box 2). Thus the micro-economic debate on the future of global supply chains spills over very critical trade and development issues.

Yet, against this pessimistic outcome, many considerations militate in favour of productive network continuing to extend their global reach. In the short run, abandoning the present global network of suppliers carries a heavy cost for the multinational firms. Off-shoring has been a central objective of many key industries, which heavily invested in their international network. Often, the new plants build off-shore are more modern and efficient than the older domestic ones, and selling them to a competitor would create a comparative disadvantage (remember the dilemma of GM with the sale of Opel).

In the longer run, the constant flow of innovations and the extension of the technological frontier are lowering the cost of communication, creating new opportunities for redesigning the international division of labour.

Box 2. Global Supply Chains, Industrialization and Development

There is a heated debate, reminiscent of the 1970s controversies about the role of trans-national corporations, between the liberal and the heterodox schools on the role of global supply chains in fostering industrialization or, at the contrary, causing "inmisering growth", low labour standards and more informality.

The establishment of export processing zones (EPZs) in developing countries and their success in attracting foreign direct investment have usually been balanced by a bias towards a low qualification and low salary profile for their employees. Because the establishments in EPzs in least advanced countries are predominantly labour intensive and foot-loose industries (e.g., apparel and garments), critics to an industrial strategy based on processing industries point to **1.** their lack of backward and forward linkages with other domestic industries; **2.** their reliance on informal markets to lower labour costs; and **3.** their sensitivity to conjunctural downturns. Partial field data from Asian countries on the impact of the crisis show declines in both average working hours and average earnings, especially in footwear and furniture industries. In addition, it seems that the crisis hurts more the formal sector, while the informal sector saw increases in average working hours (Hurst, Buttle, and Sandars 2009).

The controversy is partially based on an erroneous supposition that inward oriented (domestically integrated) and outward oriented (globally connected to supply chains) industrialization processes are mutually exclusive forms of industrialization for developing countries. It needs not being

so because capital and skill requirements are quite different, especially at the earliest stages of the export-processing industrialization. Milberg (2007) provides a survey of the pros and cons of industrial up-grading through export processing, and identifies potential spill-over. Using the tools of structural economics, Escaith (2007) shows that outward-oriented Asian economies, which inserted themselves actively in global supply chains, were able to successfully absorb their growing active population, while the more domestic based manufactures in Latin American countries could not provide enough formal jobs in the manufacturing sectors to the new entrants, pushing them instead into a growing urban informal service sector. Bacchetta, Bustamente, and Ernst (2009, 111) indicates also that EPZs actually compete more with informal activities than established national industries, "offering better, more stable employment opportunities for those previously working in the informal economy".

This logic of transferring more and more labour-intensive manufacture processing in developing countries applies easily when low-income economies export to high-income economies, because of the strong complementarities in comparative advantages. However, such complementarities are eroded when low-income economies trade with other low-income economies, because comparative advantages, cost structure and industrial specialization are rather similar. With the gravity centre of international trade moving from Developed West to Emerging East, it should be feared that, as long as these variable factors remain similar, South-South trade be driven more by absolute rather than comparative advantages, i.e., by differences in natural resource endowments rather than by efficiency opportunities.

- **Supply Chain in Global Rebalancing**

During the 1990s, large trade imbalances developed in several regions of the world; with the US running persistent deficits while Japan, Germany and later China, running surpluses. Many trust that these imbalances, financed by an increase in US liabilities, created a persistent situation of financial distortion that led to the September 2008 crisis. Rebalancing is therefore a key objective in the exit strategy, while many fears that such a rebalancing might lead to a secondary demand shock, fragilizing further the international financial system as such correction could only be possible thanks to a large correction in the bilateral exchange rates of concerned countries. From a macroeconomic perspective, the rebalancing can be analysed from several angles; the present section will focus on two particular aspects: balance of payments and global effective demand.

- **Balance of Payments**

The role of supply chains in amplifying trade flows should prove some kind of blessing when it comes to redress the "global imbalances", particularly the large trade deficit of the US economy. A back-of-the envelope calculation shows that the bilateral deficit of the USA vis-à-vis China measured with conventional trade statistics over-estimates the imbalances measured in value added content by about 60% (Table 2). This estimate, derived from an indirect measurement of value added content estimated from international input-output matrices, is based on the hypothesis of homogeneous production on the US side (i.e., the US production of final goods is similar for exports and for domestic use) and heterogeneity for China (technological dichotomy between firms producing for the domestic market and firms producing for exports, with higher import contents for the latter).

Table 2. Bilateral trade balance China-USA, Gross vs. Value Added measurement.

Billion USD	2000	2008
1. US exports to China	19	77
2. Chinese exports to USA	80	305
Balance (1-2)	(61)	(228)
3. US-VA exports to China	18	70
4. Chinese VA exports to USA	40	152
Balance (3-4)	(22)	(83)
5. Ratio (3-4)/(1-2)	0.4	0.4

Note: Asymmetric trade flows were averaged, and may differ from balance of payments values. Value added content for US exports is about 0.9, based on input-output coefficients; the respective value for China was 0.8, and adjusted to 0.5, accounting for re-exports and Export Processing Zones.

Source: Author's estimate based on COMTRADE and IDE-Jetro data

Because the domestic value added content of trade is lower than the gross commercial value recorded in

the balance of payments, closing the gap between China and the USA will be faster and, more importantly, cheaper in terms of lost welfare. This said, measuring trade in value added may, as in this case, reduce bilateral imbalances, but in other instances it will increase it, or even change its sign (Daudin *et al.* 2009; Johnson, and Noguera 2009). From a balance of payments perspective, the overall imbalance of an economy *vis-à-vis* the rest of the world will remain the same.

Measuring bilateral trade flows according to their domestic content modifies also the responses to bilateral exchange rates adjustments. In the previous example (Table 2), an appreciation of the Chinese Yuan *vis-à-vis* the US dollar will only be transmitted to the price of Chinese exports in proportion to their domestic content (50% in average). Thus, a revaluation of the Yuan would influence only moderately the bilateral balance with the USA, albeit it still may lead to a lower overall trade surplus for China, especially by increasing its total demand for imports.

- **Global Effective Demand**

Even if some bilateral imbalances may prove easier to resolve, it remains that the rebalancing will imply for the deficit countries higher national savings, to be obtained through a relative decrease in final demand and an improved competitiveness to generate net exports. These forces will lead to fundamental changes in the source of global effective demand, from "old industrialized West" to "emerging East"[31]. The implications for the industrialization prospects of other developing countries may depend more strongly on their geographical situation (distance from the emerging markets) and their existing production potential. As mentioned in Box 2, for developing countries that will remain outside regional production networks, it may be feared that rebalancing from North-South to South-South trade will be determined largely by absolute comparative advantages. Such an exit scenario implies, at least for the medium term, a relative reduction in the international demand for labour intensive manufactured consumer goods (clothing, textile and consumer electronics) and an increase in demand for commodities (agricultural, fuels and minerals) and investment goods.

Adjusting to this trend will prove especially difficult for the less-advantaged countries (i.e. resource-constrained small developing countries, especially the least-developed countries). According to WTO (2010) developed economies remain largely the dominant LDCs' export destination for manufactured articles such as clothing (95%) and some high value-added agricultural and food products (between 60 and 70%). For these developing countries suffering usually from high trade deficit, the external situation may worsen as the potential for exporting labour intensive products to the North would decline in the same time as their import bill in oil and food will increase.

Even for resource-rich less-advantaged countries, which have been benefiting from the increase in commodity prices during the 2000s, the long term impact on development is not clear, as exporting non-renewable commodities does not have the same social benefits than exporting labour-intensive merchandises and is non-sustainable in the long term. For these countries, the challenge is to manage their natural resource bonanza in a sustainable way, but it is not an easy course of action, considering the pressing needs of their population.

Confronted to such a perspective, there are basically two trade-related options opened to less-advantaged countries. One is to lower the costs of exporting. These costs are part of the supply constraints that frequently reduce the international competitiveness of LDCs and limit their trade potential. Based on World Bank estimates (Doing Business, 2009), time for exporting a standardized cargo of goods by ocean transport was 37 days for LDCs, against a world average of 25 (less than 12 days in the case of developed countries). According to Djankov *et al.* (2006), a one-day reduction in delays before a cargo sails to its export destination is equivalent to reducing the distance to trading partners by about 70 km. The issue is even made more pressing by the "distance puzzle" or "missing globalization puzzle" (the negative effect of distance on international trade is not shrinking as expected, despite technological changes) for poor countries. Carrère *et al.* (2010) estimates that low income countries exhibit a significant rising distance effect on their trade, around 18 percent between 1970 and 2006. While part of the effect could be attributed to the impact of regional trade agreements, the growing importance of vertical specialization and tightly managed global supply chains under just-in-time production increase the opportunity cost of time and distance. In this case, low–income countries that face relatively larger costs could be marginalized from global trends.

[31] "Emerging East" is an illustrative concept which includes some western hemisphere countries, like Brazil or Mexico. The concept of effective demand refers to a demand-driven model, where production responds to final demand and where some large importing countries (typically the USA and the EU in the pre-crisis scenario) play the role of world locomotive.

Because distance and transaction costs is becoming a critical factor in shaping international supply chains, all efforts should be done to decrease the economic distance with the main destination markets in developed and emerging countries. Efforts to reduce delays and domestic transaction costs in order to facilitate trade in developing countries focus principally on upgrading infrastructure and reforming administrative procedures. For the poorer countries, these efforts need to be supported by international assistance, in particular aid for trade.

The other direction for trade-policy is to secure preferential market access in emerging markets for the labour-intensive exports of less-advanced countries. Traditionally, LDCs have benefited from non-reciprocal preferences for their merchandise exports in developed country markets. More recently, a number of developing countries have granted preferences to LDCs, under a series of multilateral, bilateral and regional preferential market-access schemes. The (weighted) average tariff faced by LDCs when exporting to developing countries, nearly 12 per cent in 2006, is much higher than in developed countries (ranging from 3% for agriculture to 6.5% for clothing). Albeit 73% of the total value of LDC exports was granted duty-free status by other developing countries (90% in developed countries), it resulted principally from the favourable treatment of their exports of fuel and non-fuel minerals: only 30 per cent of their agricultural exports was accepted free of duty (93% in the case of developed countries), while the remaining agricultural exports faced an average tariff of 26 per cent (WTO 2010). This situation illustrates the wide dispersion of tariffs facing South-South trade, and the need for improving LDCs' market access in developing countries. Moreover, while the conclusion of the Doha Round (DDA) would certainly benefit LDCs trade with developed countries by removing some of the trade-distorting agricultural subsidies and consolidating the unilateral preferential treatment, it will also lead to preference erosion. The margin of preference granted to LDCs by developed countries will be reduced because their applied Most Favoured Nation tariffs, in particular tariff peaks, will be lowered under the DDA. At the contrary, there should remain an important potential for fostering preferential treatment for LDCs with most income developing countries, because the DDA cuts will principally affect their bound rates (the maximum tariff authorized under the WTO agreement) rather than the applied duties.

4. Concluding remarks

The geographical segmentation of industrial production has played a major role in shaping international economy in the past 15 years. It was at the root of the emergence of new global players, such as China or Malaysia, and the correlated dramatic reduction in absolute poverty levels. It allowed also some old industrial economies like Germany or the USA, to regain international competitiveness through increased productivity and efficiency. In the same time, the content of merchandise statistics and the economic significance of trade balances became more and more difficult to interpret.

Supply chains reshaped international trade and changed the relationship between trade and development: the surge of trade in intermediate commodities remodelled regional and international networks through a bottom-up angle and forced governments to reconsider the previous identification of industrial development with protectionism; developing countries were able to leap-frog the traditional industrialization phases by inserting themselves into complex industrial networks; the criss-crossing of manufacture networks led to large investments in transportation equipment and infrastructure in order to accommodate the huge transit of goods for processing that removed bottlenecks and favoured economic development.

Offshoring altered also the social panorama. If global economic and welfare benefits have been substantial, their distribution remains contentious. While creating numerous jobs in emerging countries, outsourcing and offshoring increased wage disparity in both developed and developing economies, fuelling an active political debate on the pro and cons of globalization.

Even if global supply chains did not alter dramatically the long run Trade-to-GDP elasticity, their role in explaining the trade collapse that followed the financial crisis of September 2008 has been determinant. And determinant is also their role in shaping the alternative exit scenarios from the Great Recession, as is their future contribution in any post-crisis scenario. The reactivation of the global supply chains is still largely dependent on the rebound of final demand in developed countries, i.e household consumption and firm's investments, which have been reduced because of higher unemployment and increased risk aversion. The road to recovery can be a slow and bumpy one for the developed economies, many of them suffering from internal or external macroeconomic deficits. It is therefore expected that the gravity centre of demand for imports will continue shifting from West to East.

The reshaping of global effective demand in any future scenario is of particular importance for the labour abundant developing countries that where relying on the strength of the global supply chain movement to attract

productive investments. Global Supply Chains are based on comparative advantages derived from costs and specialization. Complementarities based on cost differential between countries arise naturally when factors endowments are very different. This explains the specialization in trade in tasks between industrialised economies and labour abundant developing countries. Complementarities based on specialization are, at the contrary, the domain of intra-sectoral trade between complex industries, typical of the developed economies (for example, intra-EU trade in manufacture). The capacity of LDCs and least-advanced developing countries to successfully insert themselves into supply chains led by and for emerging economies is still unclear. The difference in factor endowments and production costs is not large (as in the case of mature industrialised countries), and most emerging countries still count with a large reservoir of labour.

An increase in objective and subjective transaction costs, from higher oil prices to "buy local" campaigns and murky protectionism, indicates that in the future, supply chains will probably be smaller and more regional. Let unchecked, these centripetal forces may well cause a deglobalization process which will directly affect less advanced developing economies but, in the end, will also be detrimental to both developed and emerging countries. For less advanced countries, closing the gap with the closer-knitted regional value chains means reducing transaction costs. Resource-constraint less advanced countries need more aid for trade to engage vigorously in trade facilitation programs to reduce their transaction costs, while they rely on emerging countries to provide more preferential treatments to their non-traditional exports.

On the other hand, the deglobalization outcome as a result of the present crisis is far from being certain. The technical factors that made possible the internationalization of production, from the IT revolution to innovations in engineering and business management, still promote further "flattening of the Earth". Large emerging countries are becoming new markets for final goods that reshape existing production networks, while new actors are emerging from the "not-so-emerging" countries. Thus deglobalization is probably a distant menace on objective grounds, even if it is a new global effective demand that will drive the world economy, forcing some difficult adjustments. Indeed, the 2008-2009 crises are a structural break, and the world economy will certainly not return to "business as usual". Old giants' tumbled, new global players emerged. Public opinion is also changing and the citizens' concerns on the lack of governance of the previous phase of globalization will have to be addressed, while the present gains in opening trade opportunities will have to be preserved.

References

[1] Baldwin, R. 2006. *Globalisation: the great unbundling(s)*. Economic Council of Finland, 20 September, 2006.

[2] Bénassy-Quéré, A., Decreux, Y., Fontagné, L., and Khoudour-Castéras, D. 2009. Economic Crisis and Global Supply Chains. *CEPII* No 2009-15, July 2009.

[3] Carrère, C., de Melo, J., and Wilson J. 2010. Distance and Regionalization Of Trade For Low-Income Countries. *World Bank WPS*5214, February 2010.

[4] Daudin, G., Monperrus-Veroni, P., Rifflart C., and Schweisguth, D. 2006. Le commerce extérieur en valeur ajoutée. *Revue de l'OFCE* 98: 129-165.

[5] Djankov, S., Freund, C., and Pham C.S. 2006. Trading on Time. *World Bank Policy Research Working Paper* No. 3909.

[6] Dvorak, P. 2009. Clarity is missing in Supply Chain. *The Wall Street Journal*, 18 May 2009.

[7] Escaith, H. 2007. Old and New Dualisms in Latin America and Asia: Labour Productivity, International Competitiveness and Income Distribution. *Revista Trabajo* No.5, ILO-UAM.

[8] Escaith, H. 2008. Measuring trade in value added in the new industrial economy: statistical implications. 12th Meeting of the Association française de Comptabilité nationale and *MPRA Paper* 14454, Munich.

[9] Escaith, H., and Gonguet, F. 2009. International Trade and Real Transmission Channels of Financial Shocks in Globalized Production Networks. *WTO Working Paper* ERSD-2009-06, May 2009.

[10] Escaith, H., Lindenberg, N., and Miroudot, S. 2010. International Supply Chains and Trade Elasticity in Times of Global Crisis. *WTO Working Paper* ERSD-2010-08, February 2010.

[11] Feenstra, R., and Hamilton, G. 2006. *Emergent Economies, Divergent Paths: Economic Organization and International Trade in South Korea and Taiwan*. Cambridge University Press.

[12] Grossman, G., and Rossi-Hansberg, E. 2006. The rise of offshoring: it's not wine for cloth anymore,

Proceedings, Federal Reserve Bank of Kansas City.

[13] Hertwich, E., and Peters, G. 2009. Carbon Footprint of Nations: A Global, Trade-Linked Analysis. *Environmental Sciences and Technology* Vol. 43 No 16: 6414‐6420.

[14] Hurst, R., Buttle, M., and Sandars, J. 2009. The Impact of the Global Economic Slowdown on Value Chain Labour Markets in Asia. (September 2009).

[15] IDE-JETRO. 2006. *How to make Asian Input-Output Tables*. March 2006.

[16] Johnson, R.C., and Noguera, G. 2009. *Accounting for Intermediates: Production Sharing and Trade in Value Added*. Dartmouth College and University of California, Berkeley.

[17] Krugman, P. 1995. Growing World Trade: Causes and Consequences. *Brookings Paper on Economic Activity* 1:327-377 – 1995.

[18] Lynn, B. 2009. How Detroit Went Bottom-Up. *The American Prospect*, October 2009.

[19] Milberg, W. 2007. *Export Processing Zones, Industrial Upgrading and Economic Development: A Survey*. New School for Social Research.

[20] Miroudot, S., and Ragoussis, A. 2009. Vertical Trade, Trade Costs And FDI. *OECD Trade Policy Working Paper* No. 89.

[21] Oosterhaven, J. 1996. Leontief versus Ghoshian Price and Quantity Models. *Southern Economic Journal* Vol. 62, No. 3: 750-759.

[22] Stadtler, H. 2008. Supply Chain Management – An overview, In Hartmut, S., and Kigler, C. *Supply Chain Management and Advanced Planning.* Springler-Verlag, 4th Edition.

[23] Sturgeon, T., and Van Biesebroeck, J. 2009. Crisis and Protection in the Automotive Industry: A Global Value Chain Perspective. *World Bank Policy Research Working Paper* 5060.

[24] Tanaka, K. 2009. Trade collapse and international supply chains: Evidence from Japan. *VoxEU.org*, 7 May 2009.

[25] Uchida, Y., and Inomata, S. 2010. Vertical Specialization in the Time of the Economic Crisis. IDE-Jetro 2010.

[26] Yi, Kei-Mu. 2009. The collapse of global trade: The role of vertical specialisation, In R. Baldwin, and S. Evenett (eds), The collapse of global trade, murky protectionism, and the crisis: Recommendations for the G20, *VoxEU publication* 2009.

[27] WTO. 2010. Market Access for Products and Services of Export Interest to Least-Developed Countries. *WT/COMTD/LDC/W/46 Revue*.1 , February 2010.

SOME EMPIRICAL EVIDENCE OF THE EURO AREA MONETARY POLICY

Antonio FORTE
University of Bari, **Italy**
Department of Economics and Mathematics
forteantonio80@gmail.com

Abstract

In this paper I try to find some empirical evidence of the European Central Bank's behaviour from its outset, January 1999, to the mid 2007, using a Taylor-type rule. I test a new and simple method for estimating the output gap in order to avoid problems linked with the estimate of the potential output. Moreover, I analyse the significance of some explanatory variables in order to understand what the basis of the E.C.B. monetary policy decisions are. Finally, I find an important evidence of the role of the Euro-Dollar nominal exchange rate in the conduct of the Euro Area monetary policy.

Keywords: Taylor Rule, European Central Bank, Euro-Dollar exchange rate

JEL Classification: E43, E52, E58

1. Introduction

The Taylor Rule, Taylor (1993), has been widely used in the literature in order to reproduce the monetary policy of Central Banks. This rule has been used in different ways in the literature: with different regressors, in levels or in first difference, with lagged or contemporaneous data, with expectations, and the estimations have been carried out with different techniques. This paper uses the basic version of the formula and OLS and NLS estimations in order to study the significance of many regressors, some of that completely new in literature. I introduce a new way to calculate the output gap. This new methodology seems to be, in this first attempt, quite convincing. Moreover, I tried to find what types of independent variables one should use to improve the fitting of the data and I also present a section focused on the role of inflation expectations. The rest of the paper is divided into eight sections: firstly, I introduce some works on the Taylor rule; then I explain the way in which I calculate the output gap; in the third section I present the initial findings; in the fourth paragraph I focus my attention on monetary variables; the fifth section analyses the importance of the Euro-Dollar exchange rate; the following two sections deals with the role of the inflation expectations and the interest rate smoothing; the last section concludes.

2. Taylor rules

The dispute about the best type of Taylor Rule is still unsolved. Both the backward looking and the forward looking approaches provide good results. During the last years, given the growing importance of the expectations in the theoretical framework of the economic science, it seems that the forward looking rule acquired the supremacy. Notwithstanding this, there are so many works on this issue that it is really impossible to determine what the right formula is.

At the same time, it is likely that, citing this literature, one disregards a relevant part of the works in consequence of their considerable number. Indeed, the Taylor rule can be estimated in different ways, see Carare and Tchaidze (2005) for a short excursus. In my work, I follow both the backward and the forward looking Taylor Rule strands and try to find some empirical evidence on the European Central Bank's (ECB) behaviour from the onset of its operations, January 1999. As regards the forward looking approach, I don't strictly use the expectations, for example using GMM or other types of methodologies applied in the literature, but I simply use time series for the inflation expectations obtained through consumers and experts surveys.

Given the just recalled amplitude of the literature on the Taylor rule, in this paragraph I only cite some works on the issue that I face in this paper.

In 1996, Clarida and Gertler published a work in which they study the monetary policy of the Bundesbank through a Taylor rule. This study has been one of the first for the European countries and the rule showed its usefulness also in this case. Two years later, in 1998, Clarida, Galì and Gertler presented a paper on the monetary policy in the US, Japan, Germany, Italy, France and the UK. They used GMM estimations and their work started to be a fundamental reference point in this field.

Gerlach and Schnabel (1999) analysed the EMU area. In their study they highlighted the role of the inflation expectations in the Taylor rule. They also studied the significance of other regressors, such as the Euro-Us Dollar exchange rate, but in their analysis these regressors showed a very limited role.

A work by Fourçans and Vranceanu (2002) is really interesting. They estimated many policy rules, analysing the first four years of the ECB activity. They demonstrated, using OLS and GMM, that the ECB is a conservative central bank and that the ECB is also focused on the real economy. Moreover, in line with the result of my paper, they also showed that the monetary policy disregarded the fluctuations of the M3.

Ullrich (2003) compared the Fed and the ECB from 1995 to 2002. She found that the ECB adopted a easier monetary policy in comparison with the former national central banks. This result is supported by the value of the coefficient on inflation that was above unity before the monetary union, while it decreased below unity with the ECB monetary policy.

Another work that employs the Taylor rule is the one by Sauer and Sturm (2003). They studied the first years of the Euro area using OLS and NLS. Their results leave no doubts: with the use of expectations, the coefficient on inflation is above unity. While, on the contrary, using contemporaneous data, the ECB seems to accommodate the course of the inflation rate.

In 2006 Carstensen published a work in which he tried to analyse whether the revision of the Euro area monetary policy strategy in 2002 had a real impact on the ECB behaviour. He found that this revision did not have a concrete effect on the value of the coefficients. Another important finding of this study is the very marginal role of the monetary aggregates.

Rotondi and Vaciago (2007) used a Taylor-type rule to compare the ECB with the Bundesbank. They showed that the coefficient on the inflation gap using the backward looking version of the formula is lower than the one obtained with the forward looking approach.

A paper by Parsley and Popper (2009), focused on Korea, uses GMM to estimate a policy reaction function similar to a Taylor rule, from January 1999 to April 2008. They used the exchange rate in the formula and they often find a significant relationship between this regressor and the interest rate. The exchange rate is not often used in the monetary policy rules, but some of the works I cited use the exchange rate as regressor and, recently, even Engle (2009) demonstrated the usefulness of introducing the exchange rate in an open-economy two-country model.

Starting from this literature, I present an analysis focused on the Euro area monetary policy. As I have previously said, I followed both the backward and the forward looking strands and I tried to find some empirical evidence on the ECB's behaviour from the onset of its operations, January 1999, to August 2007.

3. Output gap: a simple estimate

It is possible to rewrite the original Taylor Rule in the following way:

$$i_t = \beta_0 + \beta_1(\pi_t - \pi^*) + \beta_2 x_t \qquad (1)$$

Where "*i*" is the nominal interest rate, the monetary policy tool; the coefficient β_0 is the long run nominal interest rate, the rate that emerges when both output and inflation hit their target; β_1 is the coefficient on inflation gap, it expresses what the change in the nominal interest rate is if the inflation rate deviates from its target; π_t is the time t inflation rate and π^* is the inflation rate objective: their difference is the inflation gap; β_2 is the output gap coefficient and it expresses how a Central Bank reacts to change in the economic growth speed, and the last coefficient, x_t, represents the output gap, that is, the difference between the actual output and the potential output that an economy can reach.

In more detail, the output gap (x) is defined as the difference between the actual output and the potential output. The potential, or natural, output level is the output level that an economy can get if perfect flexibility of prices and wages exists. This is the most generic and theoretical definition of output gap (*see* for example Clarida, Galì, and Gertler 1999). But it is very difficult to calculate it. First, as it is straightforward to understand, it is impossible to know the value of the true natural output level. It is necessary to use an estimate for this value. And so, before carrying out studies about a Taylor Rule it is necessary to derive the level of the potential output in order to obtain the value of the output gap. The natural level of the output is often estimated applying a filter on the output time series.

For this reason, the estimate gained with the Taylor Rule can be influenced by the initial estimate of the output gap. The use of different time series to obtain the output gap, or a different method for calculating the natural level of output, can affect the final results. It could be possible that an economist starts a work with an *original error*.

Moreover, it is common to substitute the gross domestic product with the industrial production as a proxy of the output. And so, many economists estimate a natural level of the industrial production and an industrial production gap through the use of the Hodrick Prescott filter on this type of time series and they use this type of gap in the Taylor Rule.

In this paper, in order to eliminate this initial problems, I have used a different and simpler way to estimate the output gap: I do not substitute the gross domestic product time series with other time series and I do not make use of any type of filter.

Analyzing the attitude of the ECB towards the economic framework, I have supposed that a quarterly GDP growth between 0.5 and 0.6% could be considered as an inflation-neutral speed growth. We can see at this range as a natural and non-inflationary economic growth. I suppose that the ECB moves the nominal interest rate if the quarterly GDP growth is above or below this specific range only. It is clear that this type of method is really simple: for example, a quarterly GDP growth above 0.6% could represent a threat for price stability and so the ECB is induced to move nominal interest rate adopting a tighter monetary policy. Using this simple method it is possible not to use preliminary estimates (i.e. estimates on GDP trend or other type of calculation in order to obtain an output gap) before testing the Taylor Rule. Furthermore, this method has given very good results, as one can see in the next sections. In next sections I refer to this indicator as the GDP gap.

4. Initial findings

As I have said before, in this first section I use a backward looking Taylor Rule and the estimates are obtained with OLS regressions. I use the day-to-day rate as dependent variable. The sample examined in the paper starts in January 1999 and ends in August 2007. I decided to disregard the first period of the crisis. I got data on GDP, inflation and Euro-Dollar exchange rate from Eurostat's website and those on monetary aggregate M3 from the ECB's website.

Inflation data are monthly. I have transformed the GDP data from quarterly to monthly with a simple method. I have applied the GDP quarterly growth to each month of that quarterly. The quarterly GDP growth indicates the increase of the GDP over a period of three months in comparison with the previous three months. It means that, on average, the monthly amount of GDP is increased with the same pace. For example, if we recognize a 0,5% quarterly GDP growth in the second quarter of a year, we can assume that in each month of that quarter (April, May, June) there has respectively been a 0,5% GDP growth in comparison with the first, second and third month of the previous quarter (January, February, March). Through this simple approach I have transformed my GDP gap quarterly data in GDP gap monthly data. Then, if we consider the gap between actual value of GDP growth (gdp%) and the natural range of growth (0,5%÷0,6%= gdp%*, see the previous section) we can obtain the GDP gap for a single month. In the rest of the paper I will call this gap GDP gap (= gdp% - gdp%*) and I will use this value in the estimates of the Taylor Rule.

As regards the lags, the ECB Council decides the path of the MRO rate during the meetings held the first Thursday of each month. So, it is really impossible that the ECB has contemporaneous values of output gap and inflation gap at disposal (this is the original prevision made by Taylor, 1993). For this reason I decided to use two lags for the inflation rate and 4 lags for the GDP gap. Indeed, it seems plausible that at the beginning of a month the ECB *knows* the value of the inflation gap of two month before. For the GDP I have considered the great difficulty to calculate this indicator and the delay of its release.

The resulting formula is:

$$i_t = \beta_0 + \beta_1(\pi_{t-2} - \pi^*) + \beta_2 x_{t-4} \qquad (2)$$

The results of the first estimation are shown in Table 1.

Coefficients on inflation gap and on GDP gap are significant, but the values of these coefficients and the value of the R^2 statistic indicate that this initial estimate does not have a great economic significance. Coefficient on inflation gap (inflgap_2) has a value slightly below 0,60. This is a little value compared to the major empirical work or to the optimal value suggested by the dominant economic theory (*see*, for example, Clarida, Gertler, and Galì 1998), even if this value, according to the work of Ball (1999), is fully acceptable. This value implies that the

ECB is not very reactive towards inflation fluctuations. Coefficient on GDP gap is quite big. This value, above the unity, means that the ECB is really careful about deviations of GDP around its trend.

Table 1. Taylor Rule, basic version

OLS, 100 obs., sample 1999:05-2007:08, robust standard errors, dependent variable: day to day rate					
Explanatory Variable	Coefficient	Std. Err.	t-test	p-value	
Const	3,00640	0,153089	19,6382	<0,00001	***
inflgap_2	0,603602	0,307136	1,9653	0,05224	*
gdpgap_4	1,07067	0,373052	2,8700	0,00504	***
Mean of dependent variable = 3,0453					
Standard deviation of dependent variable = 0,91656					
R.S.S. = 72,0106					
Standard Error of residuals = 0,861613					
R^2 = 0,134156					
Adj. R^2 = 0,116304					
F-test statistic (2, 97) = 4,39156 (p-value = 0,0149)					

These results could be caused by the long sample that I have considered and by the different events occurred during this last seven years. We can remember, for example, the end of the positive economic cycle due to the internet bubble in the first part of 2001, or we can consider the global economic shock due to terrorist attacks of 11th September 2001, or we can just mention the change in the conduct of the ECB's monetary policy in 2002-2003. All these phenomena, together with many others not mentioned here, have probably induced some changes in the managing of the Euro Area monetary policy. All these facts reduce the model suitability in reproducing the actual data and, as a consequence, it is not possible to represent in the right way the ECB's monetary policy. So, we can assert that this basic Taylor Rule seems not to work well over this long sample.

It can be useful to deepen this work in search of a more accurate rule. In the next pages I will add other explanatory variables in the right hand side of my Taylor Rule in order to improve the correspondence of the estimated data to the actual values.

5. Does the first pillar have an active role?

The ECB's monetary policy is based, as the official documents say, on two pillars. On the one hand the first pillar is focused on Euro area monetary situation. On the other hand the second pillar consists of the study of the entire economic system situation.

There has been a great debate on the role and importance of the first pillar (see for example Gerlach, 2004, and Jaeger, 2003), mostly because it has been the ECB that officially and explicitly communicates, at the onset of its activity, a target value for the growth of the monetary aggregate M3. The ECB specified a precise objective for the M3 growth: a 4.5 percent annual growth for the moving three-month averages of the M3. This growth, according to the European Central Bankers, was neutral for the price level of the area. For these reasons I included M3 in my basic Taylor Rule formula to test how the ECB has reacted to fluctuations of this aggregate.

Table 2. Taylor Rule with M3

OLS, 100 obs., sample 1999:05-2007:08, robust standard errors, dependent variable: day to day rate					
Explanatory Variable	Coefficient	Std. Err.	t-test	p-value	
Const	3,21785	0,391119	8,2273	<0,00001	***
inflgap_2	0,607302	0,288817	2,1027	0,03811	**
gdpgap_4	1,01752	0,35411	2,8735	0,00500	***
M3gap_2	-0,0857523	0,137823	-0,6222	0,53529	
Mean of dependent variable = 3,0453					
Standard deviation of dependent variable = 0,91656					
R.S.S. = 70,549					
Standard Error of residuals = 0,857255					
R^2 = 0,15173					
Adj. R^2 = 0,125221					
F-test statistic (3, 96) = 3,56476 (p-value = 0,017)					

There is an ample debate on the role of this pillar given the fact that the 3-months moving average of M3, since the first months of 1999, has grown more than its established target. As a consequence, one could question if the monetary pillar really exists.

Table 2 shows results obtained with the M3 gap as an explanatory variable in the Taylor Rule. The formula estimated in this case is:

$$i_t = \beta_0 + \beta_1(\pi_{t-2} - \pi^*) + \beta_2 x_{t-4} + \beta_3(M3gap_{t-2}) \quad (3)$$

We can immediately note that coefficients on GDP gap and inflation gap have the right sign and are statistically different from zero. The reaction coefficient on M3 gap is not statistically significant and it is negative. If this value was right, we should have a Central Bank that performs an expansionary policy when M3 increases more than its optimal target (a similar result, i.e. a negative coefficient on the M3, is showed in Ullrich 2003).

The emerging of this type of relation between the day-to-day rate and M3 rises many doubts over the role of the monetary aggregate in the conduct of the monetary policy by the ECB. It seems that the ECB does not pay so much attention to the fluctuations of this monetary aggregate. This result is in line with the findings of the papers by Fourçans and Vranceanu (2002) and Carstensen (2006).

Moreover, the global results are quite disappointed. As we have noted for the basic model, it seems that a very simple rule, as the one used here, does not succeed to acceptably explain the course of the monetary policy over a long sample, even with the use of the M3 gap.

We could have two alternative problems: the Taylor rule is not useful in this context or the regressors I have chosen are not completely correct.

In sum, over this sample the Taylor Rule continues to provide useless information on the conduct of monetary policy in the Euro area. We can't use this formula as a benchmark to examine the monetary policy and to make some assumptions on the future path of the interest rate.

6. The weight of the Dollar

An ample political debate and many economic works have concerned the role of the exchange rates in the conduct of monetary policy, especially in the new framework of the European Monetary Union (*see,* for example, Carstensen 2006, Clarida, and Gertler 1996, Clarida, Galì, and Gertler 1998, Parsley, and Popper 2009, Engle 2009). It is well known that the fundamental task of the ECB is to keep inflation low and so the mandate of the ECB does not explicitly include a supervision of the exchange rate fluctuations. As a consequence, the ECB should not directly care about the fluctuations of the exchange rates, when it decides the course of the interest rates (see art. 105 of the Treaty). That is, the aim of the ECB is not to control the course of the exchange rate. However, the path of the exchange rate can obviously be part of the analysis of the second pillar of the ECB's monetary policy.

The just cited works have tried to find empirical evidences on the importance of the exchange rates in a Taylor-type rule. It is unquestionable that the exchange rates cover a crucial role in determining the speeding up and the slowing down of the economic cycle, mainly for the opened economies. But they also have an impact on the price level (through the pass-through effect). Even if these effects seem to be more mitigate than they were in the previous decades, we cannot disregard their existence and their possible impact on inflation.

Moreover, with the onset of the monetary union the role of the exchange rates for the nations inside the EMU has strongly changed. Before the adoption of the Euro every EMU nations should care about many bilateral exchange rates within the European boundaries. From January 1999 the attention has been mainly directed to the US Dollar – Euro bilateral exchange rate. Indeed, this exchange rate has now a central importance for the development of the European economic cycle.

Let us consider, for example, that all the raw material (not only the energetic raw material but also the unprocessed agricultural products) are quoted in US Dollars. The nominal value of these goods changes if the Euro-Dollar exchange rate is fluctuating and this could have an impact on the domestic inflation. Furthermore, the United States are one of the most important market for the European exports.

Besides, the importance of the Euro-Dollar exchange rate is increased by the strong linkage among Dollar and other national currencies all around the World. Indeed, there are many pegged exchange rates that have the US Dollar as an anchor (for example the Chinese yuan or the currencies of the OPEC). In this way an appreciation or a depreciation of the Euro towards the Dollar has a strong impact on the value of other import-export flows and on other import prices.

Given these facts, I decided to add the Euro-Dollar exchange rate in the basic formulation of the Taylor Rule. An important difference in comparison with the usual procedure is the direct use of the nominal exchange rate in the formula. I have chosen to use the nominal exchange rate without adopting any type of transformation. I add the time series of the nominal Euro-Dollar exchange rate in levels for three main reasons. First, the nominal exchange rate is an indicator easy and ready to use. Second, there is a psychological cause: the value of the nominal exchange rate has a great impact on the political behaviour, on the economic decisions and on the feelings of the people. Third, I think that the Taylor Rule should be an useful and simple instrument to inspect the course of the monetary policy and so I have avoided the use of transformations of the time series that should have made the rule too complex for unskilled people.

So, in my work I suppose that the Council of the ECB simply watches at this nominal exchange rate when it examines the entire economic framework of the Euro area.

I added the monthly average of the Euro-Dollar nominal exchange rate (symbol: €/$) with one lag:

$$i_t = \beta_0 + \beta_1(\pi_{t-2} - \pi^*) + \beta_2 x_{t-4} + \beta_3 €/\$_{t-1} \quad (4)$$

Table 3 shows the results obtained with this regression. I obtained a good result with the estimation: all the estimated coefficients have the right sign and they are different from zero at the usual levels of significance.

At a first sight even the graph of the actual and estimated values of the day-to-day rate (Figure 1) shows a certain degree of likeness between the two series. And this is obviously a very good goal if we consider the fact that we are testing the rule over our complete sample. But it is useful to carefully examine all the data.

First, it seems that the value of the constant is too high. This value (6.65) is higher than the values obtained in the previous regressions or in other comparable works.

But this value is strongly influenced by the coefficient on the bilateral Euro-Dollar nominal exchange rate. We must subtract the negative coefficient on the exchange rate to the value of the constant. As it is possible to understand watching the data in Table 3, the coefficient on Euro-Dollar nominal exchange rate is negative and it is multiplied by a value between 0.87 and 1.37 (respectively the minimum and the maximum value of the monthly average exchange rate in my sample). So it has a deep impact on the value of β_0. In this way the real value of the constant is reduced to a more normal value (around 3.3).

Table 3. Taylor Rule with exchange rate

OLS, 100 obs., sample 1999:05-2007:08, robust standard errors, dependent variable: day to day rate					
Explanatory Variable	Coefficient	Std. Err.	t-test	p-value	
Const	6,65496	0,970459	6,8575	<0,00001	***
inflgap_2	0,471199	0,218169	2,1598	0,03328	**
gdpgap_4	1,04028	0,239931	4,3357	0,0004	***
€/$_1	-3,29523	0,921761	-3,5749	0,00055	***
Mean of dependent variable = 3,0453					
Standard deviation of dependent variable = 0,91656					
R.S.S. = 44,6884					
Standard Error of residuals = 0,682279					
R^2 = 0,462674					
Adj. R^2 = 0,445883					
F-test statistic (3, 96) = 10,5077 (p-value < 0,00001)					

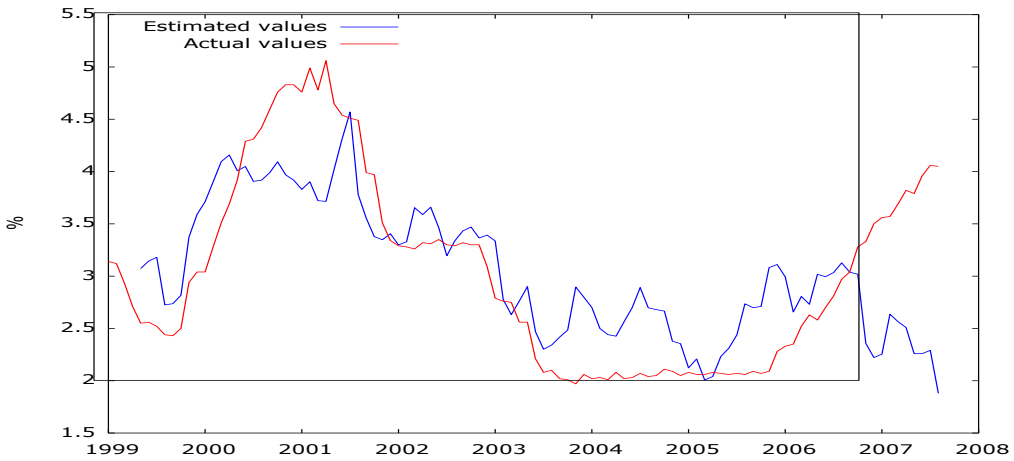


Figure 1. Day-to-day rate (actual and estimated values)

The coefficient on exchange rate has two functions. On the one hand, it impacts on the constant and so we have to jointly consider these values. On the other hand, it obviously indicates the marginal effect produced by a variation of the exchange rate on the nominal interest rate. We can infer an important information analysing this coefficient: an appreciation of about 30 cents of the exchange rate (that is, the Euro has strengthened against the dollar) induces the ECB to cut the interest rate of about 100 basis points.

Given this analysis, it is evident the importance of the nominal Euro-Dollar exchange rate in determining the decisions of the ECB's monetary policy.

This prominent attention towards this exchange rate can also explain the reduction of the coefficient on inflation gap that has a lower value than the ones estimated in the previous models.

Figure 1, as I have said before, shows the estimated and the actual values obtained with this last model. The two series diverge in a considerable way during the second half of 2006. Starting from that period, the model does not manage to well reproduce the path of the day-to-day interest rate.

Given this result, it seems that during the last period of the sample there has been a change in the managing of the Euro area monetary policy. If it was true, this fact should have caused interference on the estimated series.

So, I have thought to repeat the regression with the same model but over a shorter sample. I have eliminated the last period, in which the two series diverge, and so the sample now starts in January 1999 and ends in June 2006.

Table 4 shows the results of this last regression and Figure 2 illustrates the path of the estimated and actual values of the Euro area day to day interest rate.

Coefficients are always highly significant and they have the right sign, but the value of the coefficients has changed in comparison with the previous regression. The constant and the coefficient on the exchange rate have changed most of all. But if we remember the analysis suggested in the previous pages, we can only see a change in the marginal impact of the explanatory variables on the dependent variable. Indeed, the constant and the coefficient on the exchange rate have to be jointly evaluated. Even in this case, if we compute in the right way this linkage, we will find a constant around 2.5.

At a first sight, we could consider too small a value of 0.60 for the coefficient on inflation gap, but we should remember that in this model the ECB reacts in a very sharp way to fluctuations of the Euro-Dollar exchange rate. Following my results, the ECB should cut the interest rate by 1 per cent point if the Euro-Dollar nominal exchange rate grows by 20 basis points. If we consider these two effects at the same time, the impact of the coefficient on inflation gap and on exchange rate, and if we bear in mind the impact of this exchange rate on inflation, a value of 0.60 on inflation gap is acceptable. In addition a low value of the coefficient on the inflation gap is showed in other works too, *see* Ullrich (2003).

Table 4. Taylor Rule with exchange rate, short sample

OLS, 86 obs., sample 1999:05-2006:06, robust standard errors, dependent variable: day to day rate					
Explanatory Variable	Coefficient	Std. Err.	t-test	p-value	
Const	8,59879	0,639332	13,4496	<0,00001	***
inflgap_2	0,595707	0,116116	5,1303	<0,00001	***
gdpgap_4	0,83068	0,198511	4,1845	0,00007	***
€/$_1	-5,27647	0,569766	-9,2608	<0,00001	***
Mean of dependent variable = 2,96628					
Standard deviation of dependent variable = 0,953266					
R.S.S. = 13,1562					
Standard Error of residuals = 0,400551					
R^2 = 0,829673					
Adj. R^2 = 0,823442					
F-test statistic (3, 82) = 36,0183 (p-value < 0,00001)					

Furthermore, we can suppose that, as some economists have stressed (*see* Blinder 1997), in real world the central bankers are more cautious in comparison with what the empirical works suggest. And this happens because Central Banks do not have a perfect knowledge of the economic system. Prudence is fundamental in order to avoid an excessive or incorrect intervention that could undermine the stability of the financial system. This reasoning is especially worth for the ECB that has had to operate (and still operates) in a new economic framework. It is much more difficult for the ECB to forecast the reactions of the economic system to its monetary policy operations: this fact fully supports the choice of a high degree of prudence in opposing the fluctuations of the inflation.

The coefficient on output gap has the right sign and it has a plausible value.

It is also important to stress that the value of the R^2 statistic is the highest one among all the regressions of this paper. Indeed, both the R^2 statistic and figure 2 demonstrate that this last model mimics very well the course of the dependent variable.

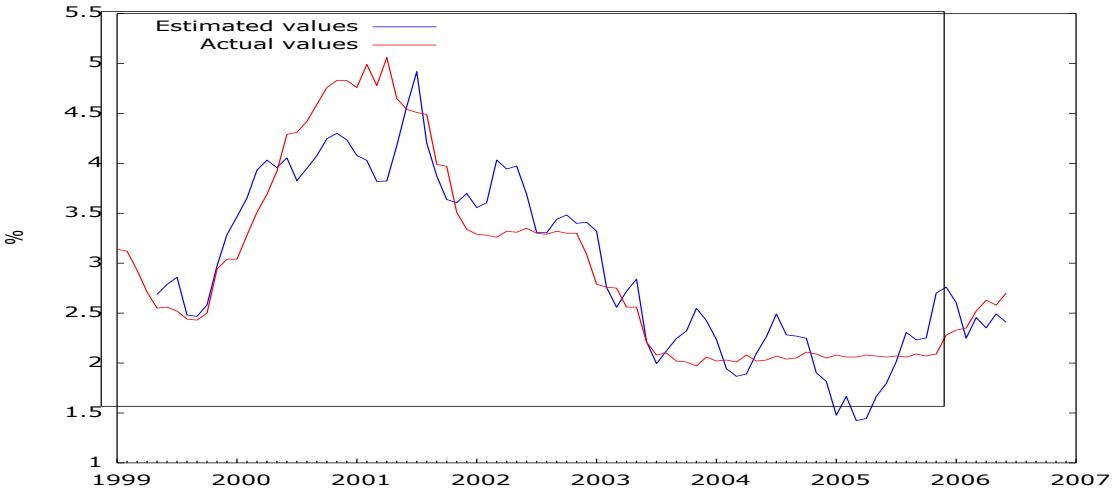


Figure 2. Day-to-day rate (actual and estimated values), short sample

This version of the Taylor Rule makes it possible to draw another important conclusion. I have previously explained the interaction between the value of the constant and the value of the coefficient on exchange rate. From this linkage a relevant result derives: the ECB can define a different value for the interest rate even in the presence of the same value of inflation gap and output gap. For example, in the basic structure of the Taylor Rule, without the Euro-Dollar nominal exchange rate, when output and inflation match their target values, central bank always sets the same interest rate. On the contrary, with the last model the value of the interest rate can be different even if inflation gap and output gap have the same value in different times. This happens because the value of the interest rate varies with the change of the exchange rate.

The example in Table 5 makes this result more evident. We suppose that inflation gap and output gap are zero (i.e. inflation and output are at the target level). In this situation the interest rate will not always be the same, but it will be higher if the Euro-Dollar exchange rate has a low value and it will be lower if the exchange rate is appreciated. In this way we can grasp the role of the exchange rate in dampening or heating the inflation rate even if output and inflation match their respective target. All these findings make it possible to say that this last model is the best, among the ones I tested till now, in reproducing the path of the day-to-day rate in the Euro area over a very long period of time.

Table 5. The role of the exchange rate in the Taylor Rule

$$i_t = \beta_0 + \beta_1(\pi_{t-2} - \pi^*) + \beta_2 x_{t-4} + \beta_3 €/\$_{t-1} \quad (5)$$

using the data of the Table 4, Equation (4) becomes:

$$i_t = 8.59879 + 0.595707(\pi_{t-2} - \pi^*) + 0.83068 x_{t-4} - 5.27647 €/\$_{t-1} \quad (6)$$

If $(\pi_{t-2} - \pi^*)$ =0 and x_{t-4} =0 the value of i_t will not be always the same,

but i_t will change with the fluctuations of €/$$_{t-1}$:

a. If $€/\$_{t-1}$ = 0.90 => i_t = 3.85

b. If $€/\$_{t-1}$ = 1.20 => i_t = 2.27

From the graphical analysis we get some further interesting features. We note that the line of the estimated values is much more unstable then the line of the actual values. Remembering the Taylor's words (1998), this difference between estimated and actual data could be a measure of the ECB's discretion. We should not forget that, even if a Central Bank uses a monetary rule, there is always a certain degree of freedom in the conduct of the monetary policy. As a consequence, it is impossible to perfectly mimic the actual data.

A last note regards Figure 1. During the second half of 2006 the line of the estimated values and the line of the actual data of the day-to-day rate diverge. It seems that the ECB's monetary policy becomes tighter in contrast with the estimated policy that shows a reduction of the interest rate. If we study the course of the lines of the three explanatory variables used in the formula, we find that probably the ECB, starting from the mid 2006, does not put any more attention on the path of the exchange rate. It seems that the ECB started to follow a different strategy in comparison with the past: a monetary policy with a smaller weight on the fluctuations of the Euro-Dollar exchange rate. Indeed, it is just before that period that the Euro-Dollar nominal exchange rate started to grow.

This result could strengthen the idea, sustained by some European politicians, that the rise of the interest rate since the end of 2006 had been inappropriate.

7. The inflation expectations

The analysis of the previous pages has highlighted two main features: first, the basic version of the Taylor Rule has not been able to reproduce in a suitable way the monetary policy of the ECB over a long period; second, the use of the Euro-Dollar exchange rate has improved the fitting of the rule.

In this paragraph I try to change the rule eliminating the inflation gap as a regressor and substituting it with a measure of the inflation expectations. It is well known the relevance of the inflation expectations in the studies on monetary policy (*see*, for example, Berck 2000, Bernanke 2004, and Cecchetti, and Debelle 2005). The role of the expectations is essential from different points of view: first, the anchoring of the inflation expectations is one of the basic tasks of the Central Banks; second, the stability of the expectations could be seen as an index of the Central Banks credibility; third, a solid stabilization of the expectations reduces the effects of negative shocks on the prices trend; and fourth, stable inflation expectations are essential to reduce the sacrifice that a Central Bank should impose on the economy in order to absorb a price shock. These four points, but there are obviously others not cited here, underline the importance of the inflation expectations in managing the monetary policy.

There are different way for using the inflation expectations in the Taylor Rule, and there are different studies that employ them (among the others *see* Gerlach, and Schnabel 1999, Sauer, and Sturm 2003, and Rotondi, and Vaciago 2007). In this paper I use two different indicators. A monthly qualitative indicator of the inflation expectations published by the European Commission and a quarterly quantitative value of the inflation expectations calculated by the ECB.

In the first attempts I employ the inflation expectations published by the European Commission. A systematic business and consumer survey is done each month and this survey provides "essential information for economic surveillance, short-term forecasting and economic research"[32]. In this survey there are two questions that are useful to my purpose. A question deals with the perception of the prices trend and another one is focused on the prices expectations. So, in the Taylor Rule I use the time series extrapolated from these questions. The questions are qualitative rather than quantitative. As a consequence, the consumers do not explicitly specify a value for the future inflation, but they only indicate what their opinion on the future speed of the prices is. A final indicator is created using these answers. This indicator is completely different from an inflation point estimate. As a consequence, in this case the coefficient on the inflation expectation has no more the significance given in the previous sections and we cannot say if the Taylor principle has been enforced or not in the Euro area. But, we can also draw important conclusions.

The Taylor Rule used in this attempt is the following one:

$$i_t = \beta_0 + \beta_1 \pi^e_{t-2} + \beta_2 x_{t-4} + \beta_3 \hat{U}/\$_{t-1} \quad (7)$$

π^e_{t-2} is the inflation expectation, that has the same lag of the inflation gap of the previous paragraphs, and x_{t-4} is, as usual, the output gap with four lags. We also have a lagged value of the bilateral Euro-US Dollar exchange rate.

The results of this regression, with inflation expectations, GDP gap and exchange rate as regressors are shown in Table 6 and in Figure 3. The coefficients have the expected sign and the coefficients on inflation expectations and exchange rate are significant. Only in the last period the two lines diverge but this is a problem of all the regressions that I have tested.

It seems that during this long period the ECB has focused the attention on these macro-indicators in order to develop the right monetary policy. The attention on the inflation expectations highlights the medium to long term vision of the ECB, in line with its mandate. Moreover, the results have underlined the key role of this inflation expectations time series in order to have a good performance of the Taylor rule over a long period. It is difficult to give an economic meaning to the value of the coefficient on the inflation expectations because it is a qualitative expectation and not a quantitative one but it is possible to affirm that the ECB is very vigilant in observing the course of this indicator. The value of the coefficient on the inflation expectations indicates that the ECB moves the interest rate by 0.059% if that value varies by 1. From this result one can wisely assert that the inflation expectations are crucial to the study of the Euro area monetary policy.

Then, the Euro-Dollar exchange rate has an important role too. In this case the ECB reduces (increases) the interest rate by 0.15% if the Euro appreciates (depreciates) by 10 cents against the Dollar.

Table 6. Taylor Rule with qualitative inflation expectations

OLS, 100 obs., sample 1999:05-2007:08, robust standard errors, dependent variable: day to day rate					
Explanatory Variable	Coefficient	Std. Err.	t-test	p-value	
Const	3,75679	1,04534	3,5939	0,00052	***
Inflexpect_2	0,0597757	0,0127212	4,6989	<0,00001	***
gdpgap_4	0,459608	0,298362	1,5404	0,12674	
€/$_1	-1,57295	0,850314	-1,8499	0,06741	*
Mean of dependent variable = 3,0453					
Standard deviation of dependent variable = 0,91656					
R.S.S. = 29,2958					
Standard Error of residuals = 0,552417					
R^2 = 0,647752					
Adj. R^2 = 0,636745					
F-test statistic (3, 96) = 26,4634 (p-value < 0,00001)					

[32] The Joint Harmonised EU Programme of Business and Consumer Surveys, User Guide, 2007, p. 2

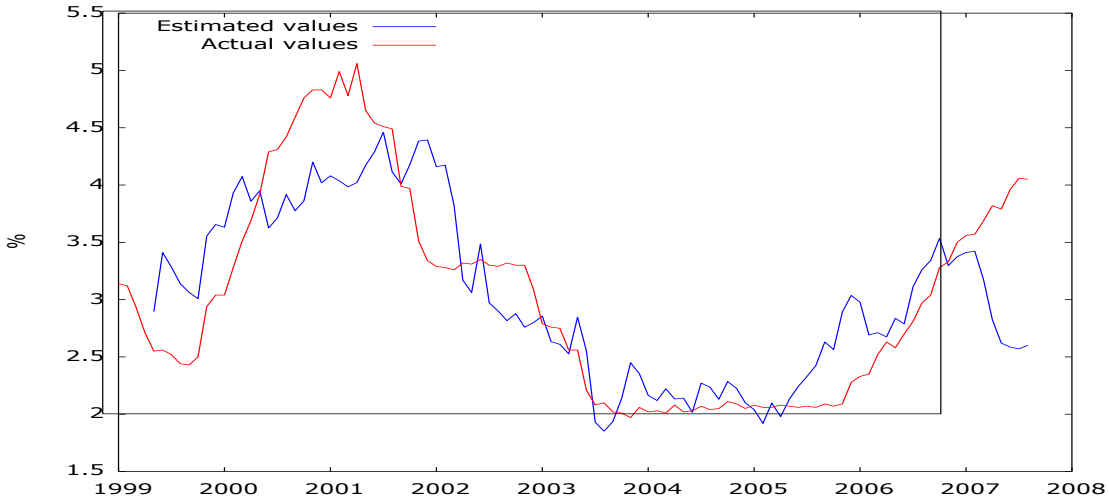


Figure 3. Day-to-day rate (actual and estimated values), T.R. with expectations

But there is a clear limit in this study: the use of this type of inflation expectations makes it impossible to directly compare these results neither with the ones obtained using the inflation gap nor with the other studies that employed a quantitative measure of the inflation expectations. I eliminate this limit at the end of this paragraph using another type of inflation expectations time series.

Before changing the inflation expectations time series, I show the results of a regression with the use of the inflation perception instead of the inflation expectations. This attempt can highlight two main aspects: firstly, it can show if the high significance level of the inflation expectations is only a fortunate case; secondly, it can also highlight if the ECB analyses the course of the inflation perception too. I used the time series of the inflation perception published by the European Commission. The methodology is the same of the inflation expectations but, in this case, the question is focused on the perception of the consumers about the course of the prices during the preceding 12 months[33].

In this case the estimated Equation is:

$$i_t = \beta_0 + \beta_1 \pi^p_{t-2} + \beta_2 x_{t-4} + \beta_3 €/\$_{t-1} \quad (8)$$

π^p_{t-2} is the inflation perception. The other regressors are the same of the preceding estimations.

The results are shown in Table 7. The coefficient on the inflation perception is not significant and the value of the p-value (0.81003) leaves no doubts. So, one can rightly affirm that the ECB is more concerned about the inflation expectations than the perception of past inflation. In theory, given the mandate of the ECB, this result is plain. But it was important to obtain an empirical result that is in line with what the theory and the ECB always affirm.

Besides, this result makes even more credible the one showed in Table 6.

Till now, one can only underline that the use of the inflation expectations increased the correspondence between the estimated and the actual values and that the expectations are statistically significant, differently from the perception. But we have not a *numerical idea* of the weight of the inflation expectations.

In order to overcome these limits, linked with the use of a qualitative inflation expectations time series, I change this time series with a quantitative one, published by the ECB[34]. This time series provides quarterly data on the inflation expectations. So, in order to harmonize the data, in the following regressions, I use quarterly data for all the macro indicators. The aim of these regressions is to estimate a coefficient that could show the quantitative attention of the ECB towards the fluctuations of this indicator.

[33] The exact question is: "How do you think the consumer prices have developed over the last 12 months?" and the possible choices are: "+ + risen a lot", "+ risen moderately", "= risen slightly", "- stayed about the same", "- - fallen", "N don't know".

[34] I used the "One Year Ahead" inflation expectations time series of the ECB survey of professional forecasters.

Table 7. Taylor Rule with qualitative inflation perception

OLS, 100 obs., sample 1999:05-2007:08, robust standard errors, dependent variable: day to day rate					
Explanatory Variable	Coefficient	Std. Err.	t-test	p-value	
Const	6,78254	0,496488	6,507	<0,00001	***
Inflpercept_2	0,00239066	0,00660928	0,241	0,81003	
gdpgap_4	0,890843	0,291248	2,897	0,00467	***
€/$_1	-3,45871	0,471812	-3,527	0,00065	***
Mean of dependent variable = 3,0453					
Standard deviation of dependent variable = 0,91656					
R.S.S. = 47,6965					
Standard Error of residuals = 0,704868					
R^2 = 0,426505					
Adj. R^2 = 0,408583					
F-test statistic (3, 96) = 23,7982 (p-value < 0,00001)					

The use of quarterly data obviously reduces the number of the observations. As a consequence, I have decided to expand the sample till the first quarter of 2008, even if this implies the use of data over the period of the subprime crisis.

Even in this case the Equation that I regress is:

$$i_t = \beta_0 + \beta_1(\pi^e_{t-1} - \pi^*) + \beta_2 x_{t-2} + \beta_3 €/\$_{t-1} \quad (9)$$

$(\pi^e_{t-1} - \pi^*)$ is the inflation gap calculated using the "one year ahead" inflation expectations of the period t-1. The other symbols have the usual meaning.

The results are shown in Table 8. All the coefficients have the expected sign and are highly significant.

Table 8. Taylor Rule with quantitative inflation expectations

OLS, 35 obs., sample 1999:III – 2008:I, robust standard errors, Dependent variable: day to day rate					
Explanatory Variable	Coefficient	Std. Err.	t-test	p-value	
Const	7,85957	1,01441	7,7479	<0,00001	***
InflexpecOYA_1	3,14009	0,724922	4,3316	0,00014	***
gdpgap_2	1,36625	0,247747	5,5147	<0,00001	***
€/$_1	-3,61714	0,912405	-3,9644	0,00040	***
Mean of dependent variable = 3,11857					
Standard deviation of dependent variable = 0,925079					
R.S.S. = 10,5444					
Standard Error of residuals = 0,583218					
R^2 = 0,637601					
Adj. R^2 = 0,60253					
F-test statistic (3, 31) = 17,24 (p-value < 0,00001)					

Figure 4 shows the estimated and the actual values. The lines show the usual divergence at the end of the sample.

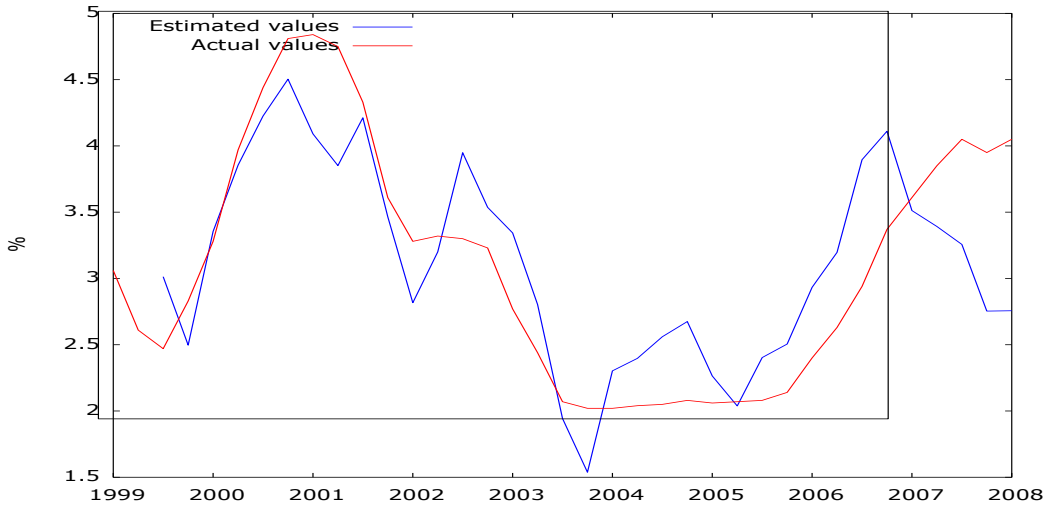


Figure 4. Day-to-day rate (actual and estimated values), T.R. with expectations, quarterly data

The results of this last regression show a very different situation in comparison with the results of the backward looking Taylor Rule used in the previous paragraphs. But, in order to directly compare the results I also regress this last Equation using real time quarterly data over the whole sample. The results of this case, shown in Table 9, are in line with the ones obtained in the previous paragraphs using monthly data.

The replacement of the real time data with the inflation expectations has radically increased the coefficient on the inflation gap (from 0.70 to 3.14). This change is in line with some other works (*see,* for example, Sauer, and Sturm 2003, or Rotondi, and Vaciago 2007). Using expectations, it seems that the ECB follows the Taylor principle. That is, the ECB increases the real interest rate when the inflation expectations go up. Moreover, the coefficient on the inflation expectations is very high. This highlights a very great attention of the ECB in defeating all the possible divergences of the inflation expectations from the target. The anchoring of the inflation expectations seems more important than a direct contrast of the actual inflation. Indeed, the coefficient calculated with real time data both with monthly and quarterly series has always been below the unity.

Table 9. Taylor Rule, basic version and quarterly data

OLS, 35 obs., sample 1999:III – 2008:I, robust standard errors, dependent variable: day to day rate					
Explanatory Variable	Coefficient	Std. Err.	t-test	p-value	
Const	5,49463	1,24749	4,4046	0,00012	***
Inflation gap_1	0,702295	0,280743	2,5016	0,01786	**
gdpgap_2	1,7495	0,315123	5,5518	<0,00001	***
€/$_1	-2,19142	1,20241	-1,8225	<0,07803	*
Mean of dependent variable = 3,11857					
Standard deviation of dependent variable = 0,925079					
R.S.S. = 16,2623					
Standard Error of residuals = 0,724286					
R^2 = 0,441085					
Adj. R^2 = 0,386996					
F-test statistic (3, 31) = 11,5731 (p-value = 2,97e-005)					

Furthermore, in the regressions with the quarterly inflation expectations the coefficient on the GDP gap is highly significant and its value is slightly above the previous ones. The Euro-Dollar exchange rate coefficient remains negative and significant as in the preceding regressions. The R^2 statistic has the highest value of all the regressions. It seems that the Taylor Rule can provide a very good fit of the data using inflation expectations.

In this Section I tested whether the inflation expectations have or not a concrete role in the conduct of the monetary policy in the Euro area. I present two types of analysis. Firstly, the regressions with the monthly data of the Business and Consumers Survey highlighted the difference between the expectations and the perceptions: the ECB only reacts against negative fluctuations of the inflation expectations, while it disregards the course of the inflation perception. Secondly, I used the *One Year Ahead* inflation expectations time series of the ECB

survey of professional forecasters in order to have a quantitative value of this *attention* towards the inflation expectations, since the Business and Consumers Survey does not supply a point estimate of the future inflation. This attempt gave a good result: it provided the evidence of the ECB great attention towards the inflation expectations. The elevated value of the coefficient on the inflation expectations means that the ECB is very alert in the surveillance of the fluctuations of the inflation expectations in order to stabilize them around the 2 per cent threshold.

8. Interest rate smoothing

A large part of the dominant theory suggests the use of a dependent lagged variable in the Taylor Rule (*see,* for theoretical and empirical examples Clarida, Galì, and Gertler 1999, Taylor 2000, Rotondi, and Vaciago 2007, or Sauer, and Strum 2003). This is due to the so called "interest rate smoothing". The smoothing can be seen has an evidence of the central bank activity or prudence. A high degree of interest smoothing means that the central bank is not so active in suddenly contrasting the signals that come out from the economy, but it prefers to act gradually towards its optimal rate.

Many studies have been undertaken in this field but their results are not always unanimous (for a short review of this literature *see* Sauer, and Sturm 2003). In my attempt I use a simple model to catch the interest rate smoothing.

The model used is the following:

$$i_t^* = \beta_0 + \beta_1(\pi_{t-1}^e - \pi^*) + \beta_2 x_{t-2} + \beta_3 €/\$_{t-1} \tag{10}$$

$$i_t = (1-\rho)i_t^* + \rho i_{t-1} \tag{11}$$

i_t^* is the optimal or target interest rate that the Central Bank should set every month. But this rate is not stable (as one can see watching the preceding figures) and so the Central Bank should change the nominal interest rate every month, often reversing its direction quickly. As a consequence, in order to eliminate this *problem*, the actual interest rate i_t gradually adjusts towards the optimal value, i_t^*, following the Equation (10), in which *ρ* is the smoothing parameter. In this way one can gauge the role of the lagged variable (that is, the role of the smoothing).

In this paragraph I use again quarterly data and the inflation expectations published by the ECB, since this specification has provided the best results. The symbols have the same meanings as in the preceding regressions. In the regression showed below, I drop out the exchange rate because I tested it and find that it does not radically improve the results. So, I decided to show only the basic rule with the smoothing parameter:

$$i_t = (1-\rho)\cdot[\beta_0 + \beta_1(\pi_{t-1}^e - \pi^*) + \beta_2 x_{t-2}] + \rho i_{t-1} \tag{12}$$

or equivalently

$$i_t = \tilde{\beta}_0 + \tilde{\beta}_1(\pi_{t-1}^e - \pi^*) + \tilde{\beta}_2 x_{t-2} + \rho i_{t-1} \tag{13}$$

where $\tilde{\beta}_i$ are the short run coefficients.

After estimating this Equation, it is possible to calculate two different types of coefficient: a short run coefficient, the product of (1-ρ) with the β_i, that is the $\tilde{\beta}_i$ coefficients, or the long run coefficients, the β_i. I show both the results. Table 10 shows the short run coefficients and the relevant statistics while Figure 5 shows the estimated values. I used the NLS estimator with robust standard errors.

Given these results, the resulting modified Taylor Rule is:

$$i_t = 0.438452 + 0.271160(\pi_{t-1}^e - \pi^*) + 0.793710 x_{t-2} + 0.889058 i_{t-1} \tag{14}$$

The coefficient of the lagged variable is highly significant and the value is in line with the relevant literature. The coefficient on the inflation gap is no more significant while the coefficient on the output gap remains highly significant.

Starting from this Equation it is straightforward to calculate the long run coefficients (see below). The values of both the long run and the short run coefficients are very high in comparison with the usual values gauged in other works. Indeed, the Taylor Rule should be:

$$i_t = 3.9521 + 2.44417(\pi^e_{t-1} - \pi^*) + 7.15431x_{t-2} \quad (15)$$

(8,090) (1,232) (3,069)

(the *t*-statistics are in the parentheses).

Table 10. Taylor Rule, interest rate smoothing

NLS, 35 obs., sample 1999:III-2008:I, robust standard errors, dependent variable: day to day rate					
Parameter	Estimate	Std. Err.	t-test	p-value	
$\tilde{\beta}_0$	0,438452	0,112024	3,914	0,00046	***
$\tilde{\beta}_1$	0,271160	0,203680	1,331	0,19280	
$\tilde{\beta}_2$	0,793710	0,144632	5,488	<0,00001	***
P	0,889058	0,0384405	23,128	<0,00001	***
Mean of dependent variable = 3,11857					
Standard deviation of dependent variable = 0,925079					
R.S.S. = 1,14261					
Standard Error of residuals = 0,191986					
R^2 = 0,96073					

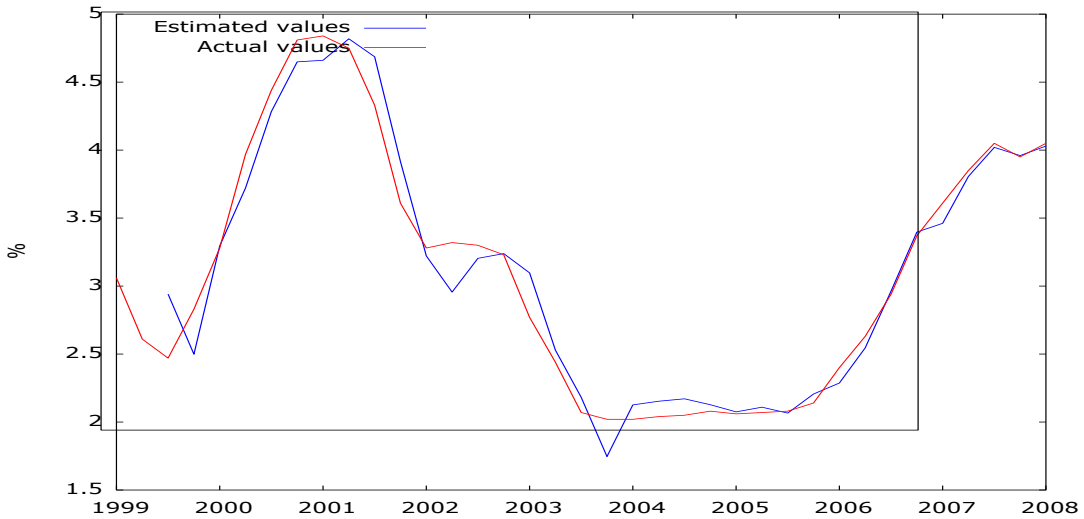


Figure 5. Day-to-day rate (actual and estimated values), T.R. with interest rate smoothing

The coefficient on the output gap is highly significant and it is quite high while the coefficient on the inflation gap is not significant. The results of this regression underline a high importance of the lagged dependent variable and of the output gap while the inflation gap is now no more significant. A similar result is shown in Sauer and Sturm (2003). But the peculiarity of this regression lays on the high value of the coefficients. A very impressive value: it implies that the ECB changes the interest rate with a very high reactivity.

At the end, the high degree of the interest rate smoothing is what we can draw from this paragraph together with the role of the output gap. It seems that the ECB has been very careful in changing the leading interest rate. It has preferred to smooth the variations of the interest rate towards the optimal interest: in this way the ECB does not transmit a wrong signal to the market or exceed in its monetary policy action.

9. Conclusions

In this paper I tried to formulate the simplest and more effective Taylor Rule that is able to reproduce the behaviour of the ECB during the first years of its activity. It is obviously arduous to summarize in one simple rule with few explanatory variables the job of a Central Bank, nevertheless I obtained some relevant findings.

I used different types of rule. The basic formula (with only two explanatory variables: the output gap and the inflation gap) has given bad results over a long period. The coefficients are highly significant but the R^2 of the regression is very low. Then, I tested the importance of the monetary aggregate M3. I showed that the role of M3 is irrelevant in deriving the right behaviour of the ECB. And this finding, as stressed in other works, is in contrast with the official general strategy announced by the same ECB. It seems that the "first pillar" does not have importance for the general strategy of the central bank.

In the subsequent attempt, I examined a Taylor Rule with three explanatory variables: inflation gap, output gap and Euro-Dollar nominal exchange rate, and I found that thanks to this rule the results have radically improved. This type of formula reproduces very well the path of the day-to-day interest rate over a long period of time. It depicts in a good way the behaviour of the European Central Bank. The interpretation of the estimated values presents some caveats, but this rule seems to work very well.

I also underlined that my GDP gap, that is, the methodology used in the paper to calculate the output gap, has been a valid tool for the analysis. It has been significant in almost all the model.

After these first findings, I expanded the analysis checking the relevance of the inflation expectations. I used two very different inflation expectations time series and, in both cases, the use of the inflation expectations improved in a very considerable way the fitting of the data. Moreover, the ECB seems to be more aggressive against the excessive fluctuations of the inflation expectations above the target than it is against the path of the actual inflation rate.

At the end I added a lagged dependent variable to test the importance of the interest rate smoothing and it seems that the interest rate smoothing is also fundamental in order to explain the monetary policy behaviour of the ECB.

References

[1] Ball, Lawrence. 1999. Efficient Rules for Monetary Policy. *International Finance*, Vol. 2, No. 1: 63-83.

[2] Berck, Jan Marc. 2000. Consumers' Inflation Expectations and Monetary Policy in Europe. *Contemporary Economic Policy, Oxford University Press,* Vol. 20: 122-132.

[3] Bernanke, Ben. 2004. The Great Moderation. Remarks by Governor Ben Bernanke at the meeting of the *Eastern Economic Association*, February, 20, 2004.

[4] Blinder, Alan. 1997. What central bankers could learn from academics and vice versa. *Journal of economic perspective,* Spring: 3-19

[5] Carare, Alina, and Tchaidze, Robert R.. 2005. The Use and Abuse of Taylor Rules: How Precisely Can We Estimate Them?. *IMF Working Paper* No. 05/148.

[6] Carstensen, Kai. 2006. Estimating the ECB Policy Reaction Function. *German Economic Review* 7: 1-34.

[7] Cecchetti, Stephen, and Debelle, Guy. 2005. Has the Inflation Process Changed?. *BIS Working Paper* No. 185.

[8] Clarida, Richard, Galì, Jordi, and Gertler, Mark. 1998. Monetary Policy Rules in practice. Some International Evidence. *European Economic Review* 42: 1033-1067.

[9] Clarida, Richard, Galì, Jordi, and Gertler, Mark. 1999. The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature, American Economic Association*, vol. 37(4): 1661-1707.

[10] Clarida, Richard, and Gertler, Mark. 1996. How the Bundesbank Conducts Monetary Policy. *NBER Working Papers,* No. 5581.

[11] Engle, Charles M. 2009. Currency Misalignments and Optimal Monetary Policy: a Reexamination. *NBER Working Paper,* No. 14289.

[12] Fourçans, André, and Vranceanu, Radu. 2002. ECB monetary policy rule: Some theory and empirical evidence. *ESSEC Working Paper* N.02008.

[13] Gerlach, Stefan. 2004. The Pillars of the ECB. *Economic Policy,* 40: 389-439.

[14] Gerlach, Stefan, and Schnabel, Gert. 1999. The Taylor Rule and Interest Rates in EMU Area: A Note. *BIS Working Paper* No. 73.

[15] Jaeger, Albert. 2003. The ECB's Money Pillar: An Assessment. *IMF Working Paper,* N. 03/82.

[16] Parsley, David C., and Popper, Helen. 2009. Evaluating Exchange Rate Management: An Application to Korea. *Hong Kong Institute for Monetary Policy Working Paper* No.28/2009.

[17] Rotondi, Zeno, and Vaciago, Giacomo. 2007. Lessons from the ECB Experience: Frankfurt Still Matters!. *Economic Notes,* Vol. 36, No. 2: 147-170.

[18] Sauer, Stephan, and Sturm, Jan-Egbert. 2003. Using Taylor Rules to Understand ECB Monetary Policy. *CESifo Working Paper Series* No. 1110.

[19] Taylor, John. B. 1993. Discretion versus policy rules in practice. *Carnegie Rochester Conference Series on Public Policy,* Vol. 39: 195-214.

[20] Taylor, John. B. 1998. Applying Academic Research on Monetary Policy Rules: an Exercise in Trasnational Economics. *The Manchester School Supplement:* 1-16.

[21] Taylor, John. B. 2000. The Monetary Transmission Mechanism and The Evaluation of Monetary Policy Rules. *Central Bank of Chile Working Paper* No. 87.

[22] Ullrich, Katrin. 2003. A Comparison Between the FED and the ECB: Taylor Rules. *ZEW Discussion Paper* No. 03-19.

MODELLING SHARE PRICES OF BANKS AND BANKRUPTS

Ivan O. KITOV
Russian Academy of Sciences, **Russia**
Institute for the Geospheres' Dynamics
ikitov@mail.ru

Abstract

Share prices of financial companies from the S&P 500 list have been modeled by a linear function of consumer price indices in the USA. The Johansen and Engle-Granger tests for cointegration both demonstrated the presence of an equilibrium long-term relation between observed and predicted time series. Econometrically, the pricing concept is valid. For several companies, share prices are defined only by CPI readings in the past. Therefore, our empirical pricing model is a deterministic one. For a few companies, including Lehman Brothers, AIG, Freddie Mac and Fannie Mae, negative share prices could be foreseen in May-September 2008. One might interpret the negative share prices as a sign of approaching bankruptcies.

Keywords: share price, modelling, CPI, prediction, the USA, bankruptcy

JEL Classification: E4, G1, G2, G3

1. Introduction

Recently, we have developed and tested statistically and econometrically a deterministic model predicting share prices of selected S&P 500 companies (Kitov 2010). We have found that there exists a linear link between various subcategories of consumer price index (CPI) and some share prices, with the latter lagging by several months. In order to build a reliable quantitative model from this link one needs to use standard and simple statistical procedures.

Following the general concept and principal results of the previous study, here we are predicting stock prices of financial companies from the S&P 500 list. In several cases, robust predictions are obtained at a time horizon of several months. In close relation to these financial companies we have also investigated several cases of bankruptcy and bailout. These cases include Lehman Brothers (LH), American International Group (AIG), Fannie Mae (FNM) and Freddie Mac (FRE). Regarding these bankruptcies, we have tested our model against its predictive power in May and September 2008. The main question was: Could the bankruptcies be foreseen? If yes, which companies should or should not be bailed out as related to the size of their debt?

In the mainstream economics and finances stock prices are treated as not predictable beyond their stochastic properties. The existence of a deterministic model would undermine the fundamental assumption of the stock market. If the prices are predictable, the participants would have not been actively defining new prices in myriads of tries, but blindly followed the driving force behind the market. It is more comfortable to presume that all available information is already counted in. However, our study has demonstrated that the stochastic market does not mean an unpredictable one.

In this paper, we analyze sixty six financial companies from the S&P 500 lists as of January 2010 as well as a few bankrupts from the financials. Some of the companies have been accurately described by models including two CPI subcategories leading relevant share prices by several months. Other companies are characterized by models with at least one of defining CPI components lagging behind related stock prices. We have intentionally constrained our investigation to S&P 500 – we expect other companies to be described by similar models.

Our deterministic model for the evolution of stock prices is based on a "mechanical" dependence on the CPI. Under our framework, the term "mechanical" has multiple meanings. Firstly, it expresses mechanistic character of the link when any change in the CPI is one-to-one converted into the change in related stock prices, as one would expect with blocks or leverages. Secondly, the link does not depend on human beings in sense of their rational or irrational behavior or expectations. In its ultimate form, the macroeconomic concept behind the stock price model relates the market prices to populations or the numbers of people in various age groups irrelevant to their skills. Accordingly, the populations consist of the simplest possible objects; only their numbers matter. Thirdly, the link is a linear one, i.e. the one often met in classical mechanics. In all these regards, we consider the model as a mechanical one and thus a physical one rather than an economic or financial one. Essentially, we work with measured numbers not with the piles of information behind any stock.

For the selected stocks, the model quantitatively foresees at a several month horizon. Therefore, there exist two or more CPI components unambiguously defining share prices several months ahead. It is worth noting that the evolution of all CPI components is likely to be defined, in part, by stochastic forces. According to the mechanical dependence between the share prices and the CPI, all stochastic features are one-to-one converted into stochastic behavior of share prices. Since the prices lag behind the CPI, this stochastic behavior is fully predetermined. The predictability of a measured variable using independent measured variables, as described by mathematical relationships, is one of the principal requirements for a science to join the club of hard sciences. Therefore, our stock pricing model indicates that the stock market is likely an object of a hard science.

A model predicting stock prices in a deterministic way is a sensitive issue. It seems unfair to give advantages to randomly selected market participants. As thoroughly discussed in (Kitov 2009b; Kitov, and Kitov, 2008; 2009ab) the models are piecewise ones. A given set of empirical coefficients holds until the trend in the difference between defining CPI is sustained. Such sustainable trends are observed in a majority of CPI differences and usually last between 5 and 20 years (Kitov and Kitov, 2008). The most recent trend has been reaching its natural end since 2008 and the transition to a new trend in 2009 and 2010 is likely the best time to present our model. As a result, there is no gain from the empirical models discussed in this paper. Their predictive power has been fading away since 2008. When the new trend in the CPI is established, one will be able to estimate new empirical coefficients, all participants having equal chances.

The remainder of the paper is arranged as follows. Section 1 introduces the model and data, which include stock prices of sixty six S&P 500 financial companies and seventy CPI components. In Section 2, empirical models are presented both in tabulated and graphical forms. For each model we have estimated standard deviation, which serves as a proxy to the model accuracy. For a few companies, the estimated models are robust over the previous 10 months. Section 3 tests these models statistically and econometrically. The Johansen (1988) and Engle-Granger (Newbold, and Granger 1967; Hendry, and Juselius 2001) tests both demonstrate that the null hypothesis of the existence a cointegrating relation between the observed and predicted time series cannot be rejected for a majority of companies. Therefore, the model is justified econometrically, and thus, all statistical inferences are valid. In Section 4, a crucial historical problem is addressed: Could one predict in May 2008 the evolution of financial stock prices? For some companies, the models estimated in the beginning of 2008 hold over the next year. Hence, the empirical modelling would have allowed accurate prediction of the evolution of stock prices, including those related to companies who filed for bankruptcy in several months. Finally, Section 5 investigates several cases of bankruptcy and bailout in the United States. It is found that many stock price trajectories would have been predicted to dive below the zero line.

The results of the presented research open a new field for the future investigations of the stock market. We do not consider the concept and empirical models as accurate enough or final. There should be numerous opportunities to amend and elaborate the model. Apparently, one can include new and improve available estimates of consumer price indices.

2. Model and data

Kitov (2009b) introduced a simple deterministic pricing model. Originally, it was based on an assumption that there exists a linear link between a share price (here only the stock market in the United States is considered) and the differences between various expenditure subcategories of the headline CPI. The intuition behind the model was simple – a higher relative rate of price growth (fall) in a given subcategory of goods and services is likely to result in a faster increase (decrease) in stock prices of related companies. In the first approximation, the deviation between price-defining indices is proportional to the ratio of their pricing powers. The presence of sustainable (linear or nonlinear) trends in the differences, as described in (Kitov, and Kitov 2008, 2009ab), allows predicting the evolution of the differences, and thus, the deviation between prices of corresponding goods and services. The trends are the basis of a long-term prediction of share prices. In the short-run, deterministic forecasting is possible only in the case when a given price lags behind defining CPI components.

In its general form, the pricing model is as follows (Kitov 2010):

$$sp(t_j) = \Sigma b_i \cdot CPI_i(t_j-\tau_i) + c\cdot(t_j-2000\ ) + d + e_j \qquad (1)$$

where $sp(t_j)$ is the share price at discrete (calendar) times t_j, j=1,…,J; $CPI_i(t_j-\tau_i)$ is the i-th component of the CPI with the time lag τ_i, i=1,..,I; b_i, c and d are empirical coefficients of the linear and constant term; e_j is the residual error, which statistical properties have to be scrutinized. By definition, the bets-fit model minimizes the RMS residual error. The time lags are expected because of the delay between the change in one price (stock or

goods and services) and the reaction of related prices. It is a fundamental feature of the model that the lags in (1) may be both negative and positive. In this study, we limit the largest lag to fourteen months. Apparently, this is an artificial limitation and might be changed in a more elaborated model. In any case, a fourteen-month lag seems to be long enough for a price signal to pass through.

System (1) contains J Equations for $I+2$ coefficients. Since the sustainable trends last more than five years, the share price time series have more than 60 points. For the current recent trend, the involved series are between 70 and 90 readings. Due to the negative effects of a larger set of defining CPI components discussed by Kitov (2010), their number for all models is (I=) 2. To resolve the system, we use standard methods of matrix inversion. As a rule, solutions of (1) are stable with all coefficients far from zero.

At the initial stage of our investigation, we do not constraint the set of CPI components in number or/and content. Kitov (2010) used only 34 components selected from the full set provided by the US Bureau of Labor Statistics (2010). To some extent, the original choice was random with many components to be similar. For example, we included the index of food and beverages and the index for food without beverages. When the model resolution was low, defining CPI components were swapping between neighbors.

For the sake of completeness we always retain all principal subcategories of goods and services. Among them are the headline CPI (C), the core CPI, i.e. the headline CPI less food and energy (CC), the index of food and beverages (F), housing (H), apparel (A), transportation (T), medical care (M), recreation (R), education and communication (EC), and other goods and services (O). The involved CPI components are listed in Appendix 1. They are not seasonally adjusted indices and were retrieved from the database provided by the Bureau of Labor Statistics (2010). Many indices were started as late as 1998. It was natural to limit our modelling to the period between 2000 and 2010, i.e. to the current long-term trend.

Since the number and diversity of CPI subcategories is a crucial parameter, we have extended the set defining components to 70 from the previous set of 34 components. As demonstrated below, the extended set has provided a significant improvement in the model resolution and accuracy. Therefore, we envisage the increase in the number and diversity of defining subcategories as a powerful tool for obtaining consistent models. In an ideal situation, any stock should find its genuine pair of CPI components. However, the usage of similar components may have a negative effect on the model – one may fail to distinguish between very close models.

Every sector in the S&P 500 list might give good examples of companies with defining CPI components lagging behind relevant stock prices. As of January 2010, there were 66 financial companies to model, with the freshest readings being the close (adjusted for dividends and splits) prices taken on December 31, 2009. (All relevant share prices were retrieved from http://www.finance.yahoo.com.) Some of the modeled companies do present deterministic and robust share price models. As before, those S&P 500 companies which started after 2004 are not included. In addition, we have modeled Fannie Mae and Freddie Mac, which are not in the S&P 500 list, and Lehman Brothers and CIT Group (CIT) which are out of the S&P 500 list. Due to the fact that the latter three companies are both bankrupts, they have been modeled over the period of their existence. Apparently, there are many more bankrupts to be modeled in the future.

There are two sources of uncertainty associated with the difference between observed and predicted prices, as discussed by Kitov (2010). First, we have taken the monthly close prices (adjusted for splits and dividends) from a large number of recorded prices: monthly and daily open, close, high, and low prices, their combinations as well as averaged prices. Without loss of generality, one can randomly select for modelling purposes any of these prices for a given month. By chance, we have selected the closing price of the last working day for a given month. The larger is the fluctuation of a given stock price within and over the months the higher is the uncertainty associated with the monthly closing price as a representative of the stock price.

Second source of uncertainty is related to all kinds of measurement errors and intrinsic stochastic properties of the CPI. One should also bear in mind all uncertainties associated with the CPI definition based on a fixed basket of goods and services, which prices are tracked in few selected places. Such measurement errors are directly mapped into the model residual errors. Both uncertainties, as related to stocks and CPI, also fluctuate from month to month.

2. Modelling financial companies

The results of modelling are presented in Table 1 and Appendix 2: two defining components with coefficients and lags, linear trend and free terms, and the standard error, σ, expressed in dollars. Negative lags, which correspond to leading share prices, are shown in bold. Figure 1 and Appendix 3 depict the observed and predicted curves. Five companies will be studied in more detail in Section 5: American International Group,

Citigroup (C), Fifth Third Bancorp (FITB), Legg Mason Inc. (LM), Moody's Corporation (MCO) and Morgan Stanley (MS).

Some financial companies have at least one defining CPI component lagging behind relevant stock price. For these companies, it is better to use the term "decomposition into" instead of "defining" CPI components. For example, share price of Aflac Incorporated (AFL) is defined by the index of financial services (*FS*) and that of transportation services (*TS*), the former lagging 2 months behind the share price and the latter leading by 6 months. Coefficient b_1 is positive. It means that the higher is the price for financial services the larger is the AFL's share price. The effect of the price index of transpiration services is opposite. Standard error for the model for the period between July 2003 and December 2009 is only $3.71. Figure 1 displays the observed and predicted prices for the period between 2003 and 2010. Before July 2003, the model does not hold and the curves deviate. Otherwise both curves are is a relatively good agreement including the sharp drop in 2008. From the statistical point of view, this is a key feature because any increase in the range of total change in the price and the defining CPIs is directly converted into higher model resolution.

Overall, standard errors in Table 1 and Appendix 2 vary from $0.77 for People's United Financial Inc. (PBCT) to ~$92 for AIG, which will be thoroughly analysed in Section 5. When normalized to the stock prices averaged over the whole period, the standard errors fluctuate less. However, for non-stationary time series with measurement errors dependent on amplitude the normalized errors are likely biased. The predicted curve is Figure 1 is very close to the observed one and foresees one month ahead. Actually, the predicted curve leads the observed one by one month.

American International Group was the first company bailed out by the US financial authorities in September 2008. This action introduced a bias into the link between AIG share price and defining CPIs, which existed before September 2008. The model listed in Table 1 is likely to be inappropriate as related to the link and not a robust one. The defining CPIs for the December 2009 model are as follows: the index for food away from home (*SEFV*) leading by 1 month and the index of prescribed drugs (*PDRUG*) leading by 13 months. In Section 5, we investigate the evolution of the bet-fit model from May 2008 to December 2009.

The model for Apartment Investment and Management Company (AIV) has both defining CPIs leading the share price: the index of pets and pet related products (*PETS*), a subcategory of the index for recreation, leads by one month and the index of prescribed drugs (*PDRUG*) is five months ahead of the price. At first glance, this set of defining CPIs does not look convincing. This might be an effect of the changing trend in the CPI. Before November 2009, the best-fit model included the index of food and beverages (*F*) and the *PDRUG,* both leading by 8 months. This set determined the best-fit model during the 12 previous months (Kitov 2010). In the smaller set of 34 CPI components used by Kitov (2010), the index of food and beverages and that of medical care (*M*) were the driving ones between November 2008 and October 2009 and provided the standard error of $2.058, but for a shorter period. With the *PDRUG*, the standard error for the same period between July 2003 and October 2009 is $2.055, i.e. only marginally better. This fact demonstrates how sensitive the model is to the defining CPIs. When more components are included, one could expect changes in the previously obtained models and lower standard errors.

The Allstate Corporation (ALL) has a model with both defining CPIs leading the share price. This model is unstable, however, and minimizes the RMS error only for the period between July 2003 and December 2009. In 2009, one of two defining components was randomly changing and one, the index for food away from home (SEFV), fixed. It is likely that the current set of defining CPIs do not include the one related to ALL. Thus, further investigations are needed.

Avalonbay Communities (AVB) has a model with one defining index (alcoholic beverages, *AB*) lagging behind the price by one month and the headline CPI less medical care (*CM*) leading by one month. This model is very stable over the previous 10 months and has a standard error of $1.36.

American Express Company (AXP) has a model predicting at a four month horizon. The defining CPIs are the index for food and beverages leading by 4 months and the index for medical care leading by 10 months. In the previous study (Kitov 2010) the model was essentially the same. So, the extended CPI set does not make a better model. The model is a robust one and minimizes the standard error for the period between July and November 2009 as well.

The model for Bank of America (BAC) is defined by the index of food and that of food away from home. The latter CPI component leads by 13 months. From our past experience, the larger is the lag the more unreliable is the model. However, both defining components provide the best fit model in the second half of 2009. Both coefficients in the BAC model are negative. This means that increasing food price forces the share price down.

The growth in the indices of food and food away from home has been compensated by linear time trend in the share price.

Table 1. Empirical 2-C models for selected S&P 500 financial companies

Company	b_1	CPI_1	τ_1	b_2	CPI_2	τ_2	c	d	σ, $
AFL	0.59	FS	**-2**	-2.37	TS	6	12.12	349.73	3.71
AIG	-191.36	SEFV	1	38.53	PDRUG	13	727.81	21116.78	92.3
AIV	-1.64	PETS	1	1.09	PDRUG	5	-1.61	-139.18	2.15
ALL	0.07	E	11	-6.86	SEFV	2	45.20	1106.19	2.82
AVB	0.57	CM	1	1.92	AB	**-1**	-12.01	-345.48	1.36
AXP	-3.81	F	4	-2.00	M	10	49.93	1115.79	2.48
BAC	-2.95	FB	3	-2.97	SEFV	13	35.43	956.52	2.53
BBT	-1.57	F	3	-0.31	FRUI	13	12.58	332.95	2.06
BEN	-7.95	FB	3	6.59	VAA	13	60.58	564.48	6.46
BK	-0.69	MEAT	13	-1.65	PETS	1	15.80	270.32	2.09
BXP	4.58	MCC	5	-5.04	PETS	3	5.63	-605.31	5.04
C	2.54	HO	5	-8.26	SEFV	2	36.70	1048.90	2.53
FITB	-4.85	SEFV	2	1.45	HS	6	21.19	621.56	1.82
HBAN	-2.15	RPR	13	1.32	FOTO	13	17.23	252.80	0.93
HCN	-1.80	PETS	2	0.76	HOSP	5	-7.06	-40.40	2.20
GS	21.06	HO	10	-29.45	SEFV	3	111.40	2496.20	13.48
JPM	-2.49	F	4	3.19	ORG	0	26.31	139.00	2.49
L	-2.49	FB	5	-1.51	TS	3	28.35	679.85	2.07
LM	-6.01	F	4	-8.17	APL	13	33.07	1754.81	6.89
PNC	1.49	CM	0	-3.44	FB	4	16.37	331.72	3.49
PSA	-4.14	SEFV	3	2.04	PDRUG	5	14.25	72.98	4.43
VNO	-11.08	SEFV	3	2.23	PDRUG	5	57.23	1113.14	5.30

Franklin Resources (BEN) is driven by the index of food (*FB*) and that of video and audio (*VAA*), both leading by several months. The former component has a negative coefficient and the latter one – positive. Bank of New York Mellon Corporation (BK) is defined by the index of meats, poultry and fist (*MEAT*) and the index of pets and per related products, both having negative coefficients. The model has a standard error of $2.09.

Boston Properties (BXP) has a model with the index of medical care commodities (*MCC*) and *PETS* leading by 5 and 3 months, respectively. This is a relatively stable model. However, the best-fit model was different before September 2009 and included the index of food and the index of miscellaneous services (*MISS*). This model had been reining since March 2009. The model obtained in (Kitov, 2010) was based on the CPI less energy (*CE*) and the index of food. It was a mediocre model with the RMS error of $5.54 compared to $5.11 obtained in this study for the same period.

Citigroup is of special interest. This company was bailed out in November 2008. For the purposes of share modelling the bailout introduces a major disturbance, because the share is not the one estimated by the free stock market any more. Accordingly, the models obtained after November 2008 are likely to be biased. In Table 1, a share of Citigroup is defined by the index of household operations (HO) and that of food away from home (SEFV). Coefficient b_1 (=+2.54) is positive and the increase in *HO* should be converted into a higher share price. The effect of *SEFV* is an opposite one with a larger coefficient b_2=-8.26. In 2007 and 2008, the index of household operations was increasing at almost the same rate as the *SEFV*, and the share fell from $50 in April 2007 to $1.5 in January 2009.

The next ten companies in Table 1 are all robust and have deterministic price models with no one of defining indexes lagging. Huntington Bancshares (HBAN) is controlled by the index of primary residence rent (*RPR*) and the index for photography (*FOTO*), both leading by 13 months. This model holds at least over the 10 months previous to December 2009. Coefficient b_1 is negative and any increase in *RPR* is converted into a

decrease in the HBAN share price 13 months later. It is instructive to track this model in 2010. It must be fading away with the transition to a new trend in the CPI. The model standard deviation is only $0.92 for the whole period. Figure 1 displays relevant curves. Except four of five points, the agreement is excellent.

Health Care REIT (HCN) has a model defined by the PETS and the index of hospital services (*HOSP*). The latter has a permanent positive trend and likely is compensated by the linear trend term with a negative slope (-40.4). The predicted and observed curves are very close. So, the model accurately predicts at a two-month horizon.

Goldman Sachs (GS) is a famous bank. The trajectory of its share price is well predicted by the index of household operations (*HO*) and the index of food away from home (*SEFV*) at a three month horizon. Since coefficient b_1 is positive the decreasing price index of household operations results in a fall in GS share price, as was observed in 2008 and 2009. In the second half of 2009, the price was on rise.

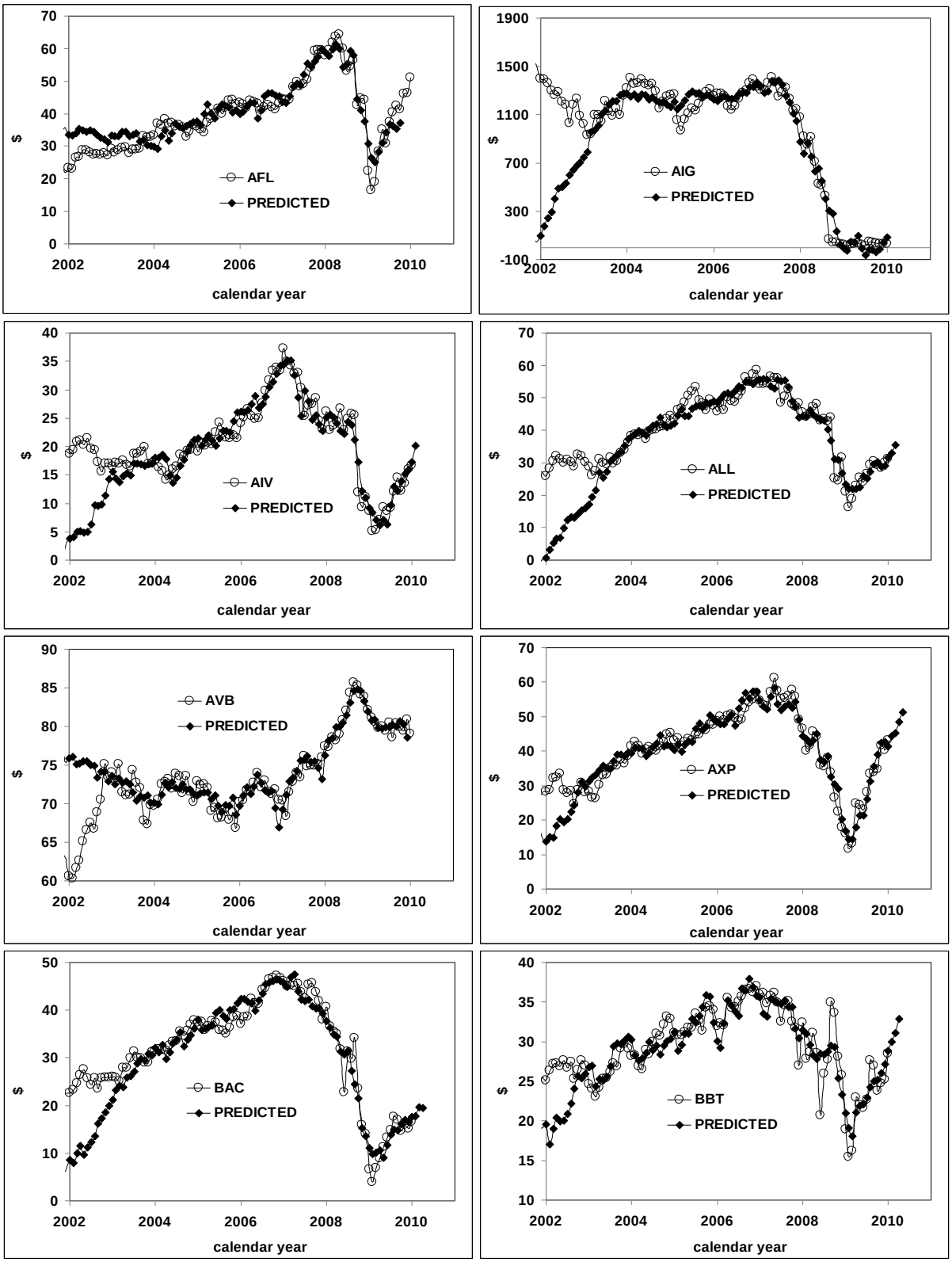
JPMorgan Chase & Co. is defined by food and beverages (*F*) and other recreation goods (*ORG*). Since the time lag of the *ORG* is zero the model can predict only contemporary share price. In Figure 1, the observed and predicted curves almost coincide before 2007. The years between 2007 and 2009 are characterized by extremely high fluctuations in the observed price. The model failed to predict this feature. In 2009, the prediction is good again, however. Therefore, the fluctuations are likely to be related to short-term forces not affecting fundamental long-term link between the price and the defining CPIs.

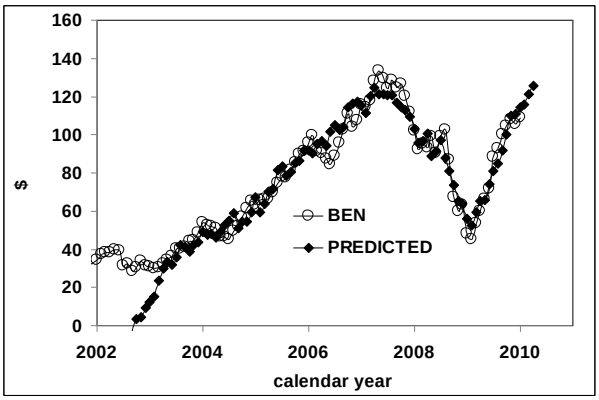
The best-fit model for Loews Corporation (L) includes the index of food (*FB*) leading by 5 months and the index of transportation services (*TS*) leading by 3 months. Both coefficients are negative and are counteracted by a positive slope c=28.35. The model for Legg Mason (LM) is based on the index of food and beverages and the index of appliances (*APL*) from the housing index, the latter leading by 13 months. Overall, the predicted time series is very close to the observed one with standard deviation of $6.89. The largest input to the standard deviation comes from a short period in 2006. Otherwise, both curves are very close even during the dramatic fall from $80 per share in the end of 2007 to $10 per share in February 2009 and during the fast recovery in 2009.

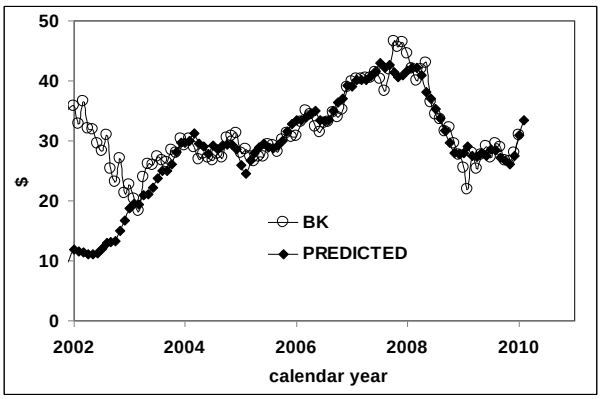
PNC Financial Services (PNC) relies on the headline CPI less medical care (*CM*) and the index of food (*FB*), the model being a contemporary to the share. Public Storage (PSA) and Vornado Realty Trust have similar models defined by the index of food away from home and the *PDRUG*. The time lags are also identical and are 3 and 5 months, respectively. Figure 1 demonstrates that the observed prices of PCA and VNO are similar with a local peak in the first half of 2008. A similar pattern is observed for AIV, which model also includes *PDRUG*. The difference between PCA and VNO is in the sensitivity to *SEFV*: b_1(PSA)=-4.14 and b_1(VNO)=-11.08.

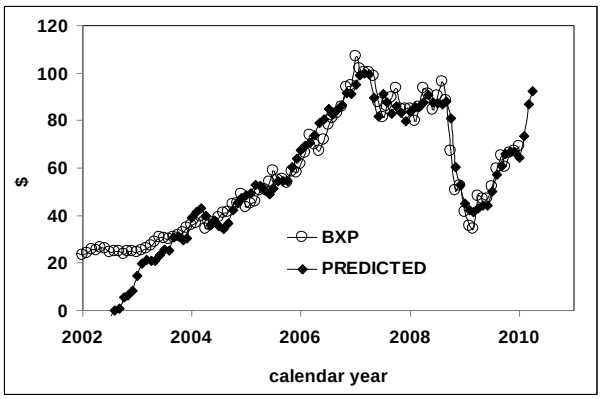
So, among the models with both defining CPIs leading relevant shares, there are examples of robust models and unstable models. For the latter companies, no fixed model is available over the past year. It is likely that these models express the lack of true defining indices in the current set of CPIs and are affected by random measurement noise. One cannot exclude that true robust models do exists for these companies.

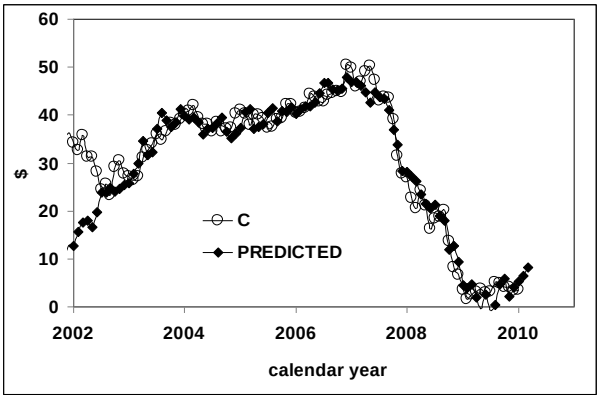
For other forty four financial companies relevant models and graphs are presented in Appendices 2 and 3. These models use quiet several defining CPIs not mentioned in Table 1. Otherwise, Table 1 contains all meaningful configurations of leading and lagging share prices and those in the Appendices are given for the sake of completeness.

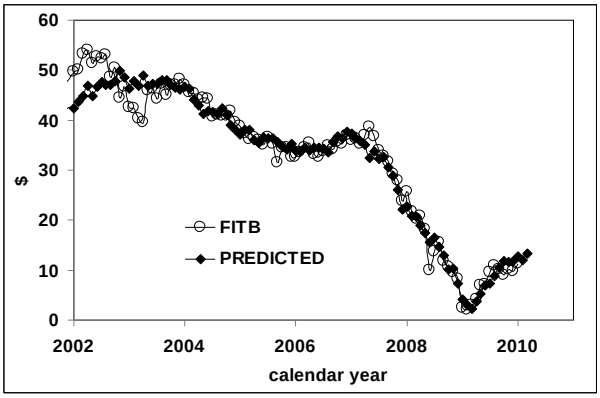

AFL
PREDICTED
AIG
PREDICTED
AIV
PREDICTED
ALL
PREDICTED
AVB
PREDICTED
AXP
PREDICTED
BAC
PREDICTED
BBT
PREDICTED
$
calendar year
2002
2004
2006
2008
2010

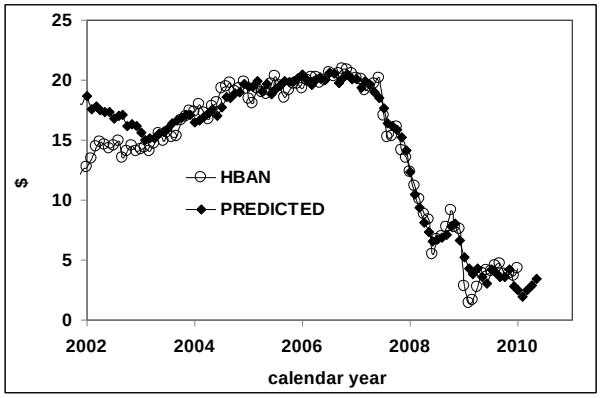

160
140
120
100
80
60
40
20
0
$
BEN
PREDICTED
2002
2004
2006
2008
2010
calendar year

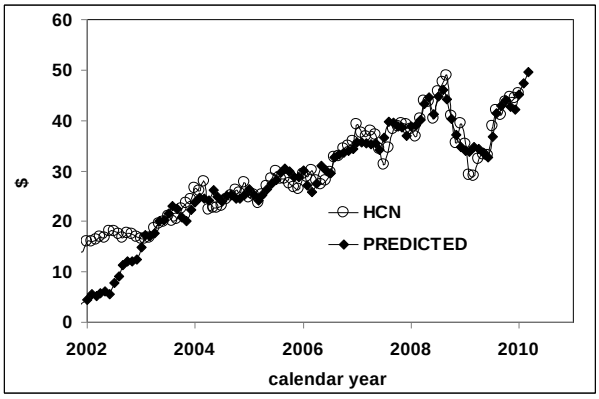

50
40
30
20
10
0
$
BK
PREDICTED
2002
2004
2006
2008
2010
calendar year

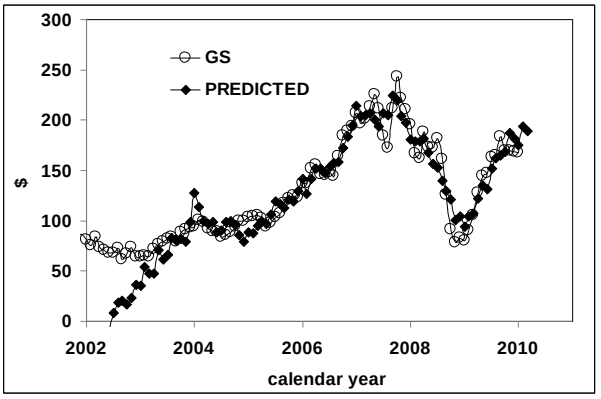

120
100
80
60
40
20
0
$
BXP
PREDICTED
2002
2004
2006
2008
2010
calendar year


60
50
40
30
20
10
0
$
C
PREDICTED
2002
2004
2006
2008
2010
calendar year


60
50
40
30
20
10
0
$
FITB
PREDICTED
2002
2004
2006
2008
2010
calendar year


25
20
15
10
5
0
$
HBAN
PREDICTED
2002
2004
2006
2008
2010
calendar year


60
50
40
30
20
10
0
$
HCN
PREDICTED
2002
2004
2006
2008
2010
calendar year


300
250
200
150
100
50
0
$
GS
PREDICTED
2002
2004
2006
2008
2010
calendar year

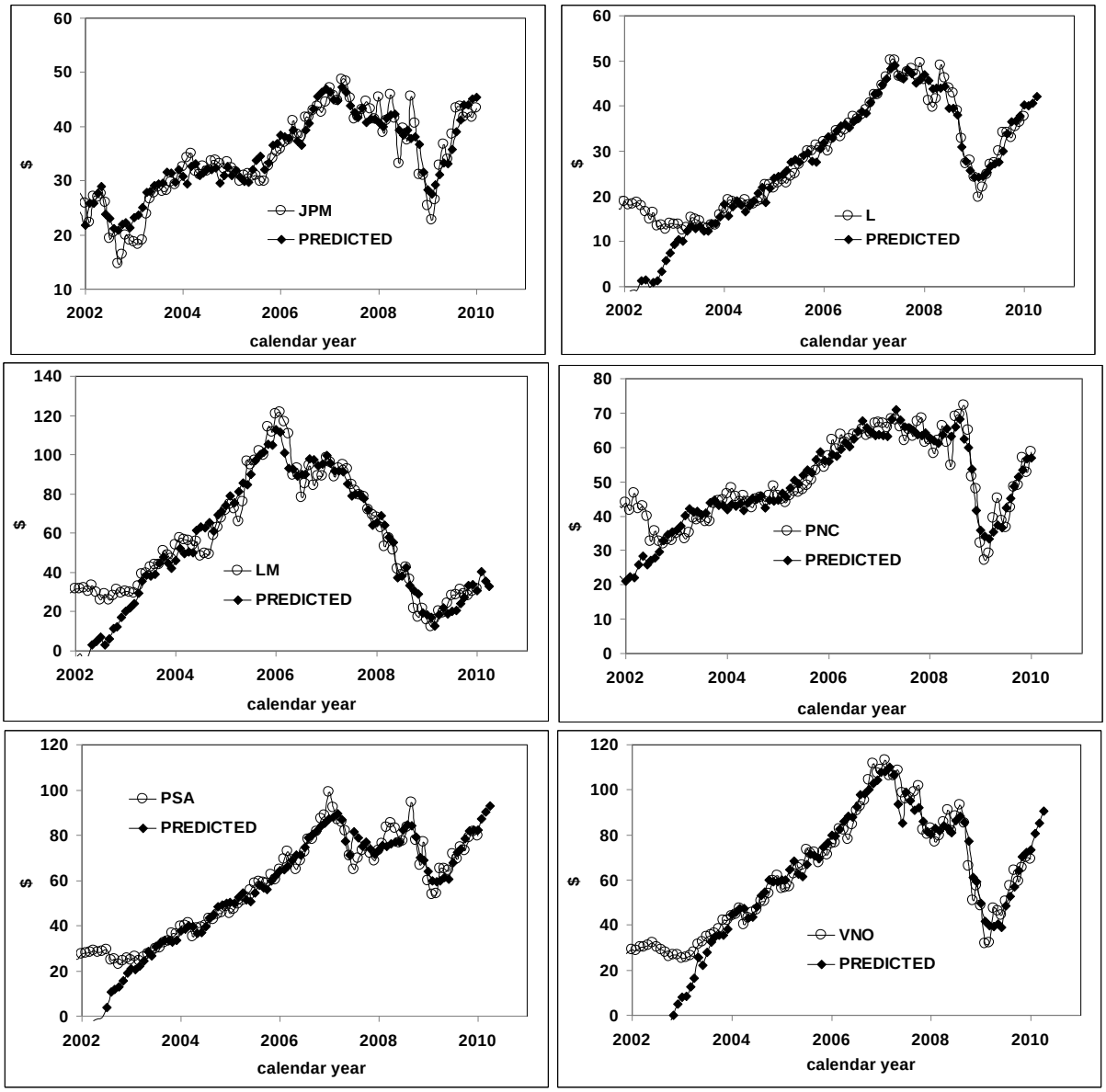


Figure 1. Observed and predicted share prices of eight financial companies from the S&P 500 list. Relevant empirical models are presented in Table 1.

The principal purpose of Section 2 consists in presenting tentative empirical models for share prices of financial companies. The current set of defining CPIs is far from a complete one and further investigation may reveal more accurate and reliable models for the same companies. However, the current models might be good enough because of high correlation between various CPI components. For a given share, the currently used CPIs may be very close to the true defining CPIs, which are not included in the set yet. Therefore, a direct statistical estimate of the model accuracy and reliability is a major task.

3. Cointegration tests

Statistical properties of the residual error are crucial for any quantitative model. Ideally, a good model involving time dependent measured variables should describe the essence of real ties. The model residual error should be a stationary (and thus, integrated of order zero) stochastic process with characteristics close to the uncorrelated white noise with Gaussian distribution. In the best case, residual errors should depend only on measurement errors, with the measurements conducted in a consistent way.

As in the previous study (Kitov 2010), we applied the Johansen cointegration test to the observed time series and those predicted in Section 2. For all studied companies, the test resulted in cointegration rank 1 or, for

two non-stationary variables, in the presence of one cointegrating relation. All results are listed in Table 2. The Johansen approach does not require both variables to be in the same order of integration.

As an alternative, we have applied the augmented Dickey-Fuller (ADF) and the Phillips-Perron (PP) tests for unit roots to all residual errors of the models in Table 1, except those with defining CPI lagging behind relevant shares. This procedure is in line with the Granger-Engle two-step method based on several specification tests applied to the residual time series. Having the same econometric power as the Johansen procedure, the Granger-Engle test allows a larger variety of specifications.

Table 2. Results of Johansen cointegration test and unit root tests as applied to the residual errors

	ADF	PP	Johansen test		
	z(t), 1%CV[1]=-3.54	z(ρ), 1%CV=-19.4	eigenvalue	trace statistics, 5%CV=3.76	rank
AIG	-4.10	-29.8	0.21	0.24	1
AIV	-6.19	-47.5	0.35	0.15	1
ALL	-5.95	-46.5	0.32	0.04	1
AXP	-4.98	-38.3	0.25	0.35	1
BAC	-4.87	-38.5	0.21	0.003	1
BBT	-5.96	-47.7	0.29	0.08	1
BEN	-4.53	-34.0	0.22	1.90	1
BK	-5.00	-40.0	0.26	0.07	1
BXP	-5.47	-39.6	0.33	0.87	1
C	-5.60	-41.0	0.36	0.27	1
FITB	-5.63	-45.8	0.26	2.39	1
GS	-4.56	-34.0	0.41	1.69	1
HBAN	-5.44	-41.5	0.29	0.29	1
HCN	-5.27	-39.0	0.37	1.91	1
JPM	-6.22	-51.2	0.30	1.15	1
L	-5.70	-43.5	0.37	0.96	1
LM	-4.75	-37.2	0.18	0.11	1
MS	-6.20	-48.4	0.33	0.09	1
PNC	-6.12	-50.0	0.33	0.47	1
PSA	-5.84	-44.3	0.33	2.18	1
VNO	-5.46	-40.0	0.33	0.83	1

[1]CV – critical value

In a sense, this Section is a fully technical one. We need only a confirmation that the regression technique used in Section 2 is applicable, i.e. the regression does not give spurious results. Both tests for cointegration unambiguously evidence the presence of long-term equilibrium relations between the actual and predicted prices. The predicted prices can be considered as weighted sums of prices for goods and services. In this regard, they are similar to the overall CPI and can be considered as independent measurements and represent just one variable. Therefore, one does not need to test both defining CPI for cointegration with relevant share price.

So, one can derive a conclusion that the deterministic pricing model provides a statistically and econometrically valid description of share prices of S&P 500 financial companies. There is a problem with the model resolution, however. As happens often in physics, in order to obtain a consistent and reliable model one should have a wider dynamic range of involved variables or to increase the accuracy of measurements. The latter is hardly possible with the past CPI readings. So, one could expect a more reliable model for the companies with share prices varying the most.

4. May 2008 vs. December 2009

The current models predicting future prices are of crucial interest for the stock market. It is always important to know which stocks will go up/down and at what rate. However, there are significant problems related to the past performance of the stock market also to be considered. One of these problems is associated with the 2008/2009 financial and economic crisis, which exposed many companies to major risks. Since the late 2007 and very actively since July 2008, stock prices of many companies have been decreasing at an accelerating speed.

The decrease costs trillions US dollars net lost after the overall asset devaluation. This is a natural challenge to our concept: Could the model predict the fall in stock prices if available in 2008?

For all investors and owners it would have been a great relief to predict, and thus, prevent or reduce the loss. Here we would like to stress again that the model is valid only when it does not disturb natural functioning of the stock market, i.e. those myriads of well-established direct and indirect interactions between economic and financial agents. When everybody shifts to one or few "salvage" stocks, their behavior becomes highly distorted, biased, and thus unpredictable. A part of the financial market is never equivalent to the whole market and this model will be worthless when used by all market players. So, we would not recommend using the model shortly after this book is published. In a sense, this publication may destroy the market configuration described by the model.

The principal question posed in this Section can be addressed quantitatively. As a first step, we move back in May 2008 and use contemporary CPI data to obtain the best-fit models for the S&P 500 share prices under study. Table 3 and Appendix 4 list these models obtained for selected financial companies. One should bear in mind that the involved prices had only limited dynamic range in the beginning of 2008 and corresponding models are not fully resolved. In this sense, the 2009 models are superior.

Then, we calculate all share prices using the 2008 models and actual CPI data between May 2008 and December 2009. In Figure 2 we compare the 2008 predictions to those obtained in January 2010 (i.e. the models for December 2009) and described in Section 2. If both models for a given share provide similar predictions then the 2008 fall was predictable, at least for the company.

A few companies in Table 3 have one defining component of the same nature as that in relevant December 2009 models. For a majority, both defining indices are different. This effect is observed despite our intention to select those 2008 models which provide the best prediction.

Table 3. Defining CPI components, coefficients and time lags for the models in May 2008.

Company	b_1	CPI_1	τ_1	b_2	CPI_2	τ_2	c	d
AFL	0.36	DIAR	0	-0.55	ITR	12	8.92	38.50
AIG	-21.11	DIAR	9	-172.66	**SEFV**[1]	2	1148.11	31872.06
AIV	1.78	VAA	3	1.23	**PDRUG**	7	-10.77	-525.90
ALL	1.85	MCC	7	-2.22	APL	9	-10.18	-207.35
AVB	1.41	EC	4	1.59	**AB**	1	-10.47	-340.92
AXP	-7.87	SEFV	4	1.20	HS	6	47.73	1098.50
BAC	-2.72	**F**	3	1.77	TS	12	8.86	115.22
BBT	-1.98	RPR	13	-1.16	MISS	12	25.57	648.22
BEN	3.28	PDRUG	12	-1.90	HOSP	3	30.13	-368.96
BK	-1.44	H	11	3.13	HO	11	-2.39	-75.70
BXP	-3.96	F	4	3.17	**MCC**	7	14.55	-136.73
C	-4.02	FB	8	-4.00	RPR	13	46.97	1376.82
FITB	-0.40	DIAR	7	0.84	PDRUG	8	-12.03	-112.79
GS	12.87	MVP	5	27.41	FOTO	12	65.74	-4221.10
HBAN	-2.07	SEFV	6	-1.66	**RPR**	13	22.63	638.78
HCN	0.20	FU	8	-2.27	VAA	0	1.07	225.28
HST	-2.81	SEFV	3	0.55	PDRUG	7	13.60	290.45
JPM	-2.54	**F**	4	-0.45	MEAT	13	20.29	484.70
L	2.02	MVP	5	-5.37	SEFV	5	33.34	643.83
LM	-2.56	MEAT	7	-12.00	**APL**	13	12.38	1528.90
PNC	2.43	DUR	10	-5.69	SEFV	4	43.97	617.15
PSA	6.13	MVP	5	-9.54	RPR	6	57.57	1103.99
VNO	2.63	**PDRUG**	7	8.56	FOTO	10	10.67	-1683.89

[1] same defining CPIs as in Table 1 are highlighted.

The May 2008 model for ALL, which includes the index of diary products (*DIAR*) and the index of intracity transportation (*ITR*), predicts the evolution of the share price relatively well till July 2009. Then the May prediction starts to diverge at an accelerating rate from the observed trajectory. The 2008 trough could be forecasted both in time and amplitude in May 2008.

For several companies, the agreement between the May 2008 and December 2009 predictions does not disappear even when models are different. For example, the 2008 model for AXP, defined by the *SEFV* and the index of household services (*HS*), does not diverge from the observed trajectory since 2008. Similar situation is observed with the model for Host Hotels&Resorts (HST). This effect shows the necessity of a complete or at least more representative set of CPIs. Otherwise, one can not distinguish between two neighboring models with defining CPI components characterized by a high degree of correlation.

The 2008 and 2010 models for BXP have one common defining variable – the index of medical care commodities (*MCC*). Nevertheless, the 2008 model fails to predict the future trajectory well. This is due to the difference between the indices for food and beverages and pets. Similar effect is observed with JPM.

The 2008 model for GS demonstrates a striking difference with the observed time series. It predicted the fall in the share price down to the zero line in the second half of 2009. In reality, no catastrophic drop happened and the price fell only to the level of $70. From the actual time series it is clear that the model for GS could not be well resolved because of very limited change in the share price by May 2008. There are several financial companies with shares predicted to fall below the zero. Some predictions were accurate enough. These companies are modeled in Section 5.

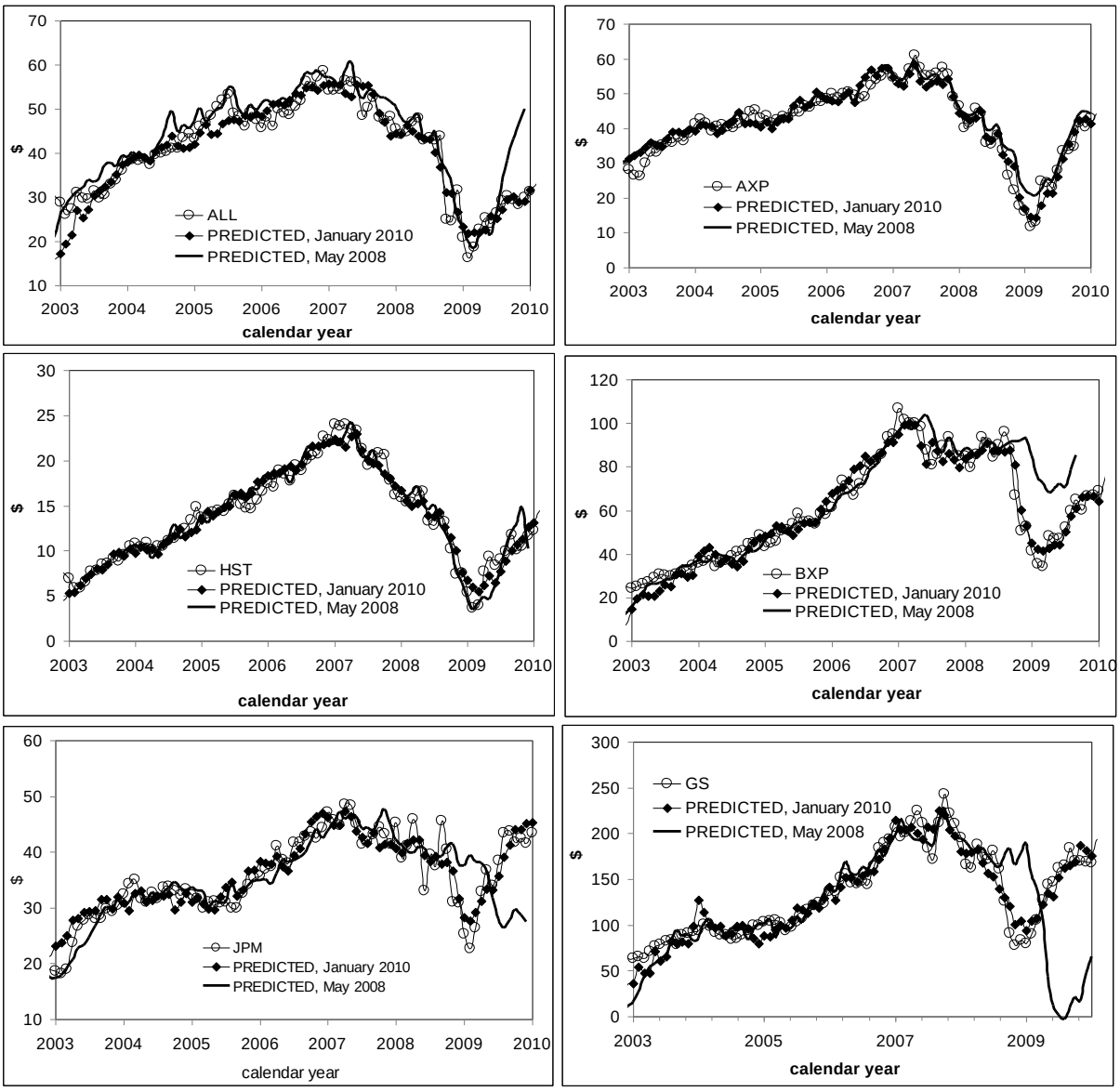


Figure 2. Comparison of stock prices predicted in May 2008 and December 2009.

The number of successful models is relatively small if to consider the initial set of S&P 500 companies. This fact raises delicate questions about the reliability of the models and the concept itself. One may assume that the successful models are just a lucky accident. The concept should be validated by modelling of more companies, extension of the set of defining CPI components, and usage of new data in the years to come.

We also take into account the fact that quantitative models, also in physics, are better resolved than all involved variables vary in wider ranges. Specifically, the difference between the 2008 and 2009 models consists in the sharp fall after July 2008. Therefore, the models obtained in 2009 are better resolved and thus superior to those from 2008. Data available in 2008 did not allow identification of right models because of high correlation between subcategories of the consumer price index. Good news is that the right models hold once and for all, but with new coefficients.

5. Predicting bankruptcy

In Section 4, we have modelled the evolution of share prices of several financial companies from the S&P 500 list between May 2008 and December 2009. It was found that some predicted share prices sank below the zero line. Under our framework, the presence of a negative stock price may be considered as an equivalent to a net debt. When long enough and without any positive prospective, such a debt would likely result in a bankruptcy.

In reality, some companies with negative predicted share prices declared bankruptcy, some were bailed out and some have been suffering tremendous difficulties since 2008. The first group is represented by Lehman Brothers who filed for Chapter 11 bankruptcy protection on September 15, 2008. The net bank debt was estimated at the level of $600 billion. More than 100 banks filed for bankruptcy since then.

Several banks were bailed out, with American International Group the first to obtain a $150 billion government bailout. The AIG bailout was presented as a major move to save the collapsing US financial system. The biggest examples of bailout are also Fannie May and Freddie Mac. All three companies had a sharp share price fall in the second half of 2008.

CIT Group Inc. (CIT) got $2.3 billion of bailout money in December 2008 and $3 billion bond holder bailout in July 2009. However, it did not help and CIT declared bankruptcy in November 2009. These companies and many others have been struggling and likely will struggle in the future trying to restructure their debts and re-enter the stock market.

Section 5 seeks to answer a number of questions:

- Was it possible to predict the evolution of total debt of the bankrupts?
- Was it possible to predict the dates of these bankruptcies?
- Is it possible to predict the date of recovery?
- It is possible to predict future bankruptcies?
- Which company had to be bailed out and when?

All S&P 500 models with negative share prices were obtained together with other models for May 2008. In this regard we should not distinguish them. The reason for a separate investigation consists in the fact that negative share prices might result in bankruptcies. This is a phenomenon no described quantitatively by our models and thus deserving special attention. Otherwise, all models were equivalent and obtained according to the same procedures. It is worth noting that the models for the same companies obtained in October 2009 are highly biased by bailouts or do not exist together with bankrupt companies.

Table 4. Models for 10 companies: May, September and December 2008, and October 2009

May 2008

Company	b_1	CPI_1	τ_1	b_2	CPI_2	τ_2	c	d
AIG	-21.11	DIAR	9	-172.66	SEFV	2	1148.11	31872
C	-4.33	FB	4	-3.63	RPR	12	46.79	1358
CIT	-4.84	F	5	11.51	SEFV	6	96.99	2610
FITB	1.46	MCC	9	-0.32	DIAR	8	13.01	227.5
FNM	9.62	RS	3	10.84	SEFV	6	24.36	733.0
FRE	-3.54	DUR	2	-9.66	RPR	13	57.75	2180
LEH	-6.27	FB	4	-1.38	HOSP	3	77.60	1411
LM	-2.57	MEAT	7	-12.02	APL	13	12.40	1532

MCO	-5.50	F	5	-5.83	RPR	9	75.37	1909
MS	7.788	R	7	-0.85	DIAR	4	1.49	-658.8

September 2008

Company	b_1	CPI_1	τ_1	b_2	CPI_2	τ_2	c	d
AIG	-22.10	DIAR	9	-178.42	SEFV	1	1198	32967
C	-4.26	FB	9	-3.62	RPR	12	46.39	1345
CIT	-0.77	DAIR	8	-8.20	RPR	11	59.95	1584
FITB	-3.07	F	12	1.06	PDRUG	8	-0.97	250.50
FNM	-15.39	SEFV	10	4.64	HS	6	68.38	1937.7
FRE	-1.14	COMM	0	-14.11	SEFV	5	90.59	2433
LEH	-7.37	FB	4	-5.29	MISS	2	102.3	2477.7
LM	-2.62	MEAT	7	-12.20	APL	13	12.46	1558
MCO	3.28	DUR	9	-9.26	RPR	9	72.78	1237.5
MS	-0.42	TPU	0	-0.95	DIAR	4	12.95	235.2

December 2008

Company	b_1	CPI_1	τ_1	b_2	CPI_2	τ_2	c	d
AIG	-22.34	DIAR	9	-173.7	SEFV	1	1169	32260
C	-3.74	FB	9	-3.76	RPR	13	44.39	1287.6
CIT	2.60	NC	12	-9.66	RPR	13	63.46	1375
FITB	-3.48	F	7	-0.93	LS	11	23.20	781.5
FNM	-5.67	F	8	-2.28	TS	0	35.61	1436
FRE	-2.21	TS	0	-8.40	RPR	13	62.87	1976.6
LEH	-5.21	F	5	-4.97	PETS	0	59.65	1323
LM	-7.27	F	5	-8.31	APL	13	39.48	1967.8
MCO	2.98	DUR	9	-9.70	RPR	10	74.58	1350
MS	-12.55	SEFV	3	2.83	HS	10	74.88	1589

October 2009

Company	b_1	CPI_1	τ_1	b_2	CPI_2	τ_2	c	d
AIG	-22.79	DIAR	9	-156.05	SEFV	0	1066	29580
C	-0.59	DIAR	4	-5.88	SEFV	5	38.84	1054.8
CIT	4.92	HFO	10	-9.37	RPR	12	62.02	1058
FITB	-4.99	SEFV	2	1.54	HS	6	21.86	630.7
FNM	-15.39	SEFV	10	4.64	HS	6	68.38	1837.7
FRE	-1.13	COMM	0	-14.12	SEFV	5	90.59	2433.8
LEH	-7.39	FB	4	-5.29	MISS	2	102.3	2477.7
LM	-5.82	FB	4	-8.18	APL	13	32.36	1722
MCO	-12.98	RPR	10	3.19	MISG	8	97.83	1981
MS	5.16	HO	10	-9.61	SEFV	3	39.62	1017

Table 4 lists 10 models with predicted negative or very close to negative prices as obtained in May, September and December 2008 as well as in October 2009. Figure 3 displays corresponding predicted and observed curves between July 2003 and December 2009. American International Group has a very stable model for the entire period as defined by the *DIAR* and *SEFV*. Theoretically, the company should suffer a rapid drop in share price from ~$1400 to the level of about -$300. In reality, this fall was stopped by a bailout with the share price hovering between $10 and $50 by the end of 2008 and through 2009. According to all four models the price should start growing in 2010. It will be an important test for our pricing concept.

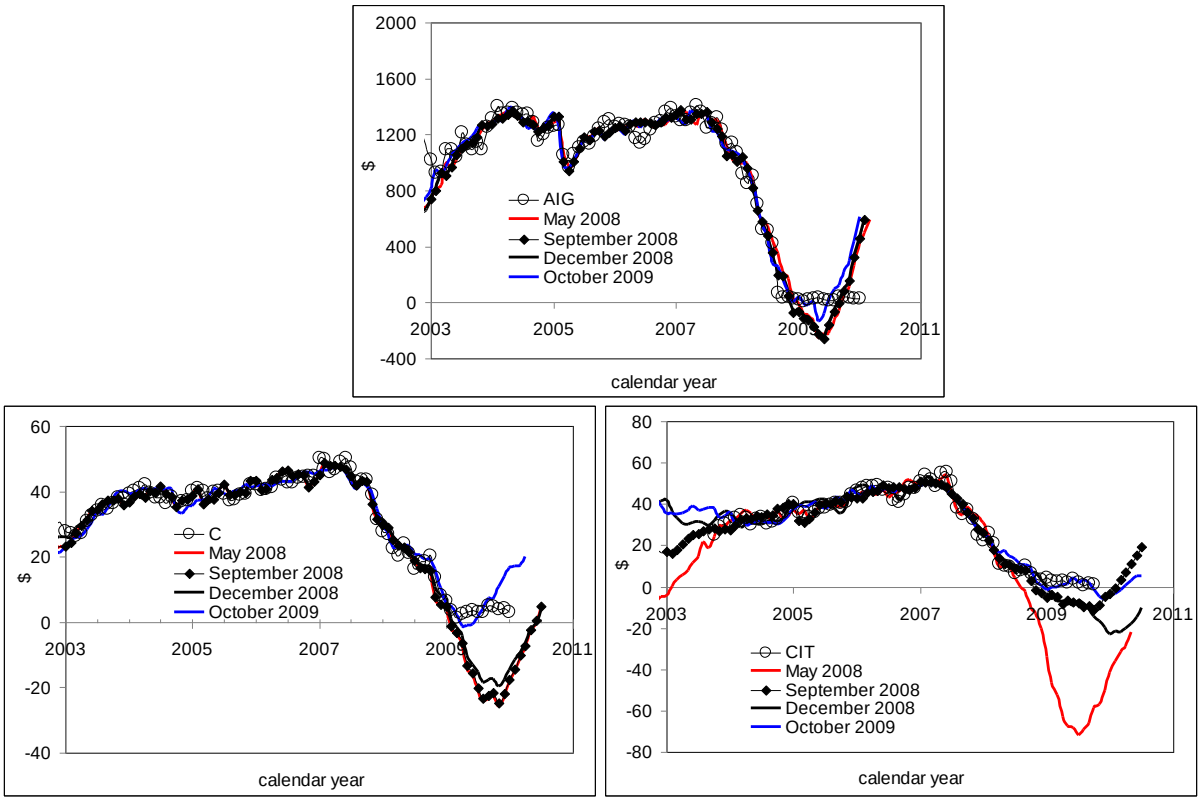
For Citigroup, the models obtained in 2008 are similar and are based on the indices of food and rent of primary residence. Figure 3 demonstrate that negative prices were expected in the end of 2008. All three models predicted the bottom price at -$30. In October 2009, the defining CPI components are different as the model tries to describe the price near $2.

The history of CIT Group (CIT) includes two attempts of bailout and a bankruptcy in November 2009 with a total debt of $10 billion. In Figure 3, the May 2008 model predicts a very deep fall in the share price. Other two models in 2008 demonstrate just a modest fall below the zero line. The bailouts have likely biased the October 2009 model and it predicts the company to recover in 2010. It would be a good exercise similar to that for the AIG model. Unfortunately, the history of CIT Group has ended with a bankruptcy, as expected.

Fanny Mae and Freddie Mac were both bailed out in September 2008. As depicts Figure 3, the models between May and December 2008 are all different. However, all of them predicted negative prices. The models for FNM imply the bottom price level of -$50 to -$60 and the pivot point somewhere in 2009. The models for FRE do predict negative prices with the bottom at -$30, but only the September model has a pivot point.

Lehman Brothers was one of the first giant companies to file for bankruptcy protection in September 2008. The May 2009 model does predict negative prices in the beginning of 2009. The September and December 2009 models are likely biased by the bankruptcy but both indicate a deep fall in the price. It is important to stress that the bottom price for LEH was predicted at -$20 with a quick return into the positive zone. Therefore, the risk might be overestimated.

The models predicted for FITB, LM, MCO and MS are presented to emphasize the problem of resolution and selection of a valid model. For these four companies there is at least one model predicting negative or very close to zero prices. In reality, no one of them has touched the zero line. Moreover, they have not been falling since the end of 2008. So, in order to obtain an accurate prediction one should the best resolution, which might be guaranteed by the higher possible dynamic range. The 2008 crisis and the following recovery allowed the biggest change in the S&P share prices. Hence, the models obtained in 2010 have to be the most resolved and thus the most reliable. Good news is that these models will be valid in the future, but with different coefficients (Kitov 2010).



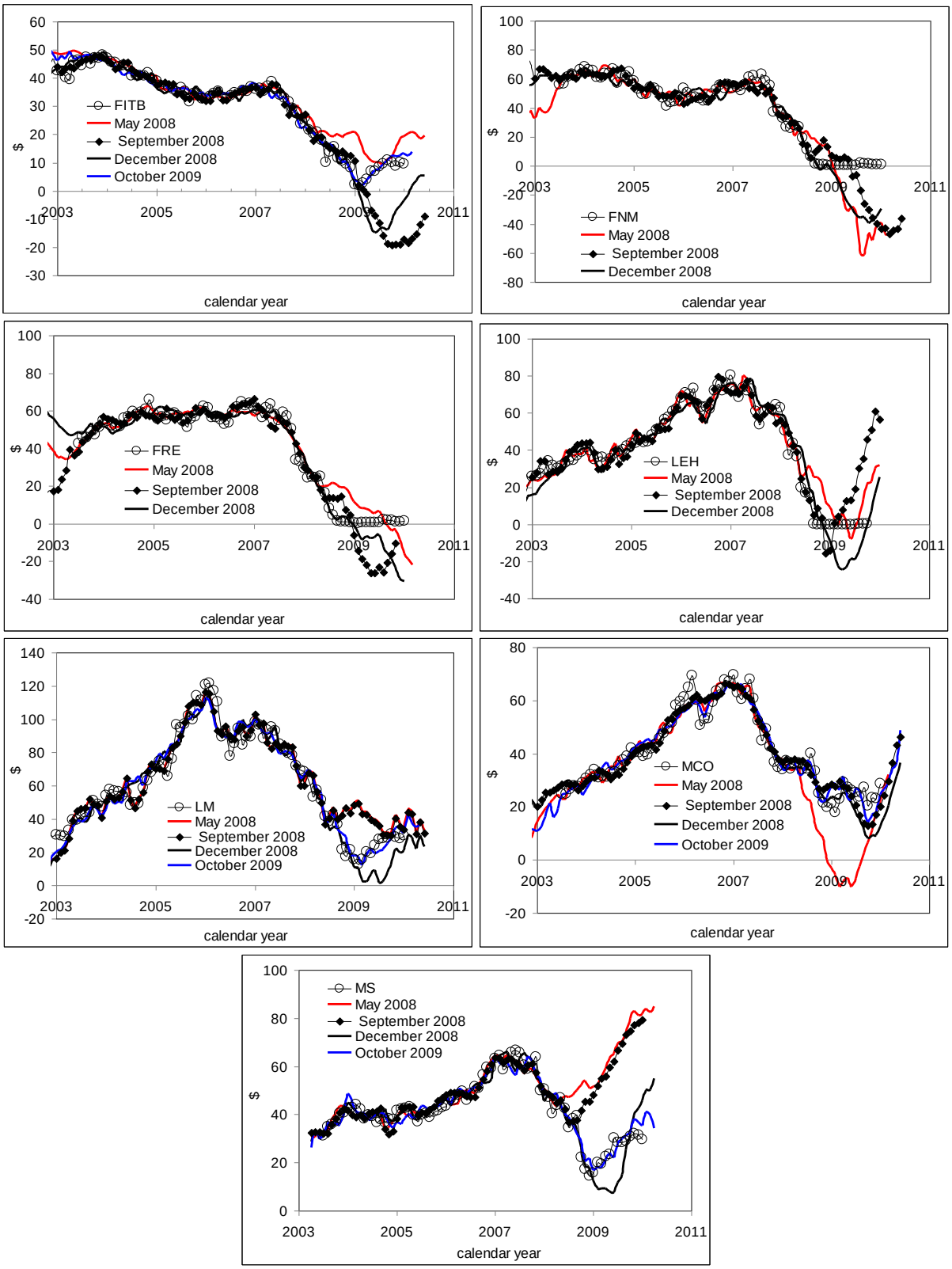


Figure 3. Comparison of stock prices for several financial companies as predicted in May, September and December 2008, and October 2009

There are six companies, all with predicted negative prices but different fate. We have a question on relative merits of the previous bank bailouts – which bank did deserve a bailout and how much would it really cost? The models in Table 4, although they are only tentative ones and should be used with all necessary precautions, might provide a measure of debt size. One can estimate the debt as a product of the number of shares and relevant market price, which was negative for the bailed out and not bailed out companies. Table 5

lists the estimated debts. Lehman Brothers had a much smaller debt than that of Citigroup, CIT and AIG. So, it would have been much easier to bail out LEH from the mathematical point of view. Also, the joint debt of AIG, FRE and FNM is less than $200 billion.

So, we have answered all questions formulated in the beginning of this Section. When having valid pricing models for the companies under consideration, one could foresee all problems before they become serious and select appropriate measures including bailouts. Moreover, taking into account the deterministic evolution of the CPI and linear trends in the CPI differences (Kitov, and Kitov 2008), one could predict major problems long before they happen and avoid most of the 2008/2009 turmoil. For this, financial companies should learn the CPI components defining the evolution of their stocks.

Table 5. Total debt as calculated from negative share prices.

Company	## Shares	Share price, $	Debt, $
LEH	$6.89 \cdot 10^8$	-20	$1.4 \cdot 10^{10}$
C	$1.1 \cdot 10^9$	-30	$3.3 \cdot 10^{11}$
CIT	$8.12 \cdot 10^9$	-20	$1.6 \cdot 10^{11}$
AIG	$1.34 \cdot 10^8$	-360	$1.0 \cdot 10^{11}$
FRE	$6.8 \cdot 10^8$	-40	$2.6 \cdot 10^{10}$
FNM	$1.11 \cdot 10^9$	-50	$5.5 \cdot 10^{10}$

6. Discussion

A deterministic model has been developed for the prediction of stock prices at a horizon of several months. The model links the shares of traded companies to consumer price indices. In this paper, we presented empirical models for financial companies from the S&P 500 list. In May 2008, the model predicted negative share prices in the second half of 2008 for Lehman Brothers, American International Group, Freddie Mac. With known defining CPI components one could predict the approaching bankruptcies. This makes of crucial importance the estimation of correct empirical models, i.e. defining CPIs, for all shares. When reversed, the model also makes it is possible to predict the evolution of various CPI subcategories.

Despite its apparent opposition to the mainstream concepts, the pricing model is deeply rooted in economics: a higher pricing power achieved by a given company should be converted into a faster growth in corresponding consumer price index. This link works excellent for many S&P 500 companies. A further improvement in the model's predictive power is likely possible using advanced methods of statistical and econometrical analysis. However, one should bear in mind that the model will work until its influence on the market is negligible. When a good portion of market participants uses the model it should fail because the market functioning will be disturbed.

Observed and predicted share prices are measured variables and the link between them is likely of a causal character during the studied period. Therefore, the mainstream stock pricing models are, in part, valid – when the evolution of the driving force is random the price is also random, but predictable.

An important possibility arises from our analysis. Using different subsets of the CPI, one can improve our tentative models for the studied companies, and easily obtain similar quantitative relationships for other companies. By extrapolating previously observed trends into the future, one may forecast share prices at various horizons. What likely is more important for a broader investor community, the proposed model also allows predicting the turning points between adjacent trends, when share prices are subject to a substantial decline.

The presented results are preliminary ones and do not pretend to provide an optimal price prediction. A comprehensive investigation with smaller components of the CPI will likely give superior results. So, we recommend refining the model in order to obtain accurate quantitative results for actual investment strategies. All in all, the lagged differences between two CPI components provide a good approximation for the evolution of many stock prices.

One may pose a question: Why did the researches in economics and finances fail to derive the model many years ago? The answer is a scientific one. There were no appropriate data. First, the partition of the headline CPI in hundreds of components is a very new development. Moreover, this process is ongoing and a researcher obtains a more adequate set of defining variables. This brings both higher resolution and reliability. Second, the reliability critically depends on the dynamic range of data. The crisis of 2008 and 2009 has resulted

in a dramatic change in both share prices and CPI components. The increased resolution and dynamic range allowed deriving a sound quantitative model. There was no chance to find the link between the share prices and CPI before the data allow. This is a general consideration applicable to all economic and financial models – adequate data must come first (Kitov 2009a).

References

[1] Bureau of Labor Statistic. 2010. Consumer price index. Table, retrieved 01.02.2010 from http://www.bls.gov/data/.

[2] Granger, C., and Newbold, P. 1967. Spurious regression in econometrics. *Journal of Econometrics*, vol. 2:111-120.

[3] Hendry, D., and Juselius, K. 2001. Explaining Cointegration Analysis: Part II. *Energy Journal*, vol. 22: 75-120.

[4] Johansen, S. 1988. Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control*, vol. 12: 231-254.

[5] Kitov, I. 2009a. Does economics need a scientific revolution?, *MPRA Paper* 14476, University Library of Munich, Germany.

[6] Kitov, I. 2009b. Predicting ConocoPhillips and Exxon Mobil stock price, *Journal of Applied Research in Finance,* vol., Issue 2(2), Winter 2009:129-134.

[7] Kitov, I. 2010. *Deterministic mechanics of pricing*. Saarbrucken, Germany: LAP LAMBERT Academic Publishing.

[8] Kitov, I., and Kitov, O. 2008. Long-Term Linear Trends In Consumer Price Indices, *Journal of Applied Economic Sciences,* vol. III, Issue 2(4), Summer: 101-112.

[9] Kitov, I., and Kitov, O. 2009a. Sustainable trends in producer price indices, *Journal of Applied Research in Finance*, vol. I, Issue 1(1), Summer: 43-51.

[10] Kitov, I., and Kitov, O. 2009b. A fair price for motor fuel in the United States, *MPRA Paper* 15039, University Library of Munich, Germany.

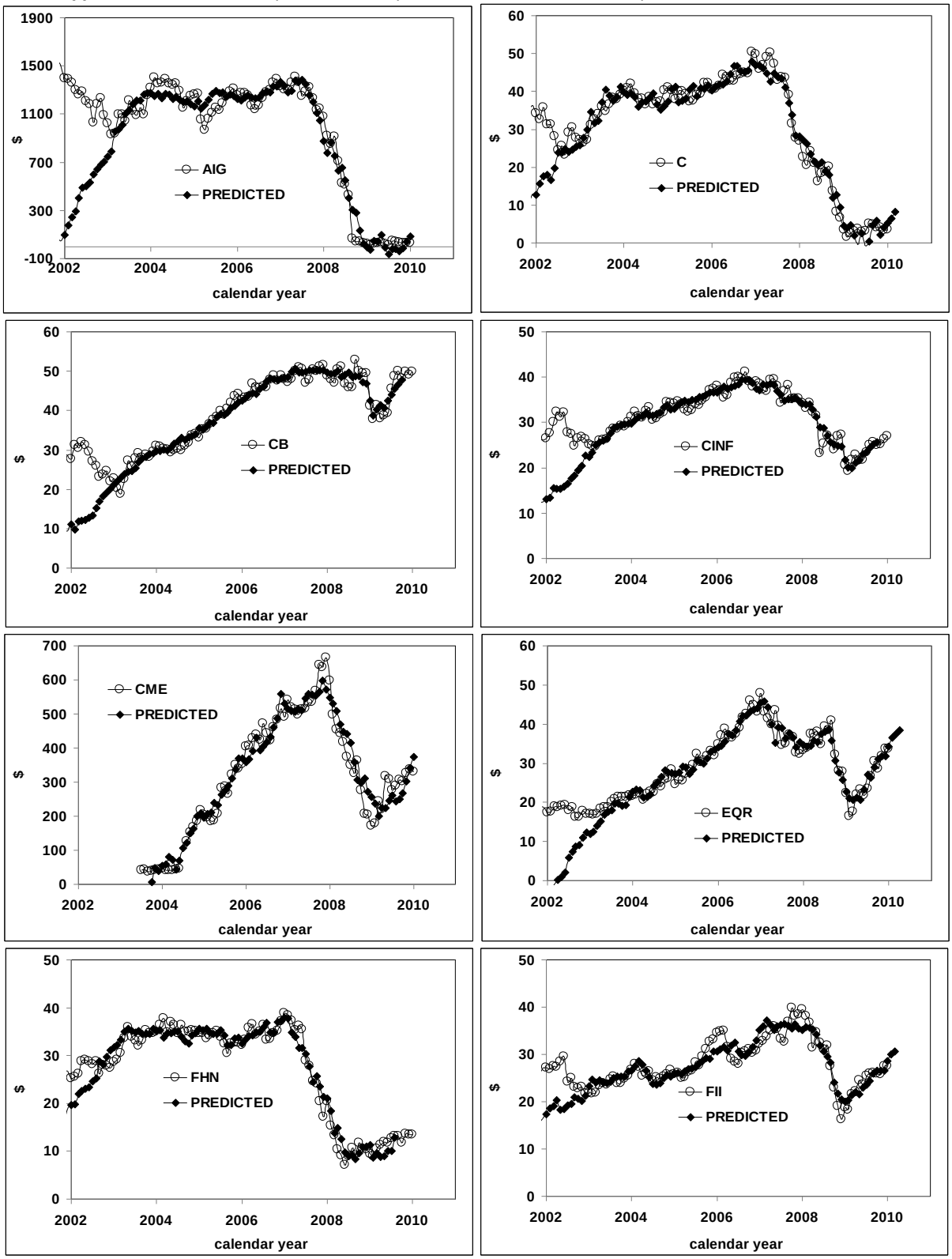
Appendix 1. List of seventy CPI components used in the study; in alphabetic order

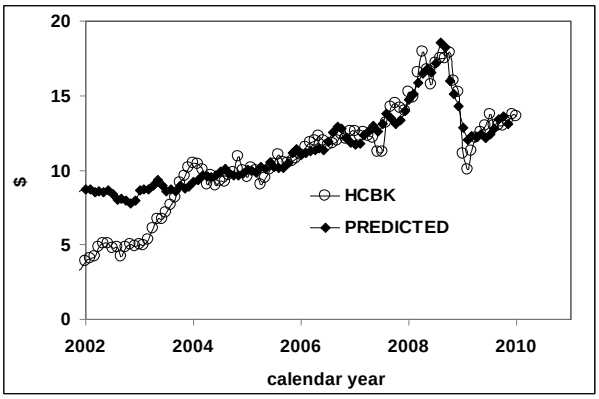
Acronym	Description	Acronym	Description
A	apparel	MAP	men's and boy's apparel
AB	alcoholic beverages	MCC	medical care commodities
APL	appliances	MCS	medical care services
C	CPI	MEAT	meats, poultry, and fish
CC	core CPI	MF	motor fuel
CE	CPI less energy	MISG	miscelleneous goods
CF	CPI less food	MISS	miscellenous services
CFSH	CPI less food and shelter	MVI	motor vehicle insurance
CFSHE	CPI less food shelter and energy	MVP	motor vehicle parts
CM	CPI less medcare	MVR	motor vehicle repaire
CO	communication	NC	new cars
COMM	commodities	NDUR	nondurables
CSH	CPI less shelter	O	other goods and services
DIAR	diary products	ORG	other recreation goods
DUR	durables	OS	other services
E	energy	PC	personal care
EC	education and communication	PDRUG	prescription drugs
ED	education	PETS	pets and related goods
F	food and beverages	R	recreation
FB	food less beverages	RENT	rent
FISH	fish	RPR	rent primary residence
FOOT	footware	RRM	recreational reading materials
FOTO	photography	RS	recreation services
FRUI	fruits and vegetables	SEFV	food away from home
FS	financial services	SERV	services
FU	fuels and utilities (housing)	SH	shelter
H	housing	SPO	sporting goods (apparel)
HFO	household furnishing and operations	T	transportation
HO	household operations	TOB	tobacco
HOSP	hospital services	TPR	private transportation
HS	housekeeping supplies	TPU	public transportation
ITR	intracity transportation	TS	transportation services
JEW	jewelry and watches	TUIT	tuition
LS	legal services	VAA	video and audio
M	medical care	WAP	women's and girl's apparel

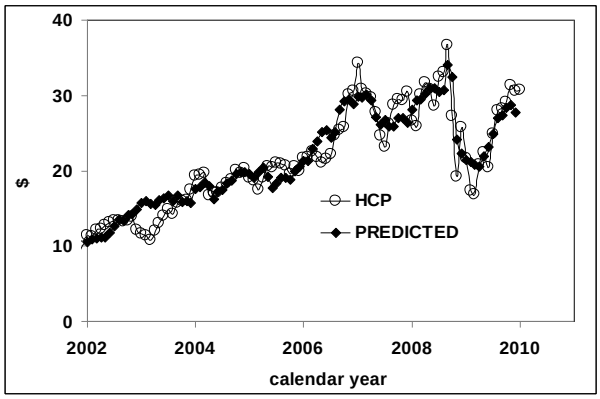
Appendix 2. Empirical 2-C models for S&P 500 financial companies

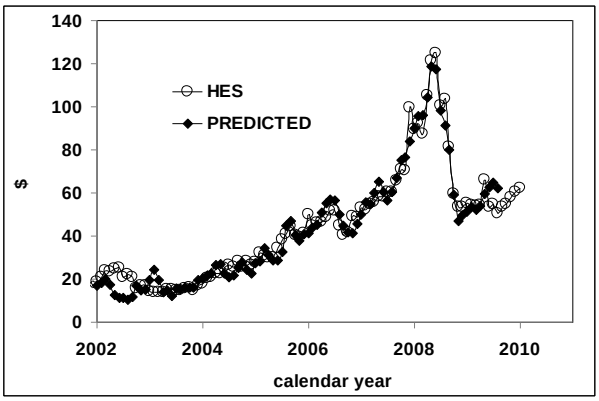
Company	b_1	CPI_1	τ_1	b_2	CPI_2	τ_2	c	d	σ, $
AIG	-191.36	SEFV	1	38.53	PDRUG	13	727.81	21166.78	92.31
C	2.94	HO	5	-8.26	SEFV	2	36.70	1048.90	2.53
CB	-1.27	F	3	-0.41	O	**-3**	16.53	313.52	1.70
CINF	-0.06	TOB	**-3**	-2.79	SEFV	**-4**	21.53	486.92	1.41
CME	-38.62	PETS	0	-30.86	AB	10	458.48	8475.35	41.43
EQR	-3.16	SEFV	3	1.12	PDRUG	4	11.48	187.90	2.14
FHN	1.52	MCC	8	-4.10	SEFV	**-5**	14.30	340.70	1.90
FII	1.49	HO	12	-1.78	PETS	2	4.02	38.95	2.01
HCBK	0.03	E	1	0.47	MISS	**-2**	-4.43	-113.67	0.83
HCP	1.31	MCC	5	0.30	FS	**-1**	-8.35	-365.78	2.10
HES	3.37	CSH	**-2**	3.81	MISS	**-5**	-47.75	-1480.90	4.87
HIG	0.53	TPU	12	-8.35	PETS	1	44.41	741.76	4.79
HST	-1.30	FB	4	-1.39	RPR	11	18.91	451.01	1.12
IVZ	2.44	HO	11	-1.52	PETS	3	-1.04	-99.66	1.79
JNS	-2.99	PETS	2	3.14	RPR	7	-4.93	-257.86	2.63
KEY	-0.36	DIAR	9	-3.92	RPR	11	28.68	763.73	1.59
KIM	3.15	RS	**-3**	-5.14	SEFV	2	25.07	454.58	2.17
LNC	-4.59	F	5	-2.35	TS	3	41.28	1212.45	3.80
LUK	0.65	TPR	**-2**	-1.42	MVI	3	7.66	332.12	3.13
MCO	-0.95	MEAT	4	-8.58	RPR	9	69.27	1664.39	3.90
MET	0.36	TPU	13	-4.41	PETS	2	26.48	364.93	2.88
MI	-2.54	SEFV	4	-2.44	RPR	13	32.35	850.83	1.64
MS	5.27	HO	8	-9.75	SEFV	2	39.55	1031.57	3.72
MTB	-3.96	FB	3	-4.65	RPR	11	57.08	1510.23	4.87
NTRS	-3.02	PETS	2	3.66	RPR	5	-5.16	-340.09	3.94
PBCT	0.65	MCC	7	-0.75	MVP	13	0.12	-82.64	0.77
PCL	-0.81	MCC	**-1**	0.57	FS	**-2**	5.27	79.01	2.02
PFG	-3.01	PETS	**-2**	0.97	FS	**-2**	14.95	85.06	3.55
PGR	0.13	FU	**-2**	-1.74	RPR	2	11.47	309.11	1.29
PLD	-3.09	PETS	**-1**	1.17	FS	**-2**	14.18	42.98	3.44
PRU	-8.13	PETS	2	0.18	TOB	0	45.25	723.38	5.44
RF	-1.29	F	**-3**	-2.04	FB	7	18.09	554.44	1.48
SLM	1.91	PETS	13	-9.31	RPR	12	53.36	1490.40	4.12
SPG	-5.78	F	3	0.98	FS	**-1**	38.82	693.54	5.36
STI	-5.70	FB	4	-0.16	TOB	5	37.05	1010.63	3.65
STT	5.12	HO	11	-4.61	PETS	4	3.78	-50.41	5.02
TMK	-8.45	SEFV	3	0.77	HOSP	2	36.15	1140.20	3.24
TROW	-1.54	FB	3	-2.09	TS	5	26.41	644.30	3.51
UNM	-0.17	FU	6	-0.91	PETS	2	8.82	110.67	1.56
USB	-1.10	FB	4	0.32	FS	0	6.26	117.85	1.67
VTR	0.45	FS	**-2**	-1.60	MISS	8	18.93	276.81	2.41
WFC	-1.15	O	**-3**	0.09	TOB	0	10.72	281.54	2.06
XL	-13.65	RPR	13	6.35	MVR	13	45.24	1413.75	4.37
ZION	-2.19	F	4	-8.09	RPR	13	67.14	1812.98	3.21

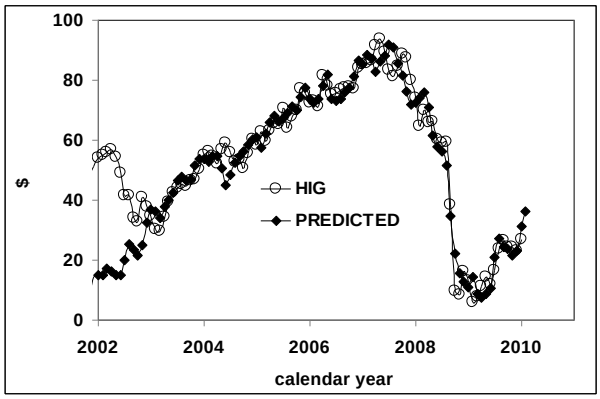
Appendix 3. Observed and predicted stock prices of S&P 500 financial companies

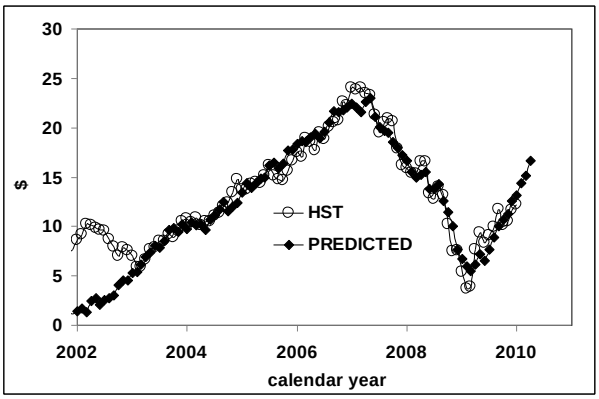


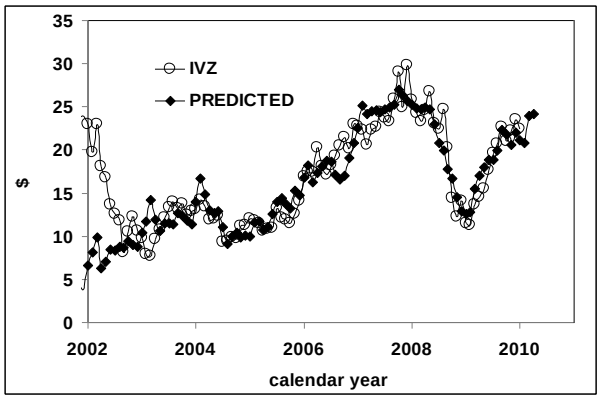

20
15
10
5
0
$
HCBK
PREDICTED
2002
2004
2006
2008
2010
calendar year

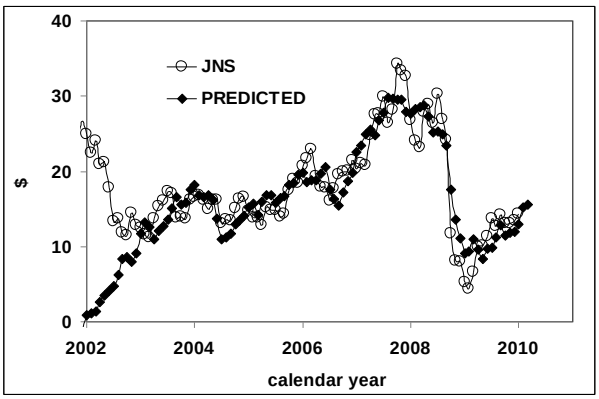

40
30
20
10
0
$
HCP
PREDICTED
2002
2004
2006
2008
2010
calendar year

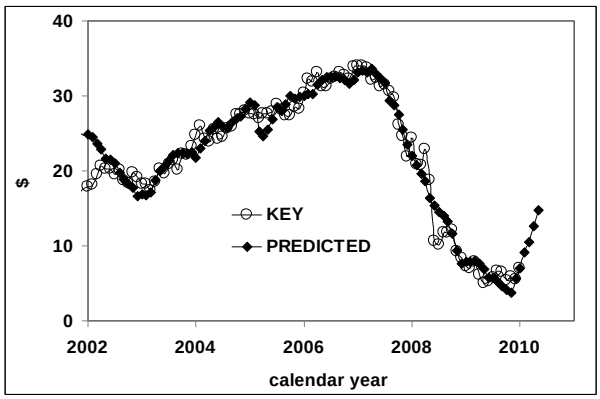

140
120
100
80
60
40
20
0
$
HES
PREDICTED
2002
2004
2006
2008
2010
calendar year

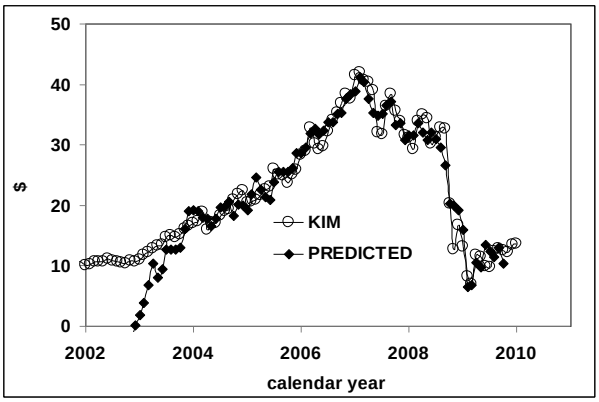

100
80
60
40
20
0
$
HIG
PREDICTED
2002
2004
2006
2008
2010
calendar year

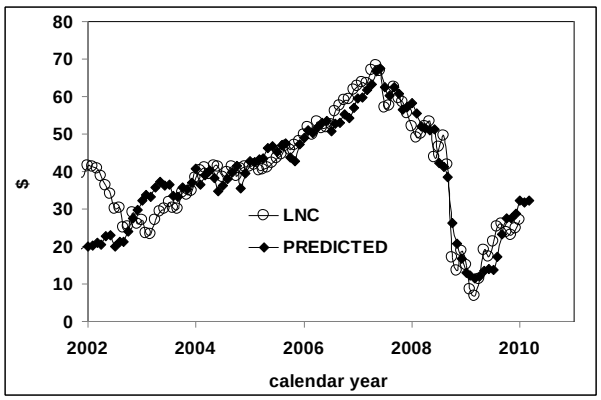

30
25
20
15
10
5
0
$
HST
PREDICTED
2002
2004
2006
2008
2010
calendar year

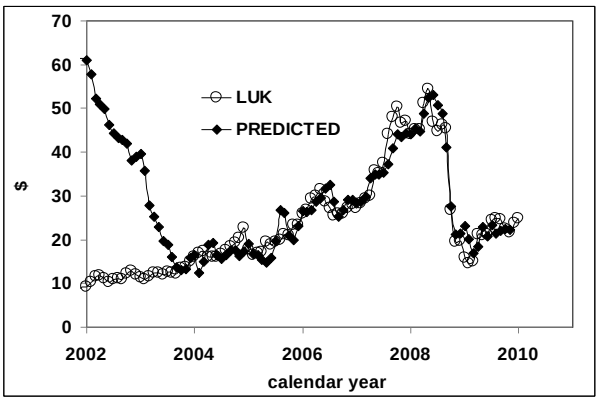

35
30
25
20
15
10
5
0
$
IVZ
PREDICTED
2002
2004
2006
2008
2010
calendar year

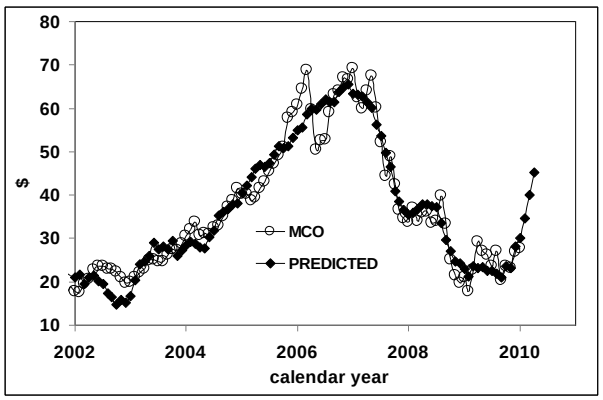

40
30
20
10
0
$
JNS
PREDICTED
2002
2004
2006
2008
2010
calendar year

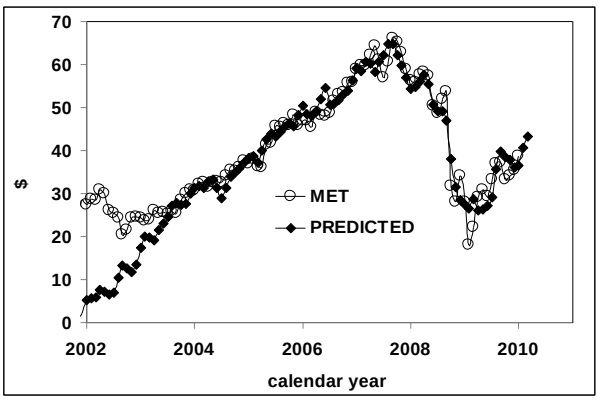

40
30
20
10
0
$
KEY
PREDICTED
2002
2004
2006
2008
2010
calendar year

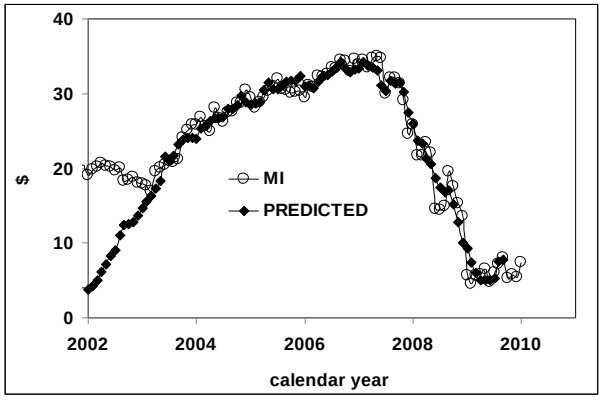

$
KIM
PREDICTED
calendar year

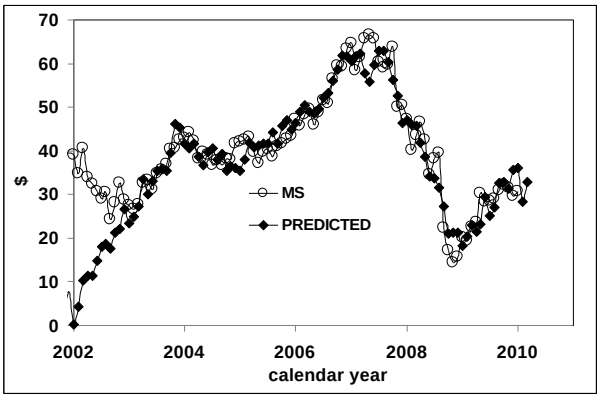

$
LNC
PREDICTED
calendar year

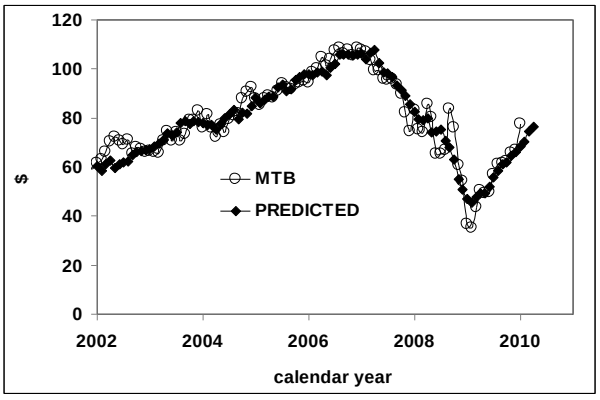

$
LUK
PREDICTED
calendar year

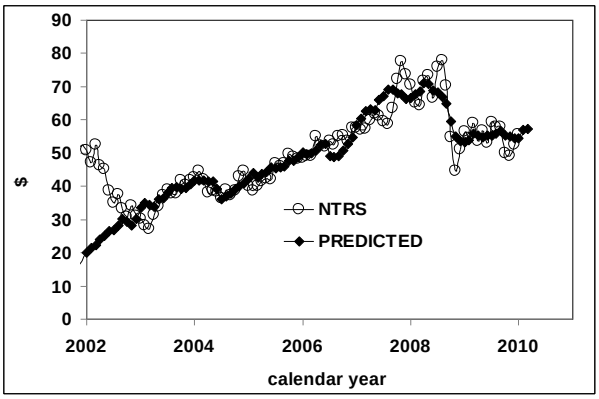

$
MCO
PREDICTED
calendar year

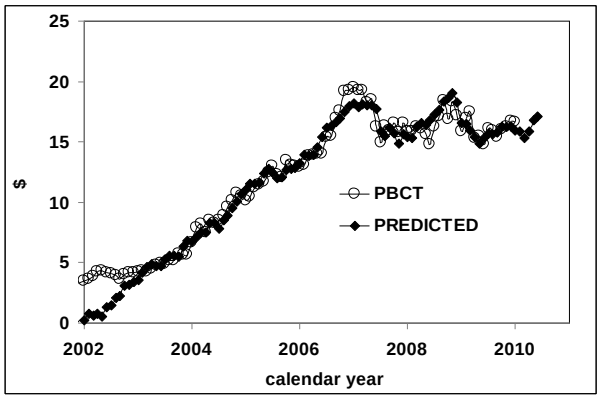

$
MET
PREDICTED
calendar year

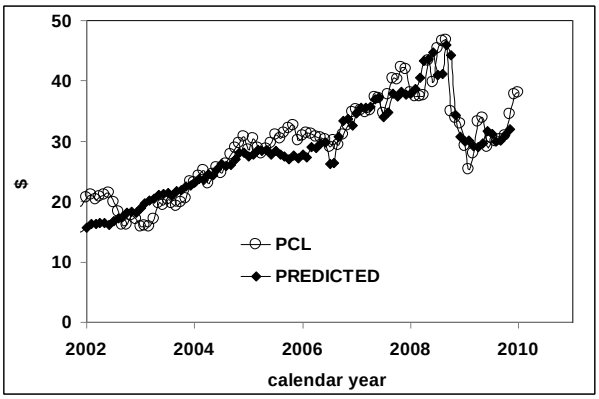

$
MI
PREDICTED
calendar year

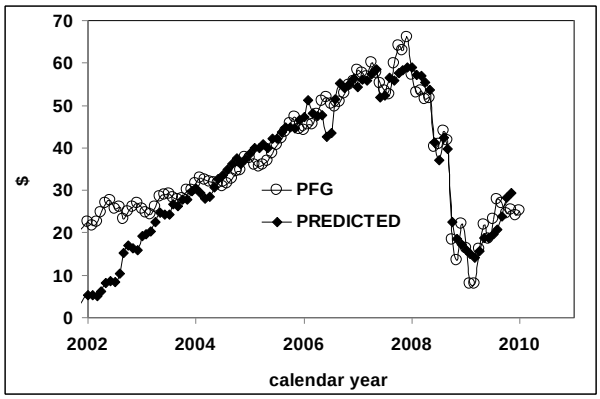

$
MS
PREDICTED
calendar year

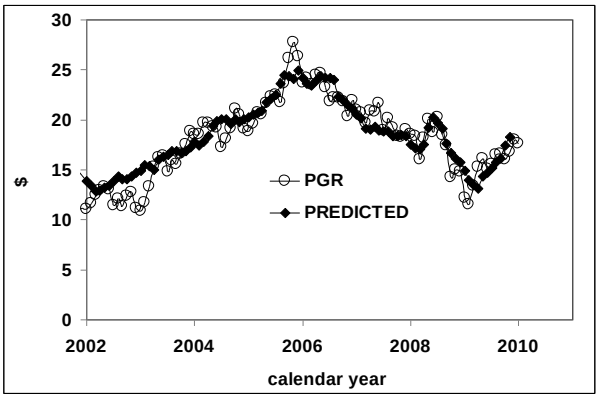

$
MTB
PREDICTED
calendar year

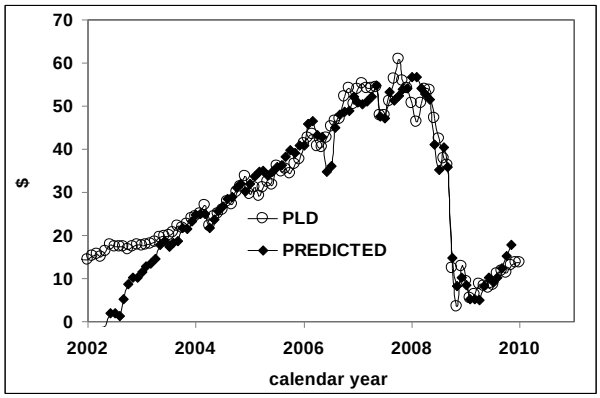

$
NTRS
PREDICTED
calendar year

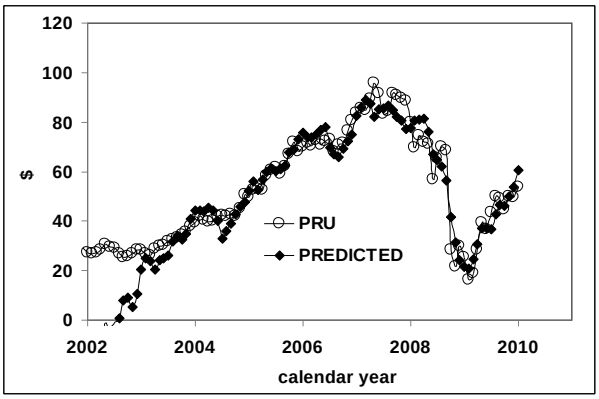

$
PBCT
PREDICTED
calendar year

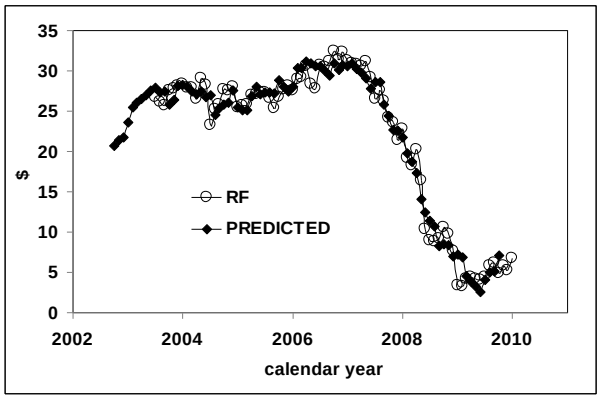

$
PCL
PREDICTED
calendar year

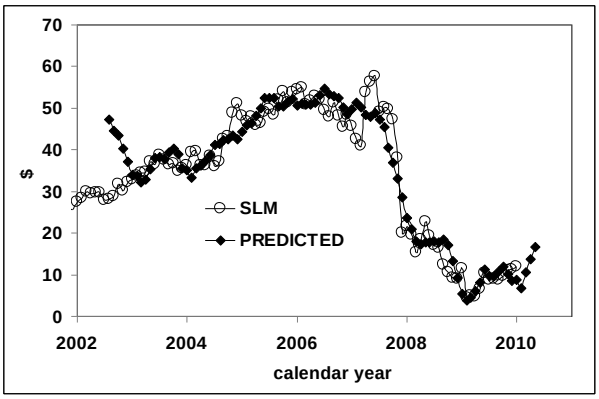

$
PFG
PREDICTED
calendar year

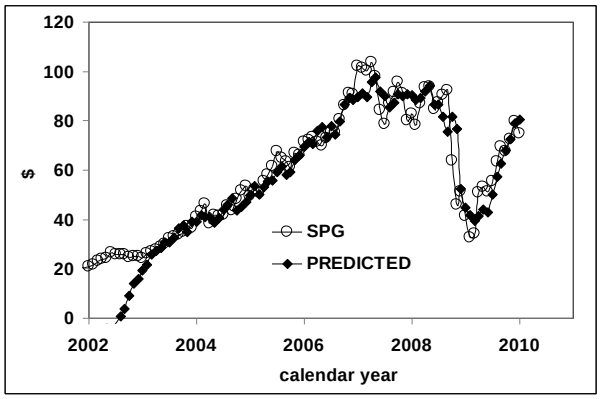

$
PGR
PREDICTED
calendar year

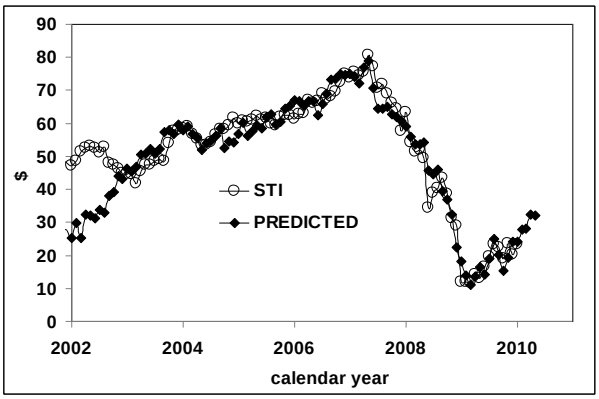

$
PLD
PREDICTED
calendar year

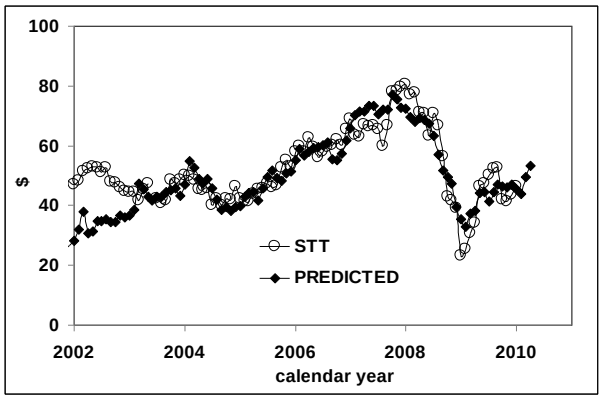

$
PRU
PREDICTED
calendar year

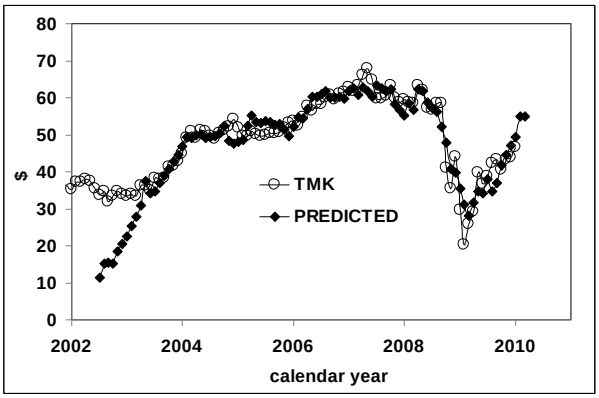

$
RF
PREDICTED
calendar year

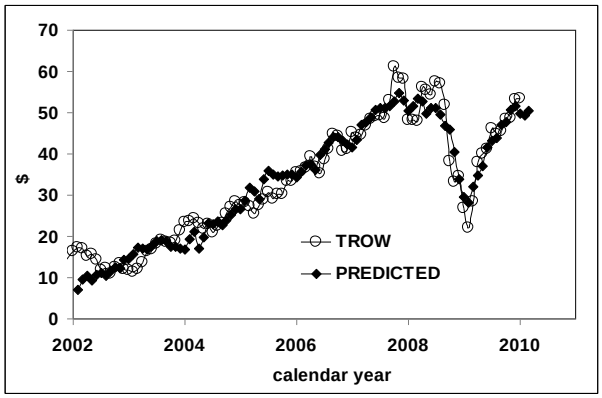

70
60
50
40
30
20
10
0
$
SLM
PREDICTED
2002
2004
2006
2008
2010
calendar year

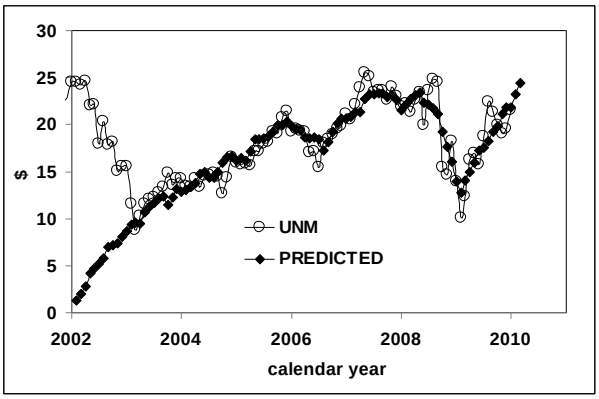

120
100
80
60
40
20
0
$
SPG
PREDICTED
2002
2004
2006
2008
2010
calendar year

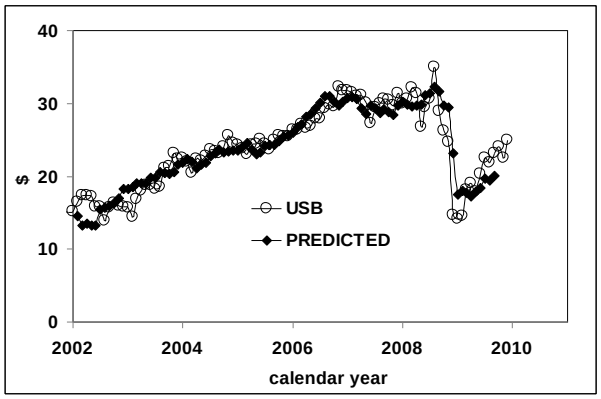

90
80
70
60
50
40
30
20
10
0
$
STI
PREDICTED
2002
2004
2006
2008
2010
calendar year

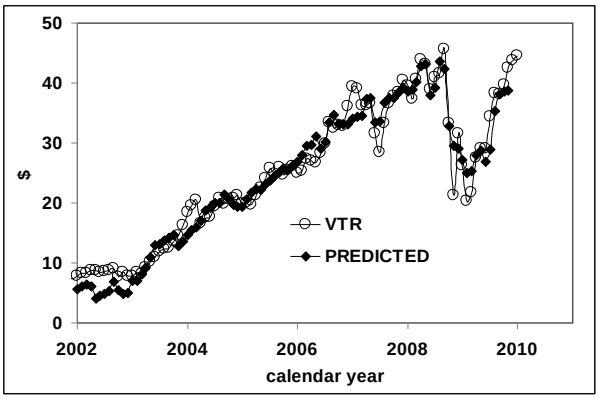

100
80
60
40
20
0
$
STT
PREDICTED
2002
2004
2006
2008
2010
calendar year

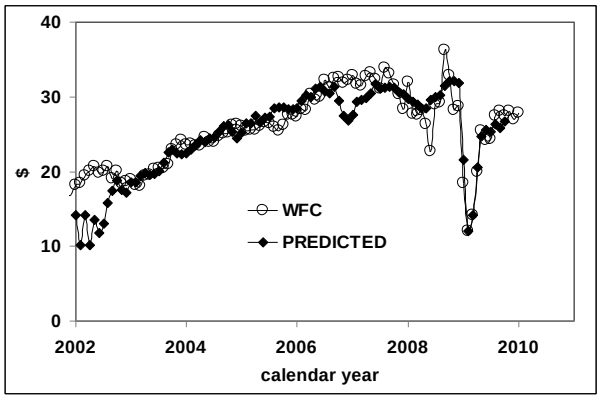

80
70
60
50
40
30
20
10
0
$
TMK
PREDICTED
2002
2004
2006
2008
2010
calendar year

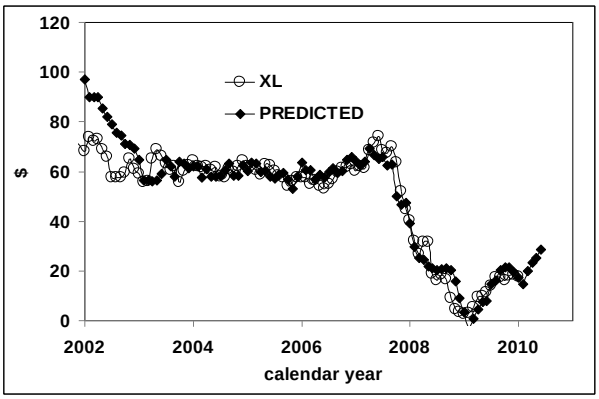

70
60
50
40
30
20
10
0
$
TROW
PREDICTED
2002
2004
2006
2008
2010
calendar year

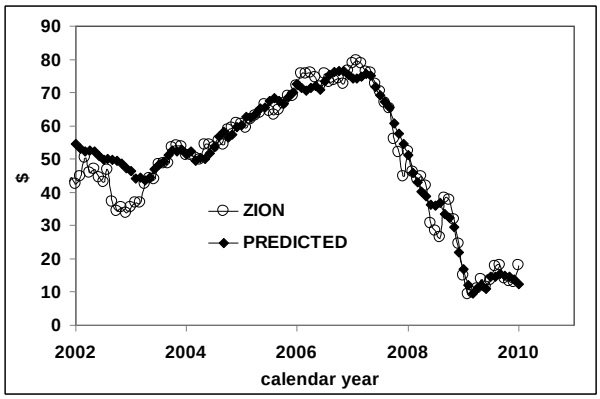

30
25
20
15
10
5
0
$
UNM
PREDICTED
2002
2004
2006
2008
2010
calendar year


40
30
20
10
0
$
USB
PREDICTED
2002
2004
2006
2008
2010
calendar year


50
40
30
20
10
0
$
VTR
PREDICTED
2002
2004
2006
2008
2010
calendar year


40
30
20
10
0
$
WFC
PREDICTED
2002
2004
2006
2008
2010
calendar year


120
100
80
60
40
20
0
$
XL
PREDICTED
2002
2004
2006
2008
2010
calendar year


90
80
70
60
50
40
30
20
10
0
$
ZION
PREDICTED
2002
2004
2006
2008
2010
calendar year

Appendix 4. Empirical models, as of May 2008.

Company	b_1	CPI_1	τ_1	b_2	CPI_2	τ_2	c	d
CB	-0.39	FISH	8	-2.48	RS	13	17.32	335.17
CINF	-2.09	F	8	-3.41	PC	1	27.78	906.09
CME	6.79	TOB	6	-169.02	SEFV	6	1014.46	23534.41
EQR	3.33	MVP	5	-5.53	RPR	6	31.91	670.54
FHN	-6.62	**SEFV**	4	1.02	PDRUG	8	25.99	810.51
FII	2.27	DUR	10	-1.54	HS	1	11.38	-48.93
HCBK	0.07	TOB	13	-0.25	PDRUG	9	3.12	45.50
HCP	4.91	R	5	0.94	OS	3	-9.23	-713.62
HES	-7.91	HO	5	17.05	SEFV	9	-44.26	-1925.87
HIG	-5.74	RPR	7	2.65	PDRUG	13	17.33	307.40
IVZ	-0.46	MEAT	7	-0.39	MVI	4	8.04	181.92
JNS	1.14	DUR	11	-2.97	FOTO	5	-4.86	183.10
KEY	-3.14	**RPR**	6	0.98	PDRUG	11	11.84	304.61
KIM	0.98	PDRUG	6	2.00	FOTO	11	0.13	-496.62
LNC	3.12	MVP	4	-10.95	SEFV	4	61.51	1455.18
LUK	-4.26	VAA	8	1.74	OS	13	-5.82	46.26
MCO	3.69	DUR	9	-9.51	**RPR**	8	75.41	1232.98
MET	-1.56	SH	11	1.13	PDRUG	12	4.74	-22.58
MI	-0.86	FB	3	-5.19	**SEFV**	4	37.48	979.40
MTB	-7.20	**RPR**	11	-3.57	MISS	12	87.49	2180.14
NTRS	-1.10	MEAT	8	0.45	DIAR	2	11.12	102.13
PBCT	-0.73	O	3	0.38	PDRUG	7	4.68	81.62
PCL	-0.22	FU	6	-0.16	TPU	0	7.81	59.81
PFG	1.57	APL	3	6.50	FOTO	13	25.34	-840.96
PGR	0.40	MVI	9	1.45	RRM	12	-5.52	-368.27
PLD	-2.16	FB	4	-1.02	**FS**	13	26.75	547.08
PRU	2.00	PDRUG	13	-1.65	HOSP	4	31.56	-71.92
RF	-3.31	RPR	13	-1.89	RRM	3	25.57	967.64
SLM	4.67	VAA	12	-6.52	**RPR**	13	41.17	706.86
SPG	15.04	R	5	-5.63	PETS	3	24.43	-1021.99
STI	-8.07	SEFV	0	1.57	PDRUG	11	35.94	893.58
STT	-3.79	SERV	11	5.71	**HO**	11	8.40	131.84
TMK	1.22	MCC	8	0.59	ITR	8	-9.63	-348.07
TROW	1.54	EC	0	0.75	PDRUG	13	-4.58	-370.49
UNM	-1.09	FB	4	-0.84	MCC	4	14.36	374.29
USB	-0.04	E	1	-1.23	MISS	12	15.68	307.26
VTR	1.04	MCC	4	0.85	**FS**	10	-7.35	-430.76
WFC	-1.62	F	13	-1.38	FB	0	18.21	491.21
XL	-2.86	O	12	2.95	PDRUG	11	-14.86	19.28
ZION	-3.22	AB	3	-7.99	**RPR**	13	70.08	1989.05

INFRASTRUCTURES AND ECONOMIC PERFORMANCE: A CRITICAL COMPARISON ACROSS FOUR APPROACHES

Gianpiero TORRISI
Newcastle University, **United Kingdom**
Gianpiero.Torrisi@ncl.ac.uk

Abstract

The paper reviews studies analysing the relationship between infrastructures and economic performance. Four different approaches are separately considered along an ideal path from three theory-based to a data-oriented models: the production function approach, the cost function approach, growth-models, and vector autoregression models. The review shows that, even with different shades and points of caution, the general idea that infrastructure has an economic enhancing effect appears to be quite robust across studies belonging to different methodological approaches.

Keywords: economic development, growth, public expenditure, public infrastructure

JEL Classification: H54, H72, O11

1. Introduction

The link between infrastructure and economic performance has been widely explored in literature utilising many different theoretical approaches and achieving also variegate results and implications for policy-makers.

To begin with, there are theoretical arguments developed in order to sustain thesis based on infrastructures' contribution to productivity considering them as initiating factor.

The first approach to address this issue consisted in considering (public) infrastructures as a free input provided by government (Aschauer 1989). This input, like all other inputs, contributes to the productive process; yet, it has the peculiar characteristic of being a public good in the proper economic sense, so that, once produced represents a structural input common to all firms' production function. From a different angle, infrastructures rather than be considered directly as inputs, could be considered as cost-saving factors.

The underlying idea is that infrastructures, providing a more favourable setting for the development of economic activity, indirectly (and positively) affect the productive process by reducing or allowing to combine more efficiently other factors (e.g. labour and capital). Thus, according to this approach, infrastructures' effects have to be analysed via the cost function, and the expected result is in favour to a reduction of production costs.

Put differently, a territory well-endowed with infrastructures increase productivity because it provides firms with a more favourable cost structure and making accessible more efficient combinations of inputs.

A more general approach consists in considering infrastructures as components of capital as a whole and, in turn, capital formation is considered as the key factor to the growth process.

According to this approach capital has to be intended in a broader sense comprising its traditional meaning (physical capital), intangible "human capital", "knowledge capital", and just infrastructures. Therefore, infrastructures contributing to capital formation belong to the key endogenous features explaining differences in the economic performance.

Although the massive body of literature developed in this field, there are still points of criticism and debate involving many aspects (e.g. infrastructure definition and measurement, productive and unproductive infrastructures, causality direction and magnitude of their impact, short-run and long-run temporal dimension of their impact).

Inside this puzzle of counter-arguments the strongest point of criticism, considering infrastructures a normal good, extremes the discourse to completely deny the effects of infrastructures on productivity.

On this approach's view the empirical evidence of a positive relationship between infrastructures and productivity has to be read in the sense that the former are just accommodating factors which demand increases as the economic system increases its activity. Hence, in order to deal with the issue free from a preconceived idea, data-oriented approach has been also adopted to analyse the relationships between infrastructures and measures of economic performance.

Models belonging to this approach are often labelled as Vector Autoregressive models (VAR); the peculiar feature of these models consists in explaining a limited numbers of variables (including infrastructures) by their own lags and lags of other variables without imposing no a priori causality among them.

Across studies, generally speaking, the existence of a positive impact going from infrastructure to productivity is confirmed, but the empirical evidence is really composite.

However, analysis's outcome depends both on the empirical setting and on methodology. Moving from this consideration, this paper reviews the four approaches mentioned above following an ideal path from the first-generation partial approaches based on production and cost function(s) to the general VAR approach aiming at emphasize the underlying idea characterising each one.

The structure of the paper is as follows. It starts with reviewing, in Section 2, studies utilising the production function approach; Section 3 reviews studies adopting the cost function approach; Section 4 presents studies implementing growth-models, and Section 5 deals with vector autoregression models. Section 6 presents some concluding remarks.

2. The production-function approach

The production function approach models the amount of output that can be produced for each factor of production, given technological constraints.

The seminal work using this approach to measure the economic impact of infrastructure goes back to Aschauer (1989) that introduced government expenditure intended as a public good into the production function.

Studies following this approach share the same underlying idea that public capital can be considered as an additional *input factor* having the characteristic of a public good in the proper economic sense (i.e. being *not rival* and *not excludable*).

A general form of production function utilised across studies is

$$Y = f(A, L, K, G) \tag{1}$$

Where the variable introduced above are defined as follows: *Y* – Is the level f economic output; *A* – is the "technological progress"; *K* – is the stock of private capital; *L* – is the stock of labour; *G* – represents a measure of public capital.

Clearly, in order to quantify the impact of various input on output we need a specific functional form. In other words, we need to explain what *f* "means".

Usually, an aggregate Cobb-Douglas production function is utilised in empirical works:

$$Y = AL^{\alpha}K^{\beta}G^{\gamma} \tag{2}$$

The most common econometric method to estimate the parameters α, β, γ is the ordinary least squares (OLS) and since one of the basic requirements of OLS method is that the relation must be linear, Equation (2), is often transformed taking natural logarithms of both the left and the right side, obtaining the following

$$\ln Y = \alpha \ln AL + \beta \ln K + \gamma \ln G \tag{3}$$

A further transformation consists in put Equation (3) in per-worker terms and assuming constant returns to scale across all inputs (i.e. assuming that $\alpha + \beta + \gamma = 1$). The result is given by the following (4)

$$\ln \frac{Y}{L} = \ln A + \beta \ln \frac{K}{L} + \gamma \ln \frac{G}{L} \tag{4}$$

According to specification (4) – and (3) - the main research question consists in estimating the parameter " γ " which represents the elasticity of output with respect to infrastructures.

Note that, given the difficulties to model technological progress, Aschauer (1989)'s attempt to measure the role of infrastructure utilising (4) introduced a constant and a trend variable as a proxy for $\ln A_t$.

As discussed rather in length in Torrisi (2009) an important issue is what "put in place of" *G*.

Put differently, when attempting at estimating (4) scholars have to decide – compatibly with data availability- not only if use, in Romp and de Haan (2007)'s words, "(the monetary value of) the public capital stock (or the monetary value of the stock of infrastructure) ", but also if use a stock or a flow (monetary) measure of public capital.

Nevertheless, at this regard Irmen and Kuehnel (2008) argue that "the analysis […] *using* the stock *measure* confirms most results that are obtained in the flow case", although they continue noting that different results arise in the welfare analysis.

Whatever the choice between different solution available three main issues arise in using production function approach, namely the fact that **1.** labour and capital are exogenous (i.e. this approach does not take into account the role of factor prices in determining their utilisation), **2.** reverse causation from income to investments and, in turn, to private capital (*see* Romp and de Haan (2007) for a general discussion).

However, most important from the point of view of this analysis, is **3.** the potential feedback from income to a demand for infrastructure.

Indeed, on the one hand exist arguments in favour to the thesis arguing that infrastructure increases the output level according to what Looney and Frederiksen (1981) in their paper call the "Hansen (1965) thesis".

On the other hand have been developed arguments in favour to the thesis asserting that infrastructure is only an accommodating factor so that the demand for infrastructure increases with the level of income (Glover, and Simon 1975; Evans, and Karras 1994; Zegeye 2000) following the same behaviour of a *normal good*: public sector spending may be a normal good. That is, as income rises the demand for public infrastructure increases so that the correlation between infrastructure and output may reflect the marginal propensity to consume public goods rather than any productivity enhancing effects of infrastructure (Zegeye 2000).

In this regard various solutions have been proposed to deal with the issue of causality.

Fernald (1999), for example, derives an appropriate test to investigate the direction of the causality between infrastructure and income.

The strategy chosen by the Author works as follows: using data for 29 sectors in the US economy regarding the period from 1953 to 1989, he finds that changes in road growth are associated with larger changes in productivity growth in industries that are more vehicle intensive.

This circumstance leads Fernald (1999) to assert that infrastructure (rather roads) are exogenous. In fact, if road were endogenous, any particular relationship between industry's vehicle intensity and its relative productivity performance should be found when road growth rate changed.

Nonetheless, Canning and Pedroni (1999) find that the causality run in *both* direction by mean of a dynamic error correction model (DECM). In short, since physical stock of infrastructure and per capita income are individually non-stationary but cointegrated, they use a DECM and then test restrictions with the final purpose to study the direction of causality. As said, it appears that causality is not unidirectional but infrastructure enhances productivity and *vice versa*.

A second approach in studying the causality direction consists in using panel data methods. The underlying idea is that pooling data across different unit allows identifying the long-run production function.

Following this approach, Canning and Bennathan (2000), find an high rate of return for electricity generating capacity and the length of paved roads.

With the same aim to capture the results of infrastructure investments (and not the results of economic growth) by mean of panel data approach, Demetriades and Mamuneas (2000) and – in another work – Esfahani and Ramìres (2003) handled the causality issue by introducing a "time-lag" between variables for public infrastructure and productivity.

In these studies, investments were compared to the productivity data several years afterwards, in order to reduce the chance of misrepresentation of economic growth impacts as productivity impacts.

Both studies cited above found that public infrastructures do have a considerable impact on increasing productivity and economic growth. In particular, Esfahani and Ramìres (2003) find that the contribution of infrastructure services to GDP is substantial and, in general, exceeds the costs of provision.

Finally, instrumental variable (IV) is another approach used to deal with causality. Calderon and Serven (2002) adopting the IV approach estimate a Cobb-Douglas production function (in first difference) using lagged values of explanatory variables. Their main finding is an average elasticity of 0.16 for different types of infrastructure.

To summarise: this section focused on the production function approach to measure the economic impact of infrastructure on productivity. The main contents expressed here can be expressed as follows: (i) although estimates vary sensibly from one study to the other, a statistically significant relationship between infrastructure investment and productivity is found in most studies; (ii) the direction of causality is still object of debate. However, most authors, using different approaches, tend to support the thesis that public capital drives productivity, and not the other way around.

Next section focus on the cost-function approach which is an alternative approach developed with the principal purpose to take into account factor prices here not considered at all.

3. The cost-function approach

One of the limitations of the production function approach is that it does not take into account the role of factor prices in determining their utilisation: it reflects only technological relations.

Indeed, private factors inserted in the production function are considered exogenous and it is implicitly assumed that they are paid according to their marginal productivity. At this regard some studies have used a translog function because it is more general than the Cobb-Douglas function (among others Puig-Junoy, 2001; Stephan, 2002; Kemmerling and Stephan, 2002; Wylie, 1996).

An alternative way to deal with this issue consists in adopting the cost function approach. According to the latter the impact of infrastructure on productivity should be analysed in terms of *cost savings*.

Studies following the cost function approach aim to examine if the cost of output *decreases* as the infrastructure endowment *increases*.

The main idea followed by this approach is that public capital can be considered as a *free input* provided by government able to reduce the cost sustained by firms.

In this setting input prices are exogenously determined, so that the variables that firms can choose to produce a given level of output at minimum private cost (*C*) are the quantities of private input.

In symbols $C(p_t^i, q_t^i, A_t, G_t) = \min \sum p_t^i q_t^i \quad subY = f(q_t^i, A_t, G_t)$ (5)

where p and *q* index respectively the input price and the quantities of private input.

The parameter of interest is the *shadow price* (s_g) of the public capital[35] which is obtained by taking the negative of the partial derivative of the cost function with respect to the public infrastructure measure (G) by mean of the Shepard's Lemma. In short, shadow price can be expressed as follows

$$s_g = -\frac{\partial C(p, q, A_t, G)}{\partial G} \tag{6}$$

It is worth noting that an alternative approach consists in assuming that firms aim to maximise their profits (π) given the output prices (p^y) and input prices. This second way can be expressed in symbols as follows

$$\pi(p_t^y, p_t^i, q_t^i, A_t, G_t) = \max p_t^y Y_t - \sum p_t^i q_t^i \quad subY = f(q_t^i, A_t, G_t) \tag{7}$$

Thus, according to this approach the amount of public capital available (*G*) is an *environmental* variable that firms take into account when they optimise their behaviour. A key point at this regard is that although the stock of infrastructure is considered externally given […], each individual firm must still decide the amount it wants to use. This […] leads to the need of a demand function for infrastructure that must satisfy the conditions of standard marginal productivity theory (Romp, and de Haan 2007).

It was noted (Sturm, Jacobs *et al.* 1995) that an important advantage of the cost function approach compared to the production function approach is that it represents a more flexible functional form.

For example, it does not require *a priori* restrictions regarding substitutability of inputs. The cost function approach allows also investigating both direct – as the production function does - and *indirect* effects of public capital, in the sense that firms can vary their demand for private inputs in light to the fact that public capital might be either a substitute or a complement to other input.

Nevertheless, all this flexibility presents also a critical implication. Indeed, It requires good-quality data in order to estimate parameters and to deal with possible multicollinearity problems.

Hence, its strength point becomes, in turn, also its weakness one and a careful consideration involving the trade-off between the two aspects should be made before adopting it.

Overall, studies using the cost function approach show that public capital is *cost reducing*. However, estimates following this approach give a smaller effect than those estimates following the production function approach.

[35] Note that conceptually the shadow price represents the cost-side equivalent of the marginal product, reflecting the reduction in variable costs of production due to an additional infrastructure investment (*see* Morrison, C.J. and Schwartz, A.E. 1996. State Infrastructure and Productive Performance. *American Economic Review* 86(5): 1095-1111).

For example, Ezcurra, Gil *et al.* (2005) (in their study regarding Spanish regional production costs in the agricultural, industrial, and services sectors for the period from 1964 to 1991) find that public infrastructure reduces private costs and increases productivity.

Their estimate shows that while agricultural and service sector behave similarly, the greatest saving in private costs are found in the industrial sector: -0.154 (dollar costs per unit of public capital) for the latter, -0.145 and -0.144 for services sector and agricultural sector respectively.

Cohen and Morrison Paul (2004), realised a study regarding the cost-saving effects of infra-state public infrastructure investment in US which is worth mentioning according to (at least) three different point of view.

First, their model distinguishes between *intra* and *inter*-state effects of public infrastructure taking into account the possibility of interaction between the two categories of infrastructure.

In general terms, they find that taking spill over effect into account raises the average elasticity from -0.15 to -0.23. More deeply, they found that the largest *intra*-state effects appear in the western part of US confirming the theoretical reasoning that *inter*-state infrastructure is not crucial for state – such California – large and relatively densely populated.

Second, regarding the relationship between public and private capital (in terms of complementarity or substitutability) the Authors argue that "the output growth motivated by cost-depressing effects of infrastructure investment may stimulate capital investment and labour employment, even though overall short run public infrastructure-private output substitutability is evident at existing output levels" (Cohen, and Morrison 2004).

Third, the study also address the issue of *causality* by means of the Hausman test[36] concluding that the null hypothesis of infrastructure exogeneity is not rejected. This result is important because it empirically confirms that infrastructure does affects costs and not the other way around.

In conclusion, even if with different shades across sectors and level of analysis, studies following the cost function approach confirm the finding of those following the production function approach: infrastructure and production are positively linked, and, generally speaking, the direction of causality goes from the former to the latter.

4. Growth-model approach

Growth models aiming to test the economic impact of infrastructure are based on the general idea that economic growth is not driven merely by exogenous factors rather by dynamics which are internal to the economic system itself.

Indeed, since the mid-1980s, many studies were developed in order to explain why difference in income both over time and across countries did not disappear as the neoclassical growth models predicted.

The main feature of this tradition is the assumption that growth is an *endogenous* phenomenon affected by economic agents' behaviour. A key feature in explaining different performance is assigned to capital formation which meaning has to be intended in a broader sense including physical capital as well as human capital, knowledge capital and infrastructure.

The general economic framework to empirically test these assumptions can be expressed as follows:

$$\Delta \ln\left(\frac{Y}{L}\right)_{0,T} = \alpha + \beta\left(\frac{Y}{L}\right)_{0} + \gamma\left(\frac{I^{G}}{Y}\right)_{0,T} + Z \qquad (8)$$

Where $\left(\frac{Y}{L}\right)_{0,T}$ represents the average per capita GDP over the period [0; *T*], $\left(\frac{Y}{L}\right)_{0}$ is the initial level of real per capita GDP and $\left(\frac{I^{G}}{Y}\right)_{0,T}$ is factor added to represent government investment (rather the average rate of public investment as percentage of GDP over the [0; *T*] period); *Z* captures a set of conditional variables such

[36] Most studies test for endogeneity and find that infrastructure can be considered exogenous, but not all the studies do so. For example, the first study here cited – Ezcurra, Gil *et al.* (2005)- does not perform the Hausman test, arguing that since it regard regional data, endogeneity was not a significant problem. At this regard, as argued in Infrastructure Canada (2007), should be noted that "this may not be a sufficient justification to rule out the endogeneity problem" (Infrastructure Canada 2007).

as private investment (as percentage of GDP), proxy for human capital (usually primary and/or secondary enrolment), political instability (assassinations, revolts and coups, and war casualties), freedom, and the ratio of government consumption to GDP.

Note that while the parameter β measures technological catch up (if negative), the parameter γ - being a measure of the impact on *growth* - is not the same as the marginal productivity of capital when the measure of economic performance (for example GDP) is considered in level.

Easterly and Rebelo (1993)'s article represents an important piece of work using public capital in an empirical growth model. The Authors run pooled regressions (using individual country decade averages for the 1960s, 1970s and 1980s) of per capita GDP growth on a set of conditional variables and on public investment in various sector (added one at time): agriculture, education, health, housing and urban infrastructure, transport and communication, industry and mining.

Their work shows that the share of public investment in transport and communication infrastructure is robustly correlated with growth (with coefficient ranging from 0.588 to 0.661 according to different specifications used) as well as almost all other variables except agricultural spending which is consistently *negatively* related with growth with a coefficient between - 0.34 and - 0.231).

Moving from Easterly and Rebelo (1993) other works have been realised adopting also regional data.

Mas, Maudos *et al.* (1996), for example, regarding Spanish regions found that the initial stock of public capital (as share of gross value added) positively affects output expressed in per-capita terms.

Crihfield and Panggabean (1995), using two stages estimation technique to take into account also capital and labour endogeneity, achieved the conclusions that public infrastructures that they considered (e.g. education, streets, highways, sewerage, sanitation) surely play a role, with the caution that their contribution may be less than that the one of other forms of investment.

With respect to the contribution of specific infrastructures, Cellini and Torrisi (2009), focusing on infrastructure specific to the tourism sector, show that this particular kind of infrastructure, separately considered, has a weak impact on several indicators of economic performance (e.g. gdp, touristic presence, hotels' structures) considered also in terms of growth rate.

However, various authors have pointed at problems associated with cross-section regressions.

To begin with, biases due to omitted variables, reverse causation (Levine, and Renelt 1990; Levine, and Zervos 1993) and sample selection (De Long 1988) could affect the results which interpretation, as pointed out by Solow (1994), is often tempted by wishful thinking.

Furthermore, cross-section regressions are often not very robust. Indeed, several models *ex-ante* reasonable given the data, achieve different conclusions about the parameter of interest.

Put in Levine and Renelt (1992)'s words, given that over 50 variables have been found to be significantly correlated with growth in at least one regression, readers may be uncertain to the confidence they should place in the findings of any one study (Levine, and Renelt 1992).

In order to deal with the issue of how robust the result concerning a certain variable is to the inclusion of other relevant variables Levine and Renelt (1992), using a variant of Leamer (1978), elaborated the so-called extreme bound analysis (EBA).

According to the EBA approach should be reported an upper and an lower bound for parameter estimates obtained in regressions using as explicative variables different subsets of the set of explanatory variables.

The relationship between a certain variable and economic growth *is not* considered *robust* either if a certain variable became statistically insignificant or if the sing of its parameter in the upper bound case is different from the one obtained in the lower bound case.

Unfortunately, one of the main results of the latter study is that "few findings can withstand slight alterations in the list of explanatory variables" (Levine, and Renelt 1992).

Going further on the empirical exploratory ground, next section focuses on vector autoregression models which represent a set of *data oriented* models, i.e. models developed to use as little theory as possible in order to manage theoretical and empirical problems affecting approaches discussed above.

5. Vector Autoregression Models

Vector Autoregression (VAR) models represent a theoretical framework used with the specific purpose to deal with theoretical limitations and significant empirical controversies over the impact of infrastructure on productivity summarised above.

Indeed, the peculiar characteristic of a VAR model is that no *a priori* causality directions are imposed among variables[37]. In a VAR model a limited number of variables is considered and explained by their own lags and the lags of the other variables, so that all variables are treated as jointly determined.

In a formal way a VAR including two variables (let they be *y* and *g*) can be expressed as follows[38].

$$y_t = \partial_0 + \sum_{i=1}^{L_y} \alpha_i y_{t-i} + \sum_{i=1}^{L_g} \gamma_i g_{t-i} + u_t \quad (8a)$$

$$g_t = \eta_0 + \sum_{i=1}^{L_y} \beta_i y_{t-i} + \sum_{i=1}^{L_g} \rho_i g_{t-i} + \varepsilon_t \quad (8b)$$

Where L_y and L_g index respectively the number of lags of *y* and *g* to be considered; each Equation contains also an error term (u_t, ε_t) that has zero expected value given past information on *y* and *g*.

An unrestricted VAR model can be estimated using OLS even if variables are integrated and possibly cointegrated (Sims, Stock *et al.*, 1990).

Note that if *n* variables are included with each p lags then $n^2 \times p$ coefficients need to be estimated, besides the deterministic variables. A way to deal with this problem consists in using Akaike's (1969, 1970) Final Prediction Error (FPE) criterion in order to select the appropriate lag specification for each explanatory variable in each Equation and save a substantial number of degrees of freedom.

Alternatively, an F test for jointly significance can be used in order to choose how many lags have to be inserted for each variable (Wooldridge 2002).

Typically, studies following this approach apply *Grainger-causality* tests to find relationships between variables. In our case researchers are particularly interested in testing if infrastructure Grainger-causes productivity - i.e. if the time series prediction of GDP (or some other measure of productivity) from its own past improves when lags of measures of infrastructure are considered - and\or *vice versa*.

At this regard should be noted that although VAR models allow us to test whether the hypothesis that infrastructure *causes* productivity is valid or whether there are feedback effects from output to public capital (Romp, and Haan 2007; Sturm, Kuper *et al.* (1996), Infrastructure Canada 2007), VAR models do not definitively solve the problem of endogeneity.

Indeed, the term "causes" in "Granger causes" should be interpreted with caution. In particular, it has nothing to say about contemporaneous causality […], so it does not allow us to determine whether [a certain variable] is […] exogenous or endogenous (Wooldridge 2002).

In a relatively recent study utilising VAR models with Spanish regional data Pereira and Sagalés (2006) founded that infrastructure investments positively affect private output and also *crowd-in* private sector inputs.

Put differently, the study shows that public investment in infrastructure and private expenditure in the same field are complementary rather than substitutes.

The same conclusion has been achieved in Karadag, Ozlem Onder *et al.* (2005) with respect to the Turkish case.

Another interesting conclusion driven by Pereira and Sagalés (2006) is that surprisingly infrastructures contribute in creating disparities between regions due to fact that new investment on infrastructure are most often directed to central regions disadvantaging peripheral regions.

Sturm, Jacobs *et al.* (1995) (using data regarding the Netherlands from 1853 to 1913) consider GDP, investment series on public infrastructure, private machinery, and equipment capital to provide evidence for *unidirectional* positive relationship from *basic* infrastructures to GDP only, while the *complementary* ones appear to be not effective.

---

[37] Note that since the VAR approach does not completely reveal the underlying production process, estimates do not represent elasticity as in the production function approach. However, in order to get specific elasticity estimates starting from a VAR model can be adopted the impulse-response functions. This method allows estimating the long-run effects of different shocks on the estimated system. Applying this method requires rewriting the VAR into its Vector Moving Average (VMA) representation and, in turn, the model needs to be stable in order to make this conversion. A sufficient condition that makes the model stable is that the variables used are stationary or co-integrated.

[38] Usually, a four-variables model (output, employment, private capital, and public capital) is utilised.

Nonetheless, Xiaobo and Fan (2001) (using data regarding Indian economy) find that infrastructure and productivity often affect *each other* in the long term (i.e. estimating the model in levels).

With respect to short term (i.e. estimating the model in first differences), instead, the Authors find that the coefficients are not statistically significant.

In conclusion, papers designed on data based models reviewed in this section confirm, although once more with different shades, that public capital investments positively impact private sector output, despite the fact that they use different datasets and theoretical constructs.

More precisely, regarding the most debated point involving (Grainger-)causality, some authors conclude that *at least* infrastructure and productivity affect *each other* but no study find evidence to support the hypothesis of strict *reverse causation* from output to infrastructure.

6. Conclusion

This paper briefly reviewed the vast literature concerning the relationship between infrastructure and productivity focusing on some critical points.

Indeed, since the first-generation studies primarily based on production function and cost function approaches a significant amount of discussion on some of the theoretical and econometric issues have been developed.

This paper reviewed some of them along an ideal path from three theory-based to data-oriented models.

The actual area of significant debate, besides the magnitude of infrastructures impact on productivity and/or the causality direction, concerns other mentioned issues of **1.** short-run and the long-run significance of their contribution and **2.** the effectiveness of different category of infrastructures.

In order to deal with these issues several studies and approaches have been developed reporting that the peculiar feature in this field is represented by *heterogeneity*: the effects of public investment differs across countries, regions, and sectors.

It is worth noting that this result is reasoned according to arguments based both on economic and political grounds.

On the economic ground should be noted that the effects of new investment depend on “past history” (i.e. the quality and the quantity of the capital stock in place): the larger the quantity and the better its quality, the lower the impact of additional investment.

However, as said above, another source of heterogeneity can be found at the institutional and political ground, even if this issue (probably) have not been well researched.

Indeed, in Estache (2006)‘s words there is strong anecdotal evidence now that politics matter. [First, because] politicians will never give up the control of a sector that buys votes in democratic societies. Moreover, in societies in which corruption is rampant, they will not give up control of a sector involving large amount of money and in which contract award processes often provide opportunities for unchecked transactions (Estache 2006).

References

[1] Akaike, H. 1969. Fitting Autoregressive Models for Prediction. *Annals of the Institute of Statistical Mathematics*, 21: 243-247.

[2] Akaike, H. 1970. Statistical Predictor Identification. *Annals of the Institute of Statistical Mathematics* 22: 203-217.

[3] Albala-Bertrand, J.M., and Mamatzakis, E.C. *et al.* 2004. The Impact of Public Infrastructure on the Productivity of the Chilean Economy. *Review of Development Economics* 8: 266-278.

[4] Aschauer, D.A. 1989. Is Public Expenditure Productive? *Journal of Monetary Economics* 23: 177-200.

[5] Barro, R.J. 1988. Government Spending In A Simple Model Of Endogenous Growth. *National Bureau Of Economic Research Working Paper* No. 2588.

[6] Belloc, M., and Vertova, P. 2006. Public Investment and Economic Performance in Highly Indebted Poor Countries: An Empirical Assessment. *International Review of Applied Economics* 20 151-170.

[7] Biehl, D. 1991. *The Role of Infrastructure in Regional Development. in Infrastructure and Regional Development*. R.W.Vickerman. London, Pion.

[8] Bonaglia, F., La Ferrara, E. *et al.* 2001. Public Capital and Economic Performance: Evidence from Italy. *IGIER Working Paper* No. 163.

[9] Calderon, C., and Serven, L. 2002. The Output Cost of Latin America's Infrastructure Gap. *Central Bank of Chile Working Paper* No. 186.

[10] Canning, D., and Bennathan, E. 2000. The Social Rate of Return on Infrastructure Investments. *World Bank Working Paper* No. 2390.

[11] Canning, D., and Pedroni, P. 1999. Infrastructure and Long Run Economic Growth. *Mimeo*.

[12] Cass, D. 1965. Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies* 32: 233-240.

[13] Cellini, R., and Torrisi, G. 2009. The Regional Public Spending for Tourism in Italy: An Empirical Analysis. *MPRA paper* No.16131.

[14] Cohen, J.P., and Morrison, Paul, C.J. 2004. Public Infrastructure Investment, Interstate Spatial Spillovers, and Manufacturing Costs. *Review of Economics and Statistics* 86: 551-560.

[15] Crihfield, J., and Panggabean, M. 1995. Is public infrastructure productive? A metropolitan perspective using new capital stock estimates. *Regional Science and Urban Economics* 25: 607-630.

[16] De Long, J.B. 1988. Productivity Growth, Convergence, and Welfare: Comment. *American Economic Review* 78: 1138-54.

[17] Demetriades, P.O., and Mamuneas, T.P. 2000. Intertemporal Output and Employment Effects of Public Infrastructure Capital: Evidence from 12 OECD Economies. *Economic Journal* 110: 687-712.

[18] Di Palma, Maurizio, and Mazziotta, C., *et al.*, Eds. (1998). Infrastrutture e sviluppo. Primi risultati: indicatori quantitativi a confronto (1987-95). Quaderni sul Mezzogiorno e le politiche territoriali. Roma, Confindustria.

[19] Easterly, W., and Rebelo, S. 1993. Fiscal policy and economic growth An empirical investigation. *Journal of Monetary Economics* 32: 417-458.

[20] Esfahani, H., and Ramìres, M.T. 2003. Institutions, Infrastructure and Economic Growth. *Journal of Development Economics* 70: 443-477.

[21] Estache, A. 2006. Infrastructure. A Survey of Recent and Upcoming Issues. *World Bank. Mimeo*.

[22] Evans, P., and Karras, G. 1994. Is government capital productive? Evidence from a panel of seven countries. *Journal of Macroeconomics* 16(2): 271-279.

[23] Everaert, G., and Heylen, F. 2004. Public Capital and Long-Term Labour Market Performance in Belgium. *Journal of Policy Modelling* 26: 95-112.

[24] Ezcurra, R., Gil, C., *et al.* 2005. Public Capital, Regional Productivity and Spatial Spillovers. *The Annals of Regional Science* 39: 471-94.

[25] Fernald, J. 1999. Assessing the Link between Public Capital and Productivity. *American Economic Review* 89: 619-638.

[26] Ferrara, E.L., and Marcellino, M. 2000. TFP, Costs, and Public Infrastructure: An Equivocal Relationship. *IGIER Working Paper* No. 176.

[27] Ghali, K.H. 1998. Public Investment and Private Capital Formation in a Vector Error-Correction Model of Growth. *Applied Economics* 30: 837-844.

[28] Glover, D.R., and Simon, J.L. 1975. The effect of population density on infrastructure: the case of road building. *Economic Development and Cultural Change* 23: 453--468.

[29] Gramlich, E.M. 1994. Infrastructure Investment: A Review Essay. *Journal of Economic Literature* XXXIII: 1176-1196.

[30] Hansen, L. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50(3): 1029-1054.

[31] Hansen, N.M. 1965. The structure and determinants of local public investment expenditures. *Review of economics and statistics* 2: 150-162.

[32] Hausman, J. 1978. Specication tests in econometrics. *Econometrica* 46(3): 1251-1271.

[33] Infrastructure Canada. 2007. *Infrastructure and Productivity: A Literature Review. Transport Infrastructure and Communities Portfolio*, Government of Canada. Research and Analysis Division Infrastructure.

[34] Irmen, A., and Kuehnel, J. 2008. Productive Government Expenditure and Economic Growth. *CESifo. Working paper* No. 2314.

[35] Kamps, C. 2006. New Estimates of Government Net Capital Stocks for 22 OECD Countries 1960-2001. *IMF Staff Papers* 53: 120-150.

[36] Karadag, M., Ozlem Onder, A., *et al.* 2005. Growth of Factor Productivity in the Turkish Manufacturing Industry at Provincial Level. *Regional Studies* 39,(2): 213-223.

[37] Kemmerling, A., and Stephan, A. 2002. The Contribution of Local Public Infrastructure to Private Productivity and its Political Economy: Evidence from a Panel of Large German Cities. *Public Choice* 113(34): 403-424.

[38] Koopmans, T.C. 1965. *On the concept of optimal economic growth*. in (Study Week on the) Econometric Approach to Development Planning. Amsterdam, North-Holland Publishing Co: chap. 4. pp. 225-87.

[39] Leamer, E.E., Ed. 1978. *Specification Searches: Ad Hoc Inference from Nonexperimental Data*. New York, Wiley.

[40] Levine, R., and Renelt, D. 1990. *Cross Country Studies of Growth and Policy: Some Methodological, Conceptual, and Statistical Problems*. World Bank.

[41] Levine, R., and Renelt, D. 1992. A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review* 82: 942-963.

[42] Levine, R., and Zervos, S. 1993. What We Have Learned About Policy and Growth from Cross-Country Regressions? *American Economic Review*, 83: 427-430.

[43] Looney, R., and Frederiksen, P. 1981. The Regional Impact of Infrastructure Investment in Mexico. *Regional Studies* 15(4): 285-296.

[44] Marrocu, E., Paci, R., *et al.* 2006. Gli effetti del Capitale Pubblico Sulla produttività delle Regioni Italiane. In: Barca, F., Cappiello, F., Ravoni, L., and Volpe, M. *Federalismo, equità, sviluppo. I risultati delle politiche pubbliche analizzati e misurati dai Conti Pubblici Territoriali*. Bologna, Il Mulino.

[45] Mas, M., Maudos, J., *et al.* 1996. Infrastructures and Productivity in the Spanish Regions. *Regional Studies, Taylor and Francis Journals* 30(7): 641-649.

[46] Mastromarco, C., and Woitek, U. 2004. *Public Infrastructure investment and efficiency in Italian Regions*. www.cide.info/conf/papers/1161.pdf.

[47] Mittnik, S., and Neumann, T. 2001. Dynamic Effects of Public Investment: Vector Autoregressive Evidence from Six Industrialized Countries. *Empirical Economics* 26: 429-446.

[48] Morrison, C.J., and Schwartz, A.E. 1996. State Infrastructure and Productive Performance. *American Economic Review* 86(5): 1095-1111.

[49] Pereira, A.M. 2000. Is All Public Capital Created Equal? *Review of Economics and Statistics* 82: 513-518.

[50] Pereira, A.M. 2001. On the Effects of Public Investment on Private Investment: What Crowds In What? *Public Finance Review* 29: 3-25.

[51] Pereira, A.M., and Sagalés, O.R. 2006. Public Infrastructures and Regional Asymmetries in Spain. Department of Economics. *College of William and Mary Working Papers* 46.

[52] Puig-Junoy, J. 2001. Technical Inefficiency and Public Capital in U.S. States: A Stochastic Frontier Approach. *Journal of Regional Science* 41(1): 75-96.

[53] Putnam, R.D., Ed. 1993. *Making Democracy Work. Civic Tradictions in Modern Italy*. Princenton Princenton University Press.

[54] Romp, W., and de Haan, J. 2007. Public Capital and Economic Growth: A Critical Survey. *Perspektiven der Wirtschaftspolitik* 8(Special Issue): 6-52.

[55] Sanchez-Robles, B. 1998. Infrastructure Investment and Growth: Some Empirical Evidence. *Contemporary Economic Policy* 16: 98-108.

[56] Sims, C., Stock, J., *et al.* 1990. Inference in Linear Time Series Models with Some Unit Roots. *Econometrica* 58: 113-144.

[57] Solow, R.M. 1956. A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* 70: 65-94.

[58] Solow, R.M. 1994. Perspectives on Growth Theory. *The Journal of Economic Perspectives* 8(1): pp. 45-54.

[59] Stephan, A. 2002. Regional Infrastructure Policy and its Impact on Productivity: A Comparison of Germany and France. *Wissenschaftszentrum Berlin*, CIC Research Unit Working Paper No. FS IV 0002.

[60] Sturm, J.-E., Jacobs, J., *et al.* 1995. *Productivity impacts of infrastructure investment in the Netherlands 1853-1913*. University of Groningen. Research Institute SOM (Systems, Organisations and Management) Research Report 95D3.

[61] Sturm, J.-E., Kuper, G.H., *et al.* 1996. Modelling government investment and economic growth on a macro level: a review. *CCSO Series* No. 29.

[62] Torrisi, G. 2009. Public Infrastructure: Definition, Classification and Measurement Issues, *Economics, Management, and Financial Market*. 4 (3).

[63] Wooldridge, J.M., Ed. 2002. *Introductory Econometrics A Modern Approach*, South-western Publishing Group.

[64] Wylie, P. 1996. Infrastructure and Canadian Economic Growth. *Canadian Journal of Economics* 29 (Special Issue): S350-355.

[65] Xiaobo, Z., and Fan, S. 2001. How Productive is Infrastructure? New Approach and Evidence from Rural India. *International Food Policy Research Institute Discussion Paper* No. 84.

[66] Zegeye, A.A. 2000. *US Public Infrastructure and its contribution to private sector productivity*. US Department of Labor Bureau of Labor Statistics.

A SURVEY ON LABOR MARKETS IMPERFECTIONS IN MEXICO USING A STOCHASTIC FRONTIER

Juan M. VILLA
Inter-American Development Bank, **United States**
juanmiguelv@iadb.org

Abstract

It is assumed that observed labor income is the result of three stages across the job search process from the reservation wage formation, the bargaining between employers and potential employees when the match, and finally a possible additional adjustment once the worker is completely hired. This paper provides a methodological proposal and an intuitive estimation of the wage gain due to the presence of labor market imperfections across those three stages. The part of the wage that is explained by labor markets imperfections is estimated by performing a stochastic frontier model with panel data belonging to the Mexican labor survey – ENOE. The results suggest that 82.7% of the variance of the wages of the subordinated workers is explained by market imperfections. Moreover, public labor offices and small firms are negatively correlated with their presence.

Keywords: labor market imperfections, wage formation, reservation wage, job search, job matching, stochastic frontier, Mexico, ENOE

JEL Classification: J21, J31

1. Introduction

The labor income of a typical worker is the observable outcome of a complex process shaped by the overall competence conditions of the labor market before a job offer is accepted and once the recruitment is completed. The basic question that still remains by theoretically and empirically unanswered is how this outcome may differ among similar workers. As the explanations might be diverse, the answer does not only entail the analysis of the current conditions of employees, but also the circumstances under which unemployed workers performed the job search task similarly with their reservation wage formation.

In a competitive market, firms or employers would not have an incentive to pay wages over those reservation wages in accordance given their productivity level but, certainly, observable wages become a representative symptom of how the competitive result is modified and a cause of why the labor market is not cleared (Fields 2007).

Why similar workers earn different wages is an issue that has been theoretically addressed but at the same time is has been barely quantified and rarely connected with a wage formation process. As an attempt to contribute to the empirical analysis of labor markets, the main purpose of this document is to quantify the proportion of the wages that is caused by the labor market imperfections in Mexico.

2. Intuition on labor markets imperfection and wages formation

2.1. A brief review of the theory

There are three stages that intervene in the determination of wage levels, from the job search process through the duration of unemployment to the job offer acceptance and an eventual and subsequent wage adjustment. At a starting point, the reservation wages works as a lower bound of the final labor income assignment. As long as agents interact depending on the exposure of the workers to the labor markets imperfections upon each one of these stages, an observable gain above that reservation wage would probably arise.

One may distinguish the origin of the heterogeneity in wages. When it depends on the supply side, there are certain characteristics from workers that are not easily detectable by linear regressions (i.e. ability, job search preferences, etc.). On the other hand, the source of the differences on the demand side emerges with the presence of similar unobservable behavior from employers, such as prejudiced preferences, biased assessments of a worker's productivity, asymmetric information or monopsony (Van Klopin, and Singell 1996).

Similarly, as corroborated by the labor literature, these market imperfections are seldom studied as a whole. Although they have been studied separately, the quantification of their effect on wages between workers and employers has been practically ignored. For instance, since Stigler (1962) addressed the information in the search stage of wage formation it became the most studied issue. Rogerson *et al* (2005) address bargaining power in an effort of obtaining conclusions on the employers – employees relation. Baldwin (1991) exhibits a

labor discrimination model among black and white workers. Arnott, and Stiglitz (1985), and Arnott *et al.* (1988) broadly assess the labor turnover conditions and van den Berg (2003) surveys the consequences of the existence of minimum wages as an inhibitor of monopsony power.

In the job search stage, reservation wages determine the core of what will be the final observed wage once the employee is hired in a new job. The existence of inherent worker properties shapes the minimum amount of money that he or she would be willing to receive to accept a vacancy according to its optimal labor supply plan. Number of children, marital status, gender, location and some others characteristics of workers contribute to the variations of that amount of money or in kind payment amongst them. Examples of these empirical findings are found in Owen (1971), Heckman and Willis (1979) or Heckman (1986), who have employed empirical labor market participation models to determine which characteristic has a higher weight (*see* also Moffitt, and Wolfe 1992; Siegfried, *et al.* 1993).

The characterization of reservation wage models has treated it as a constant variable over time (Kiefer, and Neumann 1979) even during the job search when it works as an essential input of the seeker's stopping rule as shown by Mortensen (1986). Nevertheless, evidence of a decreasing reservation wage as long as unemployment duration was found in empirical studies such as Blau (1991) who does not only include an expected wage offer but rather the expected number of workable hours as the solution of the worker's labor supply plan. Thus, even though one source of overall wages decrease throughout negative business cycles is the reaction that reservation wages have when consequently the unemployment duration increases. Besides business cycle, the lack of information about vacancies might worsen the situation of unemployed workers whose labor income comes down when they do not know where employers are and duration increases.

Information on vacancies is the first element that job seekers attempt to retrieve once the unemployment begins. The better the information one gets, the better a job offer is in terms of its remuneration (Bloemen 2005) that modifies the worker's reservations wage. Several studies have shown that job seekers prefer direct information sources, such as "relatives and friends" or "directly at the work place" (Kahn, and Low 1990) over indirect information sources like Labor Market Intermediaries (Autor 2008). The main reason why this occurs is that the cost of information is smaller using the direct sources than using the indirect sources, determined by the job seeker's characteristics and their optimal level of demand for information.

Therefore, asymmetric information is the first imperfection that job seekers and employers generate in the labor market, firstly because employers cannot predict the potential productivity of a given worker and because the latter cannot predict the real conditions of an accepted job offer (most of the jobs characteristics are revealed after recruitment and not during the job search task). For example, Mavromaras *et al*. (2009) survey over-education as the most important concern of workers and a subsequent job mismatching issue in the presence of the asymmetric information that leads to the unavoidable dissatisfaction of workers or firms. This causes undesirable and costly job turnover and hence, employers decide to hire those workers with potential specific skills in terms of trainability, productivity and the least-likely to give up (Salop, and Salop 1976).

Although wage discrimination has led the vast amount of theory assessments and empirical evidence (especially for gender) as well as significant pool of causes of labor market segregation[39]. During the job search stage of wage formation firms or employers do not only have specific preferences for certain workers but rather job seekers aim their search towards those jobs that brings them a considerable satisfaction. For instance, men and women often do quite different jobs because men have less risk aversion than women (Cobb-Clarck, and Tan 2009).

Rogerson, Shimer and Wright (2005) offer an interesting approach in the transition from the first to the second stage of this survey on wage formation and its relation with markets imperfection. They question how there may be unemployed workers and unfilled vacancies at the same time and argue that "*Workers do not encounter firms completely at random but try to locate those posting attractive terms of trade*[40]". They conclude that their job bargaining models are linked to the available matching technology in the job search and job offer process.

The availability of empirical evidence on wage differentials and job bargaining has shown that in the absence of unions, a worker's productivity determines further wage adjustments. Bartolucci (2009) manages

[39] *See* for example Carroll, and Rolph (1973), Swinton (1977), Gannon, and Zenou (1997), Heckman (1998), Aslund, and Rooth (2005).

[40] Rogerson, R. Robert, Wright, R.S. 2005. Search-Theoretic Models of the Labor Market: A Survey. *Journal of Economic Literature*. Vol. XLIII (December 2005): 959-988.

matched employer-employee data in Germany, showing that wage disparities among genders are explained in large part by differences in productivity and by a smaller part by segregation or the firms' preferences.

The existence of unions is an important element in the wage formation through its third and final stage as well. They play an essential role in sorting out wage inequalities originated by the firms' predominant bargaining power and provide an explanation of why wage gaps are sharper when unionization does not take place (Card, 2001). Nonetheless, as referenced by Bemmels (1987) for the US manufacturing industry, unions are often the source of a possible mismatching of wages and the real productivity that they represent. As Wessels (1985) affirms, "(…) *Unions must decrease employment far more than is commonly thought, or they do not substantially increase the productivity of firm*[41]".

2.2. Intuition on the methodology

The intuition behind the estimation of the part of the labor income originated by the influence of the market imperfections on the wage formation process begins with the generic specification of the classic Mincer (1974) Equation. From the latter it is only drawn its theoretical components rather than an attempt to assess the returns to schooling. The specification for each employee i in period t is:

$$lnw_{it} = \beta_0 + \sum_{j=1}^{k} \beta_j x_{jit} + e_{it} \quad (1)$$

Where the core variables of the covariates x_{jit} are commonly denominated education and experience (or rather, worker's age). Arbitrarily it is complemented by the number of hours worked per month (*whm*) and the economic sector (*sector*) where the employee actually works. In any econometric context the error term e_{it} is the part of the observable labor income lnw_{it} that is not explained by the covariates x_{jit} and hence, if two workers have the same age, education, worked hours per month and economic sector, the differences in their wage would be incorporated into their own error component. Thus, the larger is the exposure of workers to the presence of labor market imperfections, the greater the error term $e_{it}$is.

The stochastic frontier model divides e_{it} into two error terms:

$$e_{it} = v_{it} + u_{it} \quad (2)$$

Where $v_{it} \overset{iid}{\sim} N(0, \sigma_v^2)$, is the part of the wage differences among similar workers due to idiosyncratic exceptional characteristics which are also personals traits commonly described as "abilities"[42]. The error term $u_{it} \overset{iid}{\sim} N^{+}(\mu, \sigma_v^2)$, reflects the wage gained by employees as a consequence of the existence of labor market imperfections.

Thus, under competitive conditions the labor income amongst workers with the same characteristics would differ only by v_{it}, or simultaneously, by their ability gain. Therefore, after the estimation of Equation 1 the rest of the heterogeneity of the observable wage, (u_{it}), does not have any other explanation besides the role of markets in generating this differential. Regardless its origin, across the stages of wage formation it is not fully controllable by workers[43].

It is necessary to point out that regarding our theoretical context, the part of Equation 1 that reveals the reservation wage of a given worker *i* comes up when u_{it} is subtracted from it (Hofler, and Murphy 1994). An employer with perfect information would not pay u_{it}, nonetheless, it emerges over the three stages of wage formation, from the job search, the job matching and a subsequent adjustment.

Another factor in the specification of Equation 1 is time. Given that the data have an evident unbalanced panel framework because of the re-contacting failures, Greene (2003) explains two restrictions of a feasible time-invariant specification for Equation 1. First, that a time invariant heterogeneity would yield a biased estimation of u_i and supposes the independence between the whole error and the covariates of the model. Second as the model supposes a time-invariance of u_i, even though the time span of the data (two quarters), it is preferable to

---

[41] Wessels, W. 1985. The Effects of Unions on Employment and Productivity: An Unresolved Contradiction. *Journal of Labor Economics*, Vol. 3, No. 1, Part 1 (Jan., 1985): 101-108.

[42] These cognitive and non-cognitive abilities are also estimated by Heckman et al (2006) or Cawley *et al.* (2001). This is an alternative to the estimation of this term, even though this is not the focus of this assessment.

[43] See for instance Boehm and Hofler (1987) where this technique is implemented to also separate markets effects on prices.

generalize the specification of the model over a time-variant basis. Thus, substituting the random parameter e_{it} of the Equation 2 into the Equation 1 is rewritten as follows[44]:

$$lnw_{it} = \beta + \sum_{j=1}^{k} \beta_j x_{jit} + v_{it} + u_{it} \quad (2.3)$$

Where u_{it} is achieved by including one dummy for each individual. Under any specification of Equation 3, in the end it is performed a generalization of u_{it} term proposed by Battese and Coelli (1995) where it is changed into the expression,

$$e^{-\eta(t-T_i)}u_i \quad (2.4)$$

That allows the model to adapt the variation of the market imperfection term across time and individuals. If η is statistically equals to zero, the model would yield a time-invariant random effect estimation. This stochastic frontier model in addition produces a parameter $\gamma = \sigma_u^2/(\sigma_u^2 + \sigma_v^2)$ manifesting the proportion to which the markets imperfections explain the variation of wages for similar employees. Later γ becomes the center of our analysis.

3. Data and Mexican labor markets

The data for the estimation of Equation 3 belongs to the National Survey of Occupation and Employment-ENOE[45], continually carried out in 124.260 dwellings by the National Institute of Statistics and Geography-INEGI[46]. The thematic coverage of the survey entails gender and age composition of the national inhabitants and the labor composition of the working age population. It is stratified on rural and urban areas of the 32 Mexican states over a continuous periodicity and a quarterly report (it follows up most of the observations on a panel structure).

The time scheme of the ENOE involves a continuous fieldwork across the whole year, but given that the re-contact of the households does not remain stable over time, the selected periods for the estimation of Equation 3 are the first two quarters of 2009, which allows us to find more accurate information of the reported households in a previous period[47]. The loss of people quarter by quarter averages about 25% of them, and after one year there is not trace of any dwelling. For instance, 76.3% of the people that were polled in I-2009 appear again in II-2009 while the rest is expected to be in III-2009.

Now, the focus of the periods I-2009 and II-2009 is placed on the main indicators of the Mexican labor markets in order to understand the structure, their potential role and scope of the labor market imperfections. As part of our initial description, it is firstly presented the hierarchical decomposition of the population of Mexico with a Working Age Population between 15 and 65 years of age:

Table 1. Hierarchical decomposition of the Labor Force.

Hierarchical Decomposition of the Labor Force (Hierarchical rates)		
Period:	**I-2009**	**II-2009**
0. Total population	100.0	100.0
1. Population six years and above	89.4	89.5
1.1 Child population (6-14 years of age)	20.3	20.0
1.1.1 Child laborers	2.7	2.9
1.2 Population 65+ years of age	9.3	9.4
1.2.1 Employed	21.0	21.8
1.3 Working age population (15-64 years of age)	70.4	70.6
1.3.1 Inactive	37.4	37.2
1.3.2 Active	62.6	62.8
1.3.2.1 Employed	94.8	94.6
1.3.2.2 Unemployed	5.2	5.4

Source: ENOE-INEGI. I-2009 and II-2009.

[44] *See* Kumbhakar and Lovell (2000) for further explanation on this specification.

[45] Encuesta Nacional de Ocupación y Empleo.

[46] Instituto Nacional de Estadística y Geografía - http://www.inegi.org.mx/inegi/default.aspx.

[47] Using more than two quarters would lead to a small re-contract of individual and a low track of them along time. Find out more information at INEGI's homepage.

The Mexican working age population has been recognized for having the lowest rates of labor participation (IADB, 2004). The active population rate reached a level less than the Latin-American rate and its unemployment rate barely had been above 5% in the last decade. A slower business rhythm has led Mexicans to increase their labor participation between I-2009 and II-2009 that was absorbed by the unemployment rate rather than the employment rate.

Additionally, the distribution of employees among sectors and education levels reveals a concentration of workers in the service sector and among those who have reached the high school education level. During this period, the re-arrangements among these shares also shows and trend towards a mild increase of level of education of Mexican workers and a higher participation of them in the service and agricultural sectors:

Table 2. Distribution of employees – Level of Education.

Distribution of the Employed - Level of Education		
Level of education %	**I-2009**	**II-2009**
Incomplete elementary	14.9	14.4
Elementary	22.2	21.6
High school	33.8	34.0
Technical	14.4	14.8
Graduate	14.7	15.1
Not specified	0.0	0.0
Total	**100.0**	**100.0**

Source: ENOE-INEGI. I-2009 and II-2009.

Table 3. Distribution of employees by Economic Sector.

Distribution of the Employed by Economic Sector		
Sector of economic activity	**I-2009**	**II-2009**
Construction	8.4	8.3
Manufacturing	15.8	15.3
Commerce	19.6	19.6
Services	42.8	43.2
Others	1.0	1.0
Agriculture	11.7	12.0
Total	**100.0**	**100.0**

Source: ENOE-INEGI. I-2009 and II-2009.

The mobility of workers to the service and agricultural sectors was accompanied by an increase of self-employed (non-subordinated) and non-remunerated workers as shown in Table 4:

Table 4. Employment categories, shares in total employment.

Employment Categories, Shares in Total Employment		
Occupational category	**I-2009**	**II-2009**
Remunerated workers	69.0	68.0
Employers	4.2	4.3
Self-employed	20.9	21.4
Non-remunerated workers	5.9	6.2

Source: ENOE-INEGI. I-2009 and II-2009.

A possible symptom of a superfluous inflow of workers in the service sector has turned out between these two periods given the observed changes in the work categories and economic sectors share. Spite of the fact that the employment in the service sector increased, it yielded a lesser share of remunerated workers that coincided with a decrease of the share of workers in the construction and manufacturing sectors (presented in Table 3).

Regarding unemployment, Table 5 presents its rates and shares among groups of gender, age, area of residence and level of education. It is possible to notice an important share of youth and males in the overall unemployment share and a low sharing of rural areas:

Table 5. Unemployment rate among groups.

Unemployment Rates Among Groups				
	Unemployment Rate by Groups %		Group Share Among Unemployed %	
Period:	**I-2009**	**II-2009**	**I-2009**	**II-2009**
Gender				
Female	5.2	5.0	37.3	35.0
Male	5.3	5.6	62.7	65.0
Age				
15-24	9.6	10.0	38.6	39.3
25-54	4.2	4.2	56.5	55.4
55-64	2.8	3.1	4.9	5.3
Area of residence				
Urban	5.3	5.3	92.7	90.9
Rural	4.7	5.8	7.3	9.1
Level of education				
Incomplete elementary	3.3	3.4	9.3	9.0
Elementary	5.0	4.7	21.0	18.9
High school	5.9	6.1	38.5	39.4
Technical	6.7	6.9	18.7	19.3
Graduate	4.5	4.8	12.5	13.4

Source: ENOE-INEGI. I-2009 and II-2009.

Despite almost all groups obtained an increase on their unemployment rates, It can be also inferred that males, youths, urban inhabitants and technicians who mostly seek jobs. Since the share of workers holding high school education is the largest group of employees in the Mexican labor markets, they represent the highest rate of the unemployment share among level of education groups.

One starting points of the previous theoretical review is the job search stage and the information sources that job seekers account on. The ENOE enquires what actions they do to find out the information on vacancies, which are grouped into direct and indirect mechanisms differed by how costly they are (recall that direct mechanisms are the least costly).

Table 6. Unemployed workers - job search mechanisms.

Unemployed workers - job search mechanisms		Type	I-2009	II-2009
1	Directly at the work place	Direct	75.4%	76.2%
2	At a private labor office	Indirect	2.2%	1.7%
3	At a public labor office	Indirect	2.0%	1.8%
4	A temporary program of public jobs	Direct	1.3%	1.5%
5	Attempt to start a own business	Direct	0.8%	1.0%
6	On line - Internet	Indirect	7.5%	9.2%
7	A public announcement	Indirect	10.0%	9.2%
8	A labor union or group	Direct	0.7%	0.8%
9	A friend or relative	Direct	16.6%	14.9%
10	A Classified ad	Indirect	2.4%	2.6%
11	Does not know		1.1%	1.4%

Source: ENOE-INEGI. I-2009 and II-2009.

These job search mechanisms are not excluding and hence job seekers might find information on vacancies using one or more sources. It is evident that rationally, Mexicans prefer direct mechanisms that are more restrictive in terms of the flow of information between employers and unemployed workers. For instance, in Table 6 it is seen that a 75.4% and 16.6% of job seekers look for information directly at the work place and

through a friend or relative in I-2009, respectively. Even though labor markets intermediaries are less restrictive in the flow of information, they are barely used in Mexico. However, the Internet has become an important source and alternative for unemployed workers in comparison to alternate government agencies.

Table 7 presents the job search mechanisms of employees before they were hired:

Table 7. Employees – job search mechanisms.

Employed – job search mechanisms		I-2009
1	Directly at the work place	19.9%
2	At a private labor office	0.2%
3	At a public labor office	0.2%
4	A temporary program of public jobs	0.2%
5	A labor union or group	1.2%
6	On line - Internet	1.1%
7	A public announcement	7.8%
8	A friend or relative	56.0%
9	Someone offered a vacant	11.0%
10	Other	1.0%
11	Does not know	1.3%

Source: ENOE-INEGI. I-2009.

Table 6 is compared to Table 7 which presents how employed actually found information about his or her current job in the period I-2009 (not available for II-2009). This contrast demonstrates de effectiveness of direct mechanism during the job search and matching stages, particularly for those who learned about a job through a person close to them. Here the proportions are switched: as consequence of not finding any response directly at the work place, the observed proportion for friends or relatives reaches 56%[48].

Another important issue to address is the distribution of wages among education levels and economic sectors. Tables 6 and 7 present the Gini coefficients of wages earned by employees during the first two quarters of 2009 that are relevant in the study of wage differences. It is worth to notice a systematic increase of the concentration of wages across this period. Agriculture was the only sector with a decrease in the concentration of wages, though it is the most inequitable amongst the others.

Table 8. Gini coefficient by level of education.

Gini Coefficient by Level of Education.		
Level of education	**I-2009**	**II-2009**
Incomplete elementary	38.5	39.2
Elementary	35.1	36.2
High school	34.0	34.3
Technical	36.6	36.0
Graduate	37.2	38.2
Not specified	31.4	24.7
Total	**40.7**	**41.4**

Source: ENOE-INEGI. I-2009 and II-2009.

[48] *See* Appendix 1 for further details in Mexican States.

Table 9. Gini coefficient by sector of economic activity.

Gini Coefficient by Sector of Economic Activity		
Gini coefficient	**I-2009**	**II-2009**
Construction	29.7	30.3
Manufacturing	37.3	38.0
Commerce	41.0	41.5
Services	40.1	40.9
Others	36.9	38.7
Agriculture	44.4	44.1
Total	**40.7**	**41.4**

Source: ENOE-INEGI. I-2009 and II-2009.

The Gini coefficient among levels of education has a U-shaped in Mexico. The distribution of wages for low educated workers is similar to those who have a Graduate level of education, while the Gini coefficient reaches a minimum level for employees with a high school degree.

4. Estimations and results

The estimation of Equation 3 is explained in this section and the estimation of the error terms for the Mexican states is detailed in the Appendix 2. Besides the covariates included in x_{jit} previously mentioned in Section 2.1, it was also incorporated the inverse of Mills ratio (Heckman, 1979) as a correction of the selection of remunerated workers in each period separately[49]. The values of the parameter $\gamma = \sigma_u^2/(\sigma_u^2 + \sigma_v^2)$ were derived in a second stage of a post-estimation procedure of the stochastic frontier on the time-variant model. It is also calculated a pseudo R-squared as a proxy of the determination coefficient for the overall model plus for the estimations amongst states. This is carried out applying the formula:

$$pseudoR^2 = 1 - (lnL_R/lnL_{UR}) \quad (4)$$

Where L_{UR} stands for the log-likelihood of the unrestricted model (only with its constant term) and L_R for the restricted term (including covariates).

4.1. Overall model

The estimations of the model for the nationwide data yielded a pseudo determination coefficient of 19.2%, which means that the power of prediction of the estimated coefficients is not strong. Nevertheless, accordingly with these kinds of studies and the specification methods supports satisfactorily a good fit of the model. In addition, as the results are aimed at explaining a part of the Mexican's wages using a stochastic frontier analysis, the central attention here is paid on the error term. As well, the estimations of Equation 3 present a starting point of what Hofler and Murphy (1994) called *reservation wages* but for the purpose of this document they represent the base line of wages in the first stage of their formation:

[49] The inverse of Mills ratio included the variables gender, marital status, household's head, head's age and education, average of household's age, urban area).

Table 10. Stochastic Frontier Estimates.

Mexico - Stochastic Frontier Estimates			
lnw_i	**Coefficient**	**Std. Error**	
edu_i	0.049	0.00	***
age_i	0.031	0.00	***
age_i^2	0.000	0.00	***
whm_i	0.003	0.00	***
wht_2^i	0.000	0.00	***
$manufacturing_i$	0.127	0.01	***
$services_i$	0.132	0.00	***
$mills_i$	-0.203	0.00	***
$constant$	6.386	0.01	**
η	-0.010	0.00	***
σ^2	0.499	0.01	
γ	0.827	0.00	
σ_u^2	0.412	0.01	
σ_v^2	0.086	0.00	
$pseudoR^2$	19.2%		

*: Significant at 10%; **: Significant at 5%; ***: Significant at 1%.

Source: ENOE-INEGI. I-2009 and II-2009.

It turned out that the estimation of γ leads it to infer that 82.7% of the variation of the heterogeneity among Mexican workers is explained by labor markets imperfections and the rest 17.3%, is due to the exceptional conditions of the workers that are also a part of their human capital. These results also suggest that this variation has a decreasing and significant rate over time since the coefficient η obtained a negative value. These findings advise that in spite of the fact that Mexico counts on several job search mechanisms and strict labor laws (Levy, 2008), if two workers are similar in their skills and qualifications the differences between their wages are not mostly explained by their individual and specific conditions but rather the nature of the labor markets they face. Both start the wage formation process from a similar condition but as long as they interact with employers, match them or receive further adjustments, their wages get different paths over time.

Furthermore, if the statistical inference on η would have yielded a coefficient equals to zero, one had concluded that these variations were expected to be constant over time. Nevertheless, in this case the implications entail that despite the estimations were obtained using a two period data, the perspectives on the prevalence of the market imperfection will fade (ceteris paribus) in the future at a rate of -1% quarterly. As remarked before, under the absence of market imperfections the proportion of wages explained by them would be indeed equals to zero. Since the model has the restriction of non-negative values of u_i, it was not sensible to think of a negative effect of markets on wages below their base line because otherwise workers would remain inactive. Besides, our theoretical framework puts the interaction between job seekers and employers as a wage gaining process that mismatch the productivity of a given worker to the real requirements of a vacancy.

The estimation of the coefficients for the covariates of the model allowed it to estimate the ratio u_i/lnw_i in addition to the single parameter σ_u^2. This average proportion of wage that corresponds to the markets imperfection effect obtained an average of 0.052 and a standard deviation of 0.03. Even though this proportion seems small it determines most of the variation among wages while the idiosyncratic component apparently is not such volatile. The next chart plots the density distribution of u_i/lnw_icovering more than 90% of the individuals between 0 and 0.15:

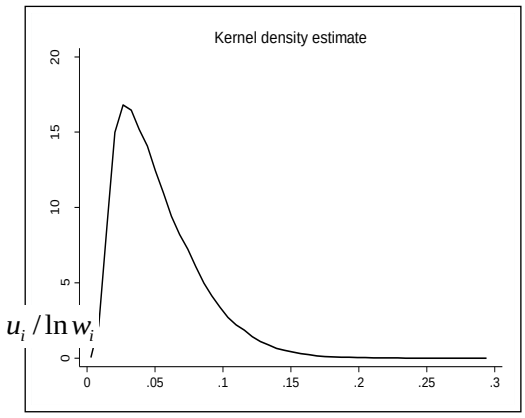


Figure 1. Non-parametric histogram of $u_i / \ln w_i$

Source: ENOE-INEGI. I-2009 and II-2009.

This non-negligible part of the wages explained by factors besides human capital and abilities is also compared to the circumstances of the employee that lead to a large or small u_i. The variations of this parameter depend on the degree in which the markets imperfections were presented during the wage formation process.

As pointed out before, the job search process has an essential role in the determination of wages and the success of job seekers accepting a wage offer above their reservation wage. It is now presented the partial correlations of u_i with respect to covariates that theoretically may mitigate or spur the presence of market imperfections:

Table 11. Partial correlations for u_i.

	Partial correlations for u_i		
Job search	Union member	0.0718	***
	Written contract	0.0424	***
	Directly at the work place	-0.0128	***
	At a private labor office	0.005	
	At a public labor office	-0.0115	***
	A temporary program of public jobs	-0.016	***
	A labor union or group	0.0107	***
	On line - Internet	0.0308	***
	A public announcement	-0.0291	***
	A friend or relative	-0.004	
	Someone offered a vacant	0.0098	***
	Other	0.0145	***
Firm size	1 person	-0.0546	***
	2 - 5 people	-0.0382	***
	6 - 10 people	-0.0086	**
	11 - 15 people	-0.001	
	16 - 20 people	0.0061	*
	21 - 30 people	0.0107	***
	31 - 50 people	0.0114	***
	51 - 100 people	0.0168	***
	101 - 250 people	0.0192	***
	251 - 500 people	0.0187	***
	500 and more people	0.0513	***

*: Significant at 10%; **: Significant at 5%; ***: Significant at 1%.

Source: ENOE-INEGI. I-2009.

In spite of the fact that the correlations are low, they manifest the direction of the behavior of each covariate jointly with u_i. For instance, besides having a written contract, being part of a labor union presented de higher positive correlation and leads the differences among wages in the third stage of wage formation. As mentioned, unions often exercise power to increase wages even if the real productivity of a given employee does not match his or her wage.

Moreover, the job search mechanisms give us the chance to infer on the evolution of wages from the job search stage. The analysis is divided into those sources with a least restrictive information flow and the others that enable the existence of asymmetric information. In this manner, even though asking for information to friends or relatives obtained a tiny and non-significant and negative partial correlation, the signs of the other search mechanisms were those expected and predictable by the theory.

The firm size draws a singular attention because it has a negative correlation among small ones and a higher positive correlation among the rest. This finding suggest that the firm size is a key factor in the markets imperfection condition in Mexico, and that larger firms measured by the number of workers are correlated with non-competitive outcomes on the markets. As large firms have better possibilities to offer higher wages and to train workers at their own expense, it is even less costly to replace a vacant.

4.2. The State level estimation

Appendix 2 presents the estimations of Equation 3 for Mexican States and makes possible the comparison among them. This sub-national analysis is aimed at understanding particular characteristics of segmented regional labor markets. The regional differences contribute to identify the sources of the labor markets imperfections replicating the previous Table 9.

The description of Table A3 starts firstly focusing on the time-varying parameter η. This estimates divide the Mexican States into three groups where the market imperfections component of wages is fixed over time (e.g. Aguascalientes, Puebla and Tabasco), some other where η determines an increasing u_i (Colima, Chiapas, and Guanajuato) and the rest where it is expected to decrease (Baja California, Chihuahua, Distrito Federal, Morelos, Sonora, Tlaxcala and Veracruz). It is evident that Chihuahua is the State with the highest negative value of the time-varying parameter, which entails a decreasing trend of the market imperfections component for the employees located there. The implications of this division of Mexican States and their characterization of the labor markets would detour our attention, but in any case this could be the beginning of a future follow up estimation of the Equation 3.

The second and most important parameter of the estimation output in Table A3, is the proportion of the variance of the heterogeneity of wages explained by market imperfections, γ. It turned out that most of the Mexican States have a variability of this component relatively high. There are few whose γ is barely lesser than the one obtained with the overall model. This component is compared to Table A2 that contains the proportions of employees that found their current job through a direct or indirect information source (as explained before).

The discussion around chart A1 has a surprising implication in terms of the use of direct and indirect job search mechanisms or information sources. Even though the observed relation between γ and the proportion of employed that found their current job through an indirect source contradicts the theory up to some extend. For low levels of the share of indirect mechanisms the parameter γ takes small values, as seen for Michoacan and Oxaca. Nonetheless, the trend line that was also plotted suggests a decreasing marginal relation between these two variables as long as the indirect share increases, as seen for Quintana and Nuevo Leon.

The States that obtained the maximum and minimum values of γ were Baja California and Oxaca, respectively. Table 9 is now replicated for these two States to address which covariates are more correlated with the component of wages attributable to the market imperfections. These correlations exhibit a common pattern: they are significant and have the same sign for employees belonging to a labor union and holding a written contract, however, the significances of the firm size variables are only relevant for Oxaca.

Table 12. Partial correlations for u_i in Baja California and Oxaca.

Partial correlations for u_i					
		Baja California		Oxaca	
Job search	Union member	-0.1797	***	-0.0903	***
	Written contract	-0.0949	***	-0.0385	*
	Directly at the work place	0.0128		0.0365	*
	At a private labor office	0.0176		0.0312	
	At a public labor office	0.0042		-0.0171	*
	A temporary program of public jobs	0.0241		0.0041	
	A labor union or group	0.0089		0.0237	
	On line - Internet	0.0321	**	0.0599	***
	A public announcement	0.0122		-0.0114	
	A friend or relative	0.0158		0.0259	
	Someone offered a vacant	0.0158		0.0595	***
	Other	0.0042		0.0809	***
Firm size	1 person	-0.008		-0.0724	***
	2 - 5 people	-0.023		-0.0469	**
	6 - 10 people	-0.0159		-0.0136	
	11 - 15 people	-0.0214		-0.0065	
	16 - 20 people	-0.0125		-0.0139	
	21 - 30 people	-0.0165		0.0022	
	31 - 50 people	-0.0129		0.0196	
	51 - 100 people	-0.0122		0.0085	
	101 - 250 people	-0.0169		0.017	
	251 - 500 people	-0.0164		0.0052	
	500 and more people	-0.0165		0.0422	**

*: Significant at 10%; **: Significant at 5%; ***: Significant at 1%.

Source: ENOE-INEGI. I-2009.

Surprisingly the signs of the first two variables are negative, in contrast of those presented in Table 9. Besides, the union membership in Baja California obtained the highest partial correlation coefficient with a negative sign. This suggest that this kind of labor condition decreases the labor imperfection wage u_i.

Regarding the sources of information during the job search stage for employees located in these two regions, it turned out that Internet is highly correlated with the market imperfection gain in both States. A possible conjecture for this finding is the existence of only one selected group of employees that have access to computers or afford all the cost that it takes to look for information online. Furthermore, the only one job search mechanism and that makes de difference between these two States is the information consultancy at public labor offices in Oxaca, which with a significance of 10% it obtained a negative sign (it is negatively correlated related with u_i)

Finally, it is pointed out the importance of the firm size in Oxaca, where the prevalence of small or single worker firms leads a negative partial correlation in comparison to the rest of the firm sizes. As exception, larger firms yielded a positive sign, as seen for the overall nationwide model.

5. Conclusions

This document presented an attempt to link the job search and job matching theory as a background of the wage formation process divided into three stages. Beginning with the reservation wage resulting from the labor supply and job search plans of potential workers; a subsequent bargaining moment and a further adjustment once they are completely hired.

The National Survey on Employment from Mexico was explored to find the evidence of labor markets imperfection plausibly presented along those three stages. Adapting former studies of reservation wage estimation it was possible to subtract the part of the observable wages that corresponds to the wage gain of workers due to markets imperfections. It was estimated that 82.7% of the variation of unexplained wages by human capital among workers are caused by asymmetric information, monopsonic power, discrimination, and so on.

The estimations were also obtained for Mexican States to identify which of them presents the highest and lowest variation of wages due to the presence of markets imperfections. This exercise led the analysis to deeply explore the relations between some theoretical variables that enable the existence of a proportion of wages that are not explained by human capital or abilities. It turned out that public labor offices have an important diminishing role on this proportion in the State of Oxaca. At the same time, it was observed that Internet is quite related to the augmentation of the mentioned wage gain.

It is important for policy makers to understand de dynamic of these labor market imperfections in terms of the firm size (although large firms are more susceptible to be regulated). Indeed, any intervention to mitigate this situation must not only count on workers' characteristics but also on the conditions of the potential working place.

References

[1] Arnott, R., Hosios, A., and Stiglitz, J. 1988. Implicit Contracts, Labor Mobility, and Unemployment. *The American Economic Review*, Vol. 78, No. 5: 1046-1066.

[2] Arnott, R., and Stiglitz, J. 1985. Labor Turnover, Wage Structures, and Moral Hazards: The Inefficiency of Competitive Markets. *Journal of Labor Economics*, Vol. 3, No. 4: 434-462.

[3] Aslund, O., and Rooth, D. 2005. Shifts in Attitudes and Labor Market Discrimination: Swedish Experiences After 9-11. *Journal of Population Economics*, Vol. 18, No. 4: 603-629.

[4] Baldwin, M. 1991. An Asymmetric Information Theory of Labor Market Discrimination. *Southern Economic Journal*, Vol. 57, No. 4: 1148-1154.

[5] Bartolucci, C. 2009. *Gender Wage Gaps Reconsidered: A Structural Approach Using Matched Employer-Employee Data*. Collegio Carlo Alberto. Notebook No. 116. September.

[6] Bemmels, B. 1987. How Unions Affect Productivity in Manufacturing Plants. *Industrial and labor Relations Review*, Vol. 40, No. 2: 241-253.

[7] Blau, D. 1991. Search for Nonwage Job Characteristics: A Test of the Reservation Wage Hypothesis. *Journal of Labor Economics*, Vol. 9, No. 2: 196-205.

[8] Bloemen, H. 2005. Job Search, Search Intensity, and Labor Market Transitions an Empirical Analysis. *The Journal of Human Resources*, Vol. 40, No. 1: 231-269.

[9] Boehm, T., and Hofler, R. 1987. A Frontier Approach to Measuring the Effect of Market Discrimination: A Housing Illustration. *Southern Economic Journal*, Vol. 54, No. 2: 301-315.

[10] Card, D. 2001. The Effect of Unions on Wage Inequality in the U.S. Labor Market. *Industrial and Labor Relations Review*, Vol. 54, No. 2: 296-315.

[11] Carroll, S., and Rolph, J. 1973. A Stochastic Model of Discrimination in the Labor Market. *Econometrica*, Vol. 41, No. 1: 97-108.

[12] Cawley, J., Heckman, J., and Vytlacil, E. Three Observations on Wages and Measured Cognitive Ability. *Labour Economics*, Vol. 8, No. 4: 419-442.

[13] Cobb-Clark, D., and Tan, M. 2009. Noncognitive Skills, Occupational Attainment, and Relative Wages. *Centre for Economic Policy Research. Discussion Paper* No. 612. July.

[14] David, H. 2008. The Economics of Labor Market Intermediation: An Analytic Framework. *Working paper* 14348. National Bureau of Economic Research, September.

[15] Fields, G. 2007. Labor Market. *ILR Collection Working Papers*. Cornell University ILR School.

[16] Gannon, F., and Zenou, Y. 1997. Segregation and Labor Discrimination in Cities. *Annales d'Economie et de Statistique*, No. 45: 233-249.

[17] Greene, W. 2003. *Distinguishing Between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Hearth Organization's Panel Data on National Health Care Systems*. Department of Economics, Stern School of Business, New York University.

[18] Heckman, J. 1998. Detecting Discrimination. *The Journal of Economic Perspectives*, Vol. 12, No. 2: 101-116.

[19] Heckman, J. 1979. Sample Selection Bias as a Specification Error. *Econometrica*, 47:153-161.

[20] Heckman, J., Stixrud, J., and Urzua, S. 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, Vol. 24, No. 3: 411-482.

[21] Heckman, J., and Willis, R. 1979. The Distribution of Lifetime Labor Force Participation of Married Women: Reply to Mincer an Ofek. *The Journal of Political Economy*, Vol. 87, No. 1: 203-211.

[22] Hofler, R., and Murphy, K. 1994. Estimating Reservation Wages of Employed Workers Using a Stochastic Frontier. *Southern Economic Journal*, Vol. 60, No. 4: 961-975.

[23] Kahn, L., and Low, S. 1990. The Demand for Labor Market Information. *Southern Economic Journal*, Vol. 56, No. 4: 1044-1058.

[24] Kiefer, N., and Neumann, G. 1979. An empirical job-search model, with a test of the constant reservation wage hypothesis. *Journal of Political Economy*, Vol. 87, No. 1: 89-107.

[25] Kolpin, V., and Singell, L. 1997. Information, Strategic Behavior, and Discrimination in the Labor Market. *Economic Theory*, Vol. 10, No. 1:175-184.

[26] Mavromaras, K. *et al*. 2009. Job Mismatches and Labour Market Outcomes. *ESRI, Working Paper* No. 314. September.

[27] Mincer, J. 1974. *Learning Schooling and Earnings*. National Bureau of Economic Research.

[28] Moffitt, R., and Wolfe, B. 1992. The Effect of the Medicaid Program on Welfare Participation and Labor Supply. *The Review of Economics and Statistics*, Vol. 74, No. 4: 615-626.

[29] Mortensen, D. 1989. The Persistence and Indeterminacy of Unemployment in Search Equilibrium. *The Scandinavian Journal of Economics*, 90th Anniversary Symposium: Whither Macroeconomics? Vol. 91, No. 2: 347-370.

[30] Owen, John D. 1971. *The demand for Leisure*. The Review of Political Enomy, Vol. 79, No. 1: 56-76.

[31] Rogerson, R., Shimer, R., and Wright, R. 2005. Search-Theoretic Models of the Labor Market: A Survey. *Journal of Economic Literature*, Vol. 43, No. 4: 959-988.

[32] Salop, J., and Salop, S. 1976. Self-Selecion and Turnover in the Labor Market. *The Quarterly Journal of Economic*, Vol. 90, No. 4: 619-627.

[33] Stigler, G. 1962. *Information in the Labor Market*. The Journal of Political Economy, Vol. 70, No. 5, Part 2: Investment in Human Beings: 94-105.

[34] Swinton, D. 1977. A Labor Force Competition Theory of Discrimination in the Labor Market. *The American Economic Review*, Vol. 67, No. 1, Papers and Proceedings of the Eight-ninth Annual Meeting of the American Economic Association: 400-404.

[35] United Nations Department of Economic and Social Affairs. Population Division. 2009. *World Population Prospects: the 2008 Revision*. March, 2009. Appendix VII.15.

[36] Van den Berg, G. 2003. Multiple Equilibria and Minimum Wages in Labor Markets with Informational Frictions and Heterogeneous Production Technologies. *International Economic Review*, Vol. 44, No. 44: 1337-1357.

[37] Wessels, W. 1985. *The Effects of Employment and Productivity: An Unresolved Contradiction*. Journal of Labor Economics, Vol. 3, No. 1: 101-108.

Appendix 1

Table A1. Unemployed workers - job search mechanisms.

Mexican States - Proportion of unemployed workers by search mechanism				
Period:	**I-2009**		**II-2009**	
State	**Direct**	**Indirect**	**Direct**	**Indirect**
Aguascalientes	71.1%	27.5%	72.0%	27.3%
Baja California	87.6%	10.8%	87.2%	8.2%
Baja California Sur	80.5%	14.5%	91.1%	8.0%
Campeche	91.7%	3.9%	95.8%	3.2%
Coahuila de Zaragoza	91.4%	6.8%	86.1%	9.7%
Colima	88.7%	9.6%	89.8%	9.2%
Chiapas	95.2%	4.5%	93.0%	6.4%
Chihuahua	86.5%	6.1%	82.6%	6.4%
Distrito Federal	64.8%	30.4%	68.5%	30.5%
Durango	90.9%	8.6%	90.4%	8.9%
Guanajuato	76.8%	21.6%	81.9%	14.3%
Guerrero	83.9%	9.0%	91.7%	6.5%
Hidalgo	81.3%	17.6%	79.2%	19.8%
Jalisco	75.5%	22.3%	74.3%	22.1%
México	79.8%	18.6%	77.0%	21.5%
Michoacán de Ocampo	93.2%	5.9%	95.6%	3.4%
Morelos	79.1%	19.3%	80.3%	17.8%
Nayarit	89.9%	5.6%	90.8%	6.5%
Nuevo León	70.8%	24.9%	73.3%	22.6%
Oaxaca	93.5%	4.8%	93.0%	6.2%
Puebla	76.2%	22.7%	74.9%	24.1%
Querétaro Arteaga	76.5%	21.0%	77.5%	20.3%
Quintana Roo	65.1%	30.1%	74.0%	24.2%
San Luís Potosí	90.3%	8.5%	85.6%	14.1%
Sinaloa	85.2%	9.2%	80.2%	9.3%
Sonora	79.7%	18.5%	82.0%	14.3%
Tabasco	91.6%	4.9%	92.8%	5.0%
Tamaulipas	83.6%	14.5%	84.6%	11.6%
Tlaxcala	92.9%	6.5%	91.8%	4.7%
Veracruz Ignacio de la Llave	84.2%	14.5%	78.2%	17.0%
Yucatán	77.0%	16.7%	78.5%	20.0%
Zacatecas	90.8%	7.6%	91.1%	5.9%

Source: ENOE-INEGI. I-2009 and II-2009.

Table A2. Unemployed workers - job search mechanisms.

Mexican States - Proportion of employed by search mechanism		
Period:	**I-2009**	
State	**Direct**	**Indirect**
Aguascalientes	89.9%	10.1%
Baja California	89.0%	11.0%
Baja California Sur	93.7%	6.3%
Campeche	92.2%	7.8%
Coahuila de Zaragoza	92.0%	8.0%
Colima	95.4%	4.6%
Chiapas	96.1%	3.9%
Chihuahua	83.0%	17.0%
Distrito Federal	87.1%	12.9%
Durango	94.2%	5.8%

Guanajuato	90.7%	9.3%
Guerrero	94.8%	5.2%
Hidalgo	93.6%	6.4%
Jalisco	89.8%	10.2%
México	88.1%	11.9%
Michoacán de Ocampo	96.6%	3.4%
Morelos	92.1%	7.9%
Nayarit	95.8%	4.2%
Nuevo León	85.6%	14.4%
Oaxaca	94.4%	5.6%
Puebla	90.8%	9.2%
Querétaro Arteaga	88.8%	11.2%
Quintana Roo	86.3%	13.7%
San Luís Potosí	89.7%	10.3%
Sinaloa	93.1%	6.9%
Sonora	92.9%	7.1%
Tabasco	94.6%	5.4%
Tamaulipas	92.7%	7.3%
Tlaxcala	92.4%	7.6%
Veracruz Ignacio de la Llave	93.5%	6.5%
Yucatán	90.0%	10.0%
Zacatecas	95.8%	4.2%

Source: ENOE-INEGI. I-2009 and II-2009.

Appendix 2.

Table A3. Unemployed workers - job search mechanisms.

Mexican States - Error term estimates						
State	σ_u^2	σ_v^2	γ	η		pseudo R^2
Aguascalientes	0.3833	0.0634	0.8580	-0.0019		0.2329
Baja California	6.4440	0.0870	0.9867	-0.0529	**	0.2839
Baja California Sur	0.3099	0.1039	0.7489	-0.0008		0.1574
Campeche	0.5076	0.0958	0.8413	0.0075		0.1936
Coahuila de Zaragoza	2.0589	0.0811	0.9621	-0.0032		0.1861
Colima	0.2480	0.0871	0.7400	0.0608	**	0.2241
Chiapas	0.2680	0.0847	0.7599	0.0611	**	0.2778
Chihuahua	2.1359	0.1051	0.9531	-0.1493	***	0.2257
Distrito Federal	0.6448	0.1029	0.8624	-0.0539	**	0.1911
Durango	0.4924	0.0829	0.8559	-0.0368		0.2236
Guanajuato	2.7899	0.0754	0.9737	0.0595	**	0.3110
Guerrero	0.2751	0.0833	0.7676	0.0088		0.1330
Hidalgo	0.7335	0.1039	0.8759	-0.0229		0.2123
Jalisco	1.1176	0.0778	0.9349	-0.0356		0.2489
México	2.5794	0.0915	0.9658	-0.0441		0.2506
Michoacán de Ocampo	0.1646	0.0749	0.6874	-0.0237		0.2485
Morelos	0.8284	0.0697	0.9224	-0.0729	**	0.2179
Nayarit	0.1650	0.0731	0.6931	-0.0093		0.2128
Nuevo León	0.8282	0.0773	0.9147	-0.0181		0.2133
Oaxaca	0.1292	0.0657	0.6629	-0.0143		0.2582
Puebla	0.5356	0.0851	0.8629	-0.0312		0.2180
Querétaro Arteaga	1.2001	0.0774	0.9394	-0.0131		0.2812
Quintana Roo	0.4988	0.1056	0.8252	0.0249		0.1584
San Luís Potosí	0.8415	0.0966	0.8970	-0.0155		0.1955
Sinaloa	0.3179	0.0788	0.8013	-0.0071		0.2063
Sonora	1.4101	0.1120	0.9264	-0.0479	*	0.2145
Tabasco	0.4073	0.0979	0.8062	0.0175		0.2170
Tamaulipas	0.3953	0.0953	0.8057	0.0178		0.1665
Tlaxcala	0.3633	0.0711	0.8363	-0.0796	***	0.2346
Veracruz Ignacio de la Llave	0.6104	0.0766	0.8885	-0.0330	**	0.1792
Yucatán	0.5364	0.0706	0.8838	0.0020		0.1771
Zacatecas	0.4047	0.0856	0.8255	-0.0111		0.2161

*: Significant at 10%; **: Significant at 5%; ***: Significant at 1%.

Source: ENOE-INEGI. I-2009 and II-2009.

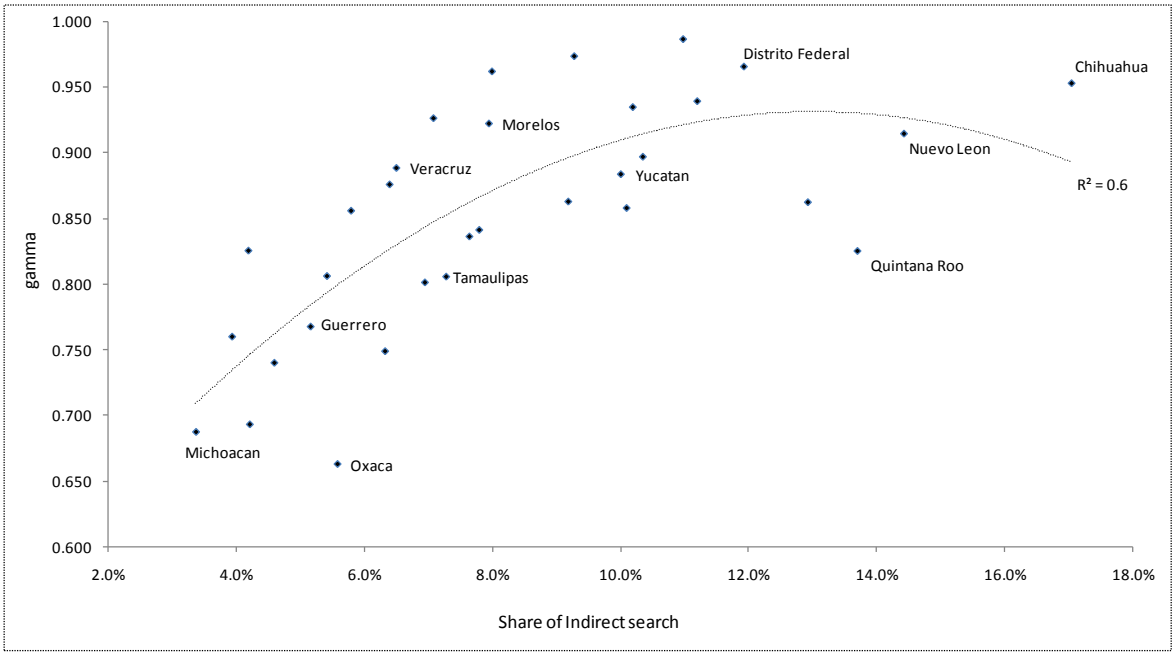


Chart A1. Plot of gamma and share of indirect search mechanisms

Source: ENOE-INEGI. I-2009 and II-2009.

Call for Papers

*A*ssociation for *S*ustainable *E*ducation, *R*esearch and *S*cience

has the honour to invite you to participate at

The First Online International Conference on "Convergence of Law and Economics" in 31st, July, 2010

The Law and Economics movement is one of the most important intellectual developments within legal scholarship in recent decades. This approach applies economic concepts to explain and clarify legal issues and aspects, not only with respect to narrowly defined "economic" fields, such as competition and economic law, but also to legal fields that concern a wide variety of non-market activities, from liability issues to family matters and crime. Thus, Law and Economics has become a central methodology for both positive and normative legal analysis; already it has influenced regulation, legislation, and case law. Many academic journals publish articles in the field, and Law and Economics has become an integral part of legal and economic education at the most prestigious universities.

The annual conference of **ASERS** dedicated to Law and Economics – *Convergence of Law and Economics* – intend to become an important forum for the exchange of research findings and ideas. The conference has the purpose of providing assistance to law and economics scholars and bringing their scholarship to a wider audience, including policy makers, legislators and judges and intend to encourage further research in the different areas of Law and Economics such as: Behavioural Law and Economics, Commercial Law, Contract Law, Criminal Law, Environmental Law and Economics, Experimental Law and Economics, Family Law, Gender and Discrimination, Health Law, Intellectual Property and Innovation Policy, Labour and Employment Law, Litigation, Dispute Resolution and the Legal Process, Property Law, Public and Administrative Law, Tort Law and Compensation.

We invite to submission original research contributions describing new results, original ideas and applications related to the topics of the conference. Papers should be submitted electronically at these e-mail addresses conferences@asers.eu and CC to constantinescu_madalina2002@yahoo.co.uk in MS Word and also in PowerPoint (*See* Instructions for Authors).

All papers will be reviewed and published in the Conference Proceeding under an ISBN reference on CD. The Proceeding will be indexed and listed in various reference search engines. The best papers selected by the Program Committee will be published in **Journal of Advanced Research in Law and Economics** after a double-blind peer-reviewing and the payment of 100€ as submission fee charged by the journal.

Important Dates:

- **10 July, 2010** – Abstract submission deadline;
- **15 July, 2010** – Notification of acceptance/rejection;
- **25 July, 2010** – Deadline for payments: 50€ for attendance at Conference;
- **25 July, 2010** – Full paper submission in MS Word and PowerPoint format;
- **31 July, 2010** – On line International Conference.

Call for Papers

*A*ssociation for *S*ustainable *E*ducation, *R*esearch and *S*cience

has the honour to invite you to participate at

The Second Online International Conference on „Competitiveness and Economic Development: Challenges, Goals and Means in a Knowledge based Society"

in 1st, December, 2010

The main goal of this Conference is to provide an opportunity for scholars and professionals from various economic, management, business related fields from all over the world to provide an interdisciplinary forum for discussion and debate about competitiveness and economic development most vital issues.

The Program Committee seeks contributions, which topics include, but are not limited to:

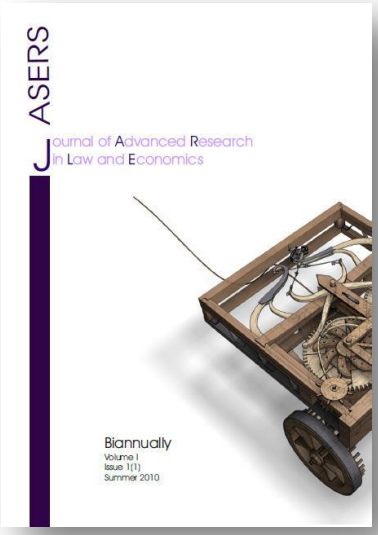
- Competitiveness of businesses and environment in a knowledge based economy;
- Sustainable development and ccompetitiveness on micro, macro and international level;
- Governance and development in the knowledge-based society ;
- Competition, innovation and growth;
- Competitively of smes;
- Competitiveness based on quality and innovation;
- Applied quantitative methods for development;
- New technologies for competition and development;
- Trade and services, opportunities and challenges;
- Elearning for knowledge-based society;
- *….Any other related topic.*

We invite to submission original research contributions describing new results, original ideas and applications related to the topics of the conference. Papers should be submitted electronically at these e-mail addresses conferences@asers.eu and CC to andystefanescu73@yahoo.com in MS Word and also in PowerPoint (*See* Instructions for Authors).

All papers will be reviewed and published in the Conference Proceeding under an ISBN reference on CD. The Proceeding will be indexed and listed in various reference search engines The best papers selected by the Program Committee will be published in Journal of Advanced Research in Management, Journal of Advanced Studies in Finance or Theoretical and Practical Research in the Economic Fields after a double-blind peer-reviewing and the payment of 100€ as submission fee charged by the journals.

Important Dates:

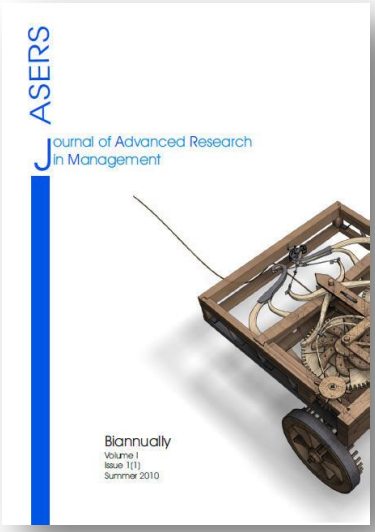
- **1 October, 2010** – Abstract submission deadline;
- **15 October, 2010** – Notification of acceptance/rejection;
- **1 November, 2010** – Deadline for payments: 50€ for attendance at Conference;
- **10 November, 2010** – Full paper submission in MS Word and PowerPoint format;
- **1 December, 2010** – On line International Conference.



Journal of Advanced Research in Law and Economics – Biannually

Editor in Chief: PhD Mădălina **Constantinescu**
Co-Editors: PhD Russell **Pittman** and PhD Eric **Langlais**

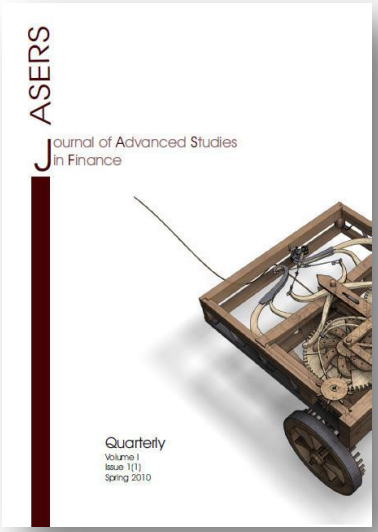
Journal of Advanced Research in Law and Economics provides readers with high quality and empirical research in law and economics. The *Journal* publishes analytical studies on the impact of legal interventions into economic processes by legislators, courts and regulatory agencies. Finally, important 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**,** starting with its first issue, will be indexed in CEEOL , IndexCopernicus 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 starting with its first issue, will be indexed in CEEOL, IndexCopernicus, and EBSCO databases.
Web: http://www.asers.eu/journals/jarm.html email: jarm@asers.eu

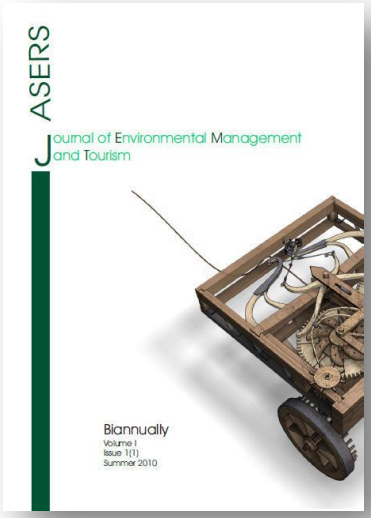


Journal of Advanced Studies in Finance – Biannually

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, behavioural 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.

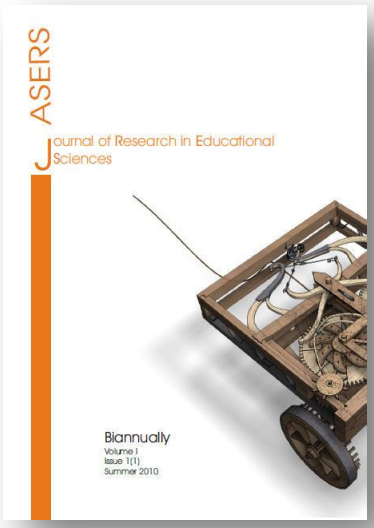
Journal of Advanced Studies in Finance, starting with its first issue, will be indexed in CEEOL, IndexCopernicus and EBSCO databases.
Web: http://www.asers.eu/journals/jasf.html email: jasf@asers.eu



Journal of Environmental Management and Tourism – Biannually

Editor in Chief: PhD Cristina **Barbu**

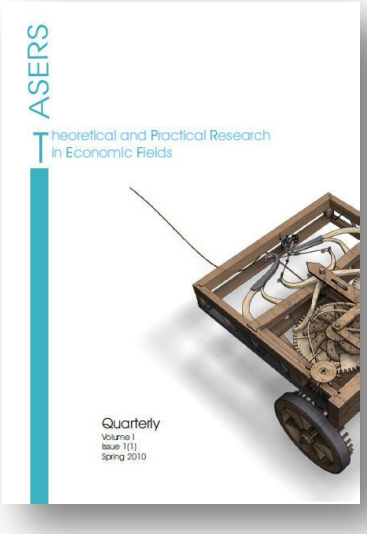
Journal of Environmental Management and Tourism will publish original research and seeks to cover a wide range of topics regarding environmental management and engineering, environmental management and health, environmental chemistry, environmental protection technologies (water, air, soil), pollution reduction at source and waste minimization, energy and environment, modelling, simulation and optimization foenvironmental protection; environmental biotechnology, environmental education and sustainable development, environmental strategies and policies, etc.
Journal of Environmental Management and Tourism starting with its first issue, will be indexed in CEEOL and very soon in IndexCopernicus 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 scholary 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 boundries of knowledge in education and are focusing on research, theory, current issues and applied practice in this area.
Journal of Research in Educational Sciences starting with its first issue, will be indexed in CEEOL, IndexCopernicus 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 starting with its first issue, will be indexed in CEEOL, IndexCopernicus and EBSCO databasesWeb: 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