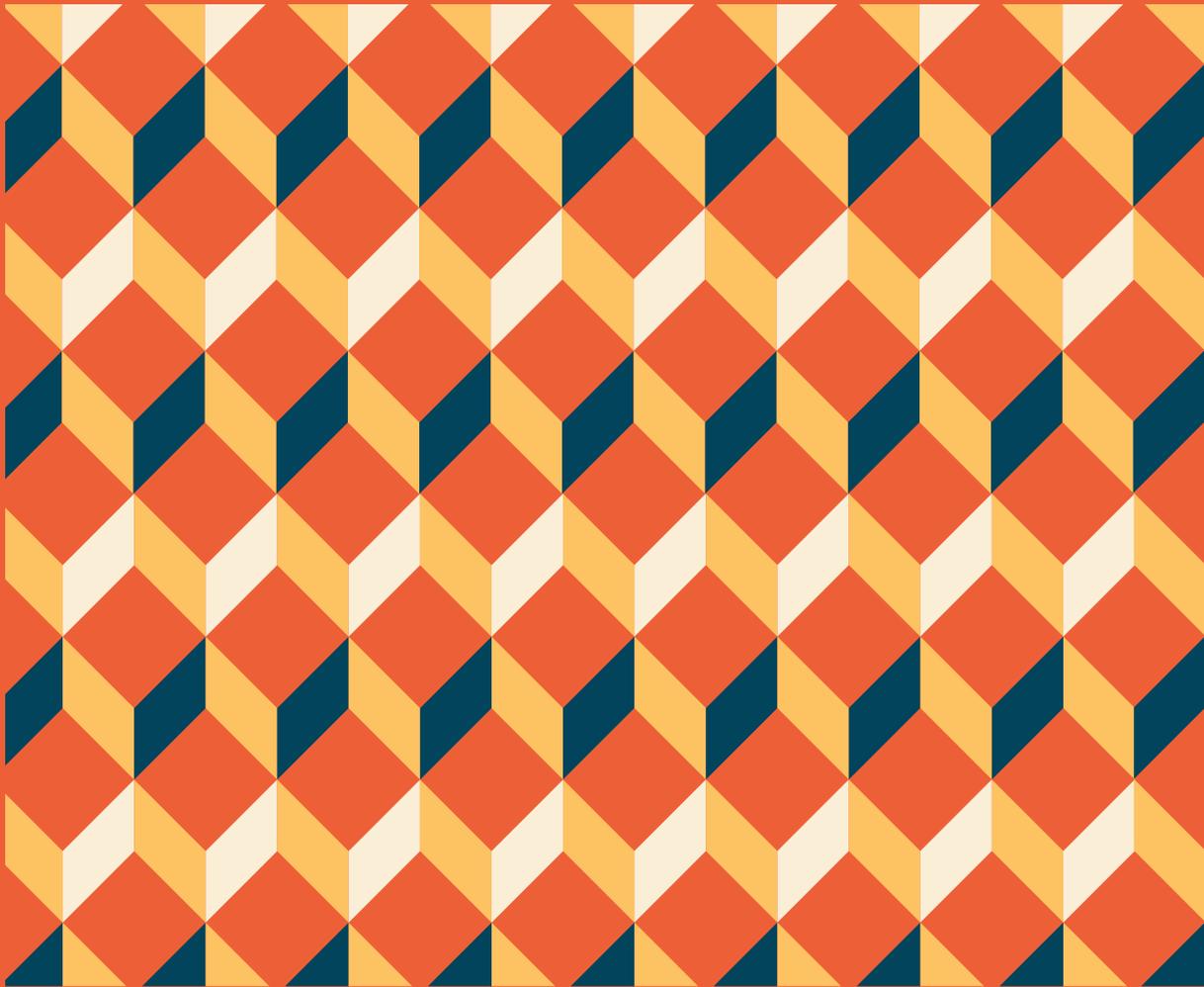Edited by
Pasquale De Muro, Saverio M. Fratini and Alessia Naccarato

ECONOMICS, POLICY AND LAW
Proceedings of the Research Days
Department of Economics

1 Collana
Dipartimento di Economia

ECONOMICS, POLICY AND LAW

Proceedings of the Research Days
Department of Economics

Edited by
Pasquale De Muro, Saverio M. Fratini and Alessia Naccarato

1 Collana
Dipartimento di Economia

Contents

Introduction

Department of Economics Research Days 2019

P. De Muro, S.M. Fratini and A. Naccarato

This book collects some of the papers presented during the Research Days held at the Department of Economics of Roma Tre University, on May 16 and 17, 2019. Although this was the first time that such an initiative was organized in our Department, thanks to our colleagues we found a remarkable success of participation. Besides two plenary sessions – opening and closing – eight parallel sessions took place, during which as many as 30 contributions were presented. Below is a list of the sessions and papers that were presented in each of them.

A.1 – *Fiscal Policy*
Chairperson: A. Di Maio
- G. Bloise, H. Polemarchakis, Y. Vailakis – *Sustainable debt*
- A. Baldini and M. Causi – *Fiscal multipliers of public consumption in Italy*
- R. Ciccone – *On the social irrationality of the limits to public debt*
- M. Deleidi, F. Iafrate and E.S. Levrero – *Public investment fiscal multipliers: An empirical assessment for the European Countries*

A.2 – *Policy and Integration*
Chairperson: F. De Filippis
- R. Crescenzi, G. de Blasio and M. Giua – *Cohesion Policy incentives for collaborative industrial research: evaluation of a Smart Specialisation forerunner programme*
- F. Benassi and A. Naccarato – *The foreign presence in Italy: characteristics of territorial integration*

- V. Piergigli – *The integration of TCN's in the supranational law: limits and potentiality of the European Union*
- S. Terzi and F. Petrarca – *A multidimensional performance indicator to evaluate the effectiveness of Italian university education*

B.1 – *Income Distribution*
Chairperson: S. Fadda
- T. Bechini and P. De Muro – *A multidimensional analysis of inequality in Italy, before and after the Great Recession*
- S.M. Fratini – *Incomes from capital in alternative economic theories*
- E. Pierucci – *Redistribution and international/intranational risk sharing: channels and determinants*
- F. Gastaldi, P. Liberati, E. Pisano and S. Tedeschi – *Regressivity-Reducing VAT Reforms*

B.2 – *Finance and Banking*
Chairperson: F. Fiordelisi
- M. Atripaldi – *Supervision by Bundesanstalt für Finanzdienstleistungsaufsicht (Bafin) of financial and banking activities in Germany*
- G. Scarano – *Financialisation of non-financial corporations, globalization and capital accumulation*
- M. Tirelli – *On the optimal investment finance of small businesses*
- D. Venanzi – *What are the risk drivers in banking system? The fundamental betas of European banks*

C.1 – *Macroeconomics*
Chairperson: R. Ciccone
- M.C. Cucciniello, M. Deleidi and E.S. Levrero – *The price puzzle for the US economy: an empirical assessment of the cost channel*
- M. Deleidi, W. Paternesi Meloni and A. Stirati – *Structural change, labour productivity and the Kaldor-Verdoorn law: evidence from European countries*

- A. Trezzini and A. Palumbo – *A historical analysis of the debate on capacity adjustment in the 'modern classical approach': dealing with complexity in the theory of growth*
- D. Girardi, W. Paternesi Meloni and A. Stirati – *Persistent effect of autonomous demand expansions*

C.2 – *Energy and Environment*
Chairperson: V. Termini
- F. Benedetto, L. Mastroeni and P. Vellucci – *Do oil and agricultural commodity prices co-move?*
- V. Costantini, F. Crespi, E. Paglialunga and G. Sforna – *System transition and structural change processes in the energy efficiency of residential sector: evidence from EU countries*
- F. Longobucco – *Ecological contract and green economy*
- L. Mastroeni and P. Vellucci – *Replication in Energy Markets: Use and Misuse of Chaos Tools*

D.1 – *Bayesian Networks and Sentiment Analysis*
Chairperson: F. Lagona
- F. Benedetto – *Sentiment Analysis for Brand Monitoring in Twitter Social Streams*
- C. Conigliani, T. Petitti and V. Vitale – *Bayesian networks for the analysis of inpatient admissions*
- D. Marella and P. Vicard – *Toward an Integrated Bayesian Network Approach to Measurement Error Detection and Correction*

D.2 – *International Trade*
Chairperson: A. Simonazzi (Sapienza – Università di Roma)
- M. Agostino, A. Giunta, D. Scalera and F. Trivieri – *Italian firms in global value chains: Updating our knowledge*
- J. Baliè, D. Del Prete, E. Magrini, P. Montalbano and S. Nenci – *Trade Policy, Global Value Chains and Developing Countries*
- S. Nenci and L. Salvatici – *New features, forgotten costs and counterfactual gains of the international trading system*

The ten contributions published here, therefore, represent just over a third of the total amount of works discussed during the Research Days. So, on the one hand, we thank the authors for responding to the call-for-papers from which this volume arose – thus renouncing to an editorial collocation of their works that could be most valuable according to the current evaluation criteria of research and scholars. On the other, we regret that some sessions are not represented here.

However, the papers published here cover a wide range of topics, concerning theoretical and applied economics, economic policy, statistical methods, institutional and legal issues. Baldini and Causi (ch.1) and Ciccone (ch.3) deal with fiscal policy from both an empirical and a theoretical point of view. The economic and statistical assessment of relevant social phenomena and public interventions is the topic considered by Benassi and Naccarato (ch. 2) and Petrarca and Terzi (ch. 8). Fratini (ch. 4) and Trezzini and Palumbo (ch. 9) provide contributions in the field of economic theory, at both a micro and macro level. Longobucco's paper (ch. 5) tackles some legal aspects related to ecological economics. European integration is the topic of Piergigli's paper (ch. 6). Finally, issues related to finance and banking system are addressed in the papers by Scarano (ch. 7) and Venanzi (ch. 10).

All the papers have benefited from a peer-review process.

Rome, October 2020

List of contributors

Baldini, Andrea. Department of Economics, Roma Tre University. andrea.baldini86@gmail.com

Benassi, Federico. Italian National Institute of Statistics, Rome. benassi@istat.it

Causi, Marco. Department of Economics, Roma Tre University. marco.causi@uniroma3.it

Ciccone, Roberto. Department of Economics, Roma Tre University. roberto.ciccone@uniroma3.it

De Muro, Pasquale. Department of Economics, Roma Tre University. pasquale.demuro@uniroma3.it

Fratini, Saverio M. Department of Economics, Roma Tre University. saveriomaria.fratini@uniroma3.it

Longobucco, Francesco. Department of Economics, Roma Tre University. francesco.longobucco@uniroma3.it

Naccarato, Alessia. Department of Economics, Roma Tre University. alessia.naccarato@uniroma3.it

Palumbo, Antonella. Department of Economics, Roma Tre University. antonella.palumbo@uniroma3.it

Petrarca, Francesca. Department of Methods and Models for Economy, Territory and Finance MEMOTEF, 'La Sapienza' University of Rome. francesca.petrarca@uniroma1.it

Piergigli, Valeria. Department of Economics, Roma Tre University. valeria.piergigli@uniroma3.it

Scarano, Giovanni. Department of Economics, Roma Tre University. giovanni.scarano@uniroma3.it

Terzi, Silvia. Department of Economics, Roma Tre University. silvia.terzi@uniroma3.it

Trezzini, Attilio. Department of Economics, Roma Tre University. attilio.trezzini@uniroma3.it

Venanzi, Daniela. Department of Economics, Roma Tre University. daniela.venanzi@uniroma3.it

Chapter 1

Fiscal multipliers of public consumption in Italy

Andrea Baldini and Marco Causi

1.1 Summary and introduction[1]

In this contribution we study Italian public spending multipliers for the period 1998-2014 and we estimate the effects of final government consumption shocks on real private Gdp in a quarterly framework.

The discussion about fiscal multipliers (FMs) is related to the effectiveness of active fiscal policies and it has revived after the blast of the Great Recession. Before 2008 the dominant opinion was that active fiscal policies were ineffective as a consequence of a Barro-Lucas effect of crowding out consumption and investment expenditure of the private sector. The Great Recession has undermined this conventional wisdom. Fiscal policy has been again proposed and implemented as a tool to foster economic activity, in particular when monetary policy loses effectiveness at the zero interest lower bound, or criticised for the underestimated effects on real economy of fiscal consolidations (DeLong and Summers 2012; Blanchard and Leigh 2013). With FMs we measure costs and benefits of active (expansive or contractionary) fiscal policies. This political question has driven a flourishing of studies for their empirical estimation.

The multiplier is by no means a single figure, but rather varies not only between countries but also within the same economy (Parker 2011). The measure of Gdp reaction to a fiscal shock may be influenced by the structural and cyclical conditions of the economy, and for this reason could change over time. FM is not a fixed parameter: it is contingent on the state of the economy. For this reason it is important to focus the attention and the research strategies not only on the size of FMs but also on the path that they have during time. There is a majority consensus in literature that FMs are higher during recessions but lower, also in recession, when conditions of weak public finance and sovereign crisis prevail. Our sample contains after 2008 periods that are characterized by these contingencies: the Great Recession, the 2011 double dip and the 2012-13 episode of fiscal consolidation, when both recession and weak public finance conditions were present. So, Italy seems a potential interesting laboratory for evaluating the changes of FMs between states of the economy.

However, a truly central point for this investigation is the choice of a correct empirical strategy. The contingent states of the economy are not long lasting, so a problem arises with the number of available observations and the statistical significance of their econometric relations. There are also many channels of interdependency between government expenditures and economic activity (automatic stabilizers, changes of interest rates, of expectations, and so on) and these widespread sources of endogeneity affect the estimation of FMs. The empirical literature has widely discussed these two methodological issues: how to identify truly exogenous fiscal shocks and which are the optimal strategies for the econometric estimation of FMs.

The goal of our contribution is to propose a coherent and simple method to identify fiscal shocks through a factor model. Factors are used to remove the influences that macroeconomic latent variables could have on the statistical relations between variables and to estimate FMs with a parsimonious specification. We extract factors as in standard econometric literature ('diffusion indexes', Stock and

Watson 2002a, 2002b) from a dataset that contains the main macroeconomic variables of the Italian economy, and we use a structural VAR augmented with factors (FAVAR) to compute FMs (Forni *et al.* 2000, 2005). Factors can be interpreted as structural elements that affect all the macroeconomic variables and can be used as instruments to clean distortions due to omitted variable bias and other possible endogeneity sources. The challenge is to identify a model that explains the economy reaction to public spending using latent factors as control variables. Our guess is that this approach could represent a solution to solve the issue of endogeneity between macroeconomic and fiscal variables and to correctly specify the model, common problems in the fiscal multiplier and SVAR literature. The parsimonious specification within a FAVAR model allows an investigation of the variability over time of FMs, using expanding window regression from the beginning of 2008 and looking closely at what happened in the following years, in sub-samples having a small number of observations.

After the introduction, in section 1.2 we have a synthetic review of literature. Section 1.3 is dedicated to explain the model and the shock identification strategy. In Section 1.4 we present the dataset used and describe the main evidences. In Section 1.5 we provide the estimations of Italian FMs. In section 1.6 we comment the results and suggest some conclusions.

1.2 Literature review

Empirical literature on FMs has grown at an exponential rate during the past ten years. Independent scholars and public natio-nal and international institutions have tried to evaluate them with the aim of proving or disproving the effectiveness of fiscal policies.

The problem is that FMs depend on nearly every aspect of the economic system, both in public and private sectors. So not only models rooted in different theoretical environments but also similar models can arrive at different conclusions. In a survey of US literature Ramey (2011)

shows that estimates of FMs for the US economy lie in the range of 0.5 to 2.0. The existence of sharp differences between FMs empirical estimations has been described as a morass (Leeper *et al*. 2017).

A tentative guide in the morass is roughly composed of seven factors from which the variability of estimations can originate (Gechert and Will 2012; Gechert and Rannenberg 2014; Sims and Wolff 2017): (i) model specification, in particular various type of DSGE models, with neoclassical or neo-keynesian characteristics, vs. structural macro-econometric models or VAR models; (ii) different specifications in the class of VAR models; (iii) procedure for the identification of the exogenous fiscal instrument; (iv) single country analysis vs. panel and, in the latter case, composition of panel; (v) short run vs. long run multipliers, at least in models that allow this distinction; (vi) quality of data for fiscal variables; (vii) sample periods and frequency of data.

If we do not care about punctual estimations but look at general aspects, a consensus exists that FMs are influenced by a set of characteristics of the economy, are different for different fiscal instruments and that they are state-dependent (Ilzetzki *et al*. 2013).

Trade openness influences multipliers through the elasticity of imports to domestic demand, so countries with a lower propensity to import tend to have higher FMs (Barrell *et al*. 2012). Labour market rigidity is a second structural factor affecting multipliers, because reduced wage flexibility tends to amplify the response of output to demand shocks (Woodford 2011). The exchange rate regime is a third factor, because with flexible rates the exchange movement can offset the effects of discretionary fiscal shocks (Born *et al*. 2013). FMs are reported to be highly sensitive with a positive sign to the fraction of population facing binding credit constraints and to wealth inequalities (Brinca *et al*. 2016). They are higher in developed vs developing countries (Karras 2014). Last but not the least, multipliers are linked to the efficiency of public administrations, that affects the timing of implementation of public policies; they also depend on the size of automatic stabilizers, whose dimension and efficiency reduce the

impact of discretionary fiscal shocks (Dolls *et al.* 2012).

FMs may change across different fiscal instruments. From the side of expenditure, the empirical literature suggests a hierarchy, with investment multipliers higher (and more persistent) than public consumption multipliers, low multipliers for social transfers and not significant multipliers for public wages (Wolff *et al.* 2006). This evidence however is not always clear-cut: as an example, the most recent empirical exercises on Italian FMs do not find a relevant difference in the size of public investment and public consumption impacts on Gdp (Carreras *et al.* 2016; De Nardis and Pappalardo 2018).

FMs are dependent from the state of the economy, in particular the business cycle and the directions of monetary policy. These two aspects are linked and distinct at the same time: they are linked because monetary policy reacts to business cycle but they are distinct because they involve different transmission channels. Large evidence exists that the multiplier effect on Gdp of a government expenditure shock is larger in recessions compared to expansions. This evidence emerges in studies concerning both single countries and panels, notwithstanding the various methodological frameworks adopted[2].

For what concerns monetary policy, both in US and in Eurozone the effectiveness of fiscal shocks depends on the degree of cooperation of monetary policy and FMs are higher during periods characterized by persistent low interest rates (Kilponen *et al.* 2015; Bonam *et al.* 2017). A corollary issue concerns what happens to FMs during periods of financial stress and banking crisis, when tight credit restrictions constrain the behaviour of consumers and producers. Empirical analysis report that in these circumstances the size of FMs shows tendencies toward higher values (Corsetti *et al.* 2012a; Hernandez de Cos and Moral-Benito 2016).

---

[2] For US see Auerbach and Gorodnichenko (2012), Arin *et al.* (2015), Caggiano *et al.* (2015). For Italy see Caprioli and Momigliano (2013). For OECD countries see Auerbach and Gorodnichenko (2013), Riera-Crichton *et al.* (2015), Jordà and Taylor (2016). For Eurozone countries see Batini *et al.* (2012), Boitani and Perdichizzi (2018). For G7 countries see Baum *et al.* (2012).

It is fair to say that the least consensual and most controversial topic is related to the potential influence on FMs of what are generally defined as weak public finance conditions, namely situations of high public debt ratio to Gdp and risk of sovereign crisis. When sovereign risk takes an important weight in expectations, a successful fiscal consolidation can determine a reduction of interest rates and an improvement of the state of confidence. These potential gains can counterbalance the output costs of fiscal consolidation, so reducing the size of multipliers and the contractionary impacts of tightening fiscal policy (Kirchner *et al.* 2010). A theoretical extreme situation could occur if the gains from fiscal consolidation prevail and outweigh the costs: if this happens FMs should change their algebraic sign. This is the so-called hypothesis of expansionary fiscal consolidation based on non-Keynesian effects of fiscal policy (Alesina and Perotti 1995; Forni *et al.* 2010).

After the Great Recession a different stream of literature has grown arguing the possibility that fiscal consolidations during a recession could be self-defeating, i.e. that they can end up with higher debt-to-Gdp ratios because multipliers in 'bad times' are higher, produce negative impacts underestimated by policy makers and moreover through hysteresis mechanisms the contraction can be transmitted to potential Gdp and become permanent (Fatás and Summers 2016).

A small numbers of empirical studies find a reversal in the algebraic sign of FMs when weak public finance conditions prevail (Corsetti *et al.* 2012a). The predominant empirical evidence is however that weak public finance conditions reduce the size of FMs, also during recessions, but without a sign reversal. Estimated values of FMs remain positive, though smaller, implying that a successful fiscal consolidation in situation of high public debt and sovereign crisis can obtain some gains through interest rates and confidence, so reducing the output costs of fiscal contraction, but it is far from

operating without pain[3].

FMs may change over time (Perotti 2002). A decline in the size of government spending multipliers and a reduction of their persistence are reported in the long run both in the US and in the Eurozone. In the US the turning point seems around 1980 (Bilbiie *et al.* 2008; Leeper *et al.* 2017), in the Eurozone around 1990 (Kirchner *et al.* 2010). In the Italian case Cimadomo and D'Agostino (2016) suggest a U-shaped time evolution of FMs, with a decline from the beginning of the 90s and an increase after the 2008 recession.

The empirical estimation of FMs raises two main challenges. The first is about the identification of fiscal shocks. The two main methods utilised in the empirical literature are the narrative approach, choosing the episodes of exogenous fiscal shocks through a qualitative and historical analysis, and the SVAR methodology. The problem of this second stream is that estimations could be overspecified: overspecification could arise from the high number of variables and lags used in a small sample environment. It follows the need to add different kind of control variables, avoiding that their nature and number could determine a severe loss in degrees of freedom.

The second challenge is the frequency of data and the problem of evaluating the stability of FMs through time. What seems important from this point of view is a research strategy that allows the evaluation of different sub-samples in the investigated period. In this perspective an alternative to VAR is to calculate multipliers with the local projection technique (Jordà 2005).

Our contribution tries to deal with these challenges with an innovation of SVAR methodologies. To our knowledge the use of a FAVAR model for the estimation of FMs is an innovation in the literature. Diffusion indexes have been sometimes used as instruments for robustness analysis, but not to identify the exogenous component of government expenditure.

---

[3] See Auerbach and Gorodnichenko (2013), Ilzetzki *et al.* (2013), Guajardo *et al.* (2014), Hernandez de Cos and Moral-Benito (2016), Boitani and Perdichizzi (2018).

1.3 Estimation strategy

FMs literature uses different SVAR model, with distinctions along two main issues: the SVAR specification and the shock identification strategy. We propose to use a factor model to specify the VAR model, and a Cholesky decomposition after the estimation of exogenous fiscal shock.

1.3.1 Factors model

A factor model allows a proper consideration of the unobserved heterogeneity that can determine coefficients estimation bias in a SVAR framework. When we extract factors we are considering latent and common characteristics of the economy; the use of these variables enable to look at pervasive influences of the economic system instead to choose arbitrarily the regressors of the VAR model. From an econometric point of view we can consider factors as control variables, in order to remove the omitted variable bias and to estimate the variable reaction to different type of shocks. Through factors it is possible to make a decomposition of every time series in two parts, the common and idiosyncratic one: the first is affected by macroeconomic shocks, while the second represents the specific dynamics of the variable.

A factor model could be described with the following equation

$$X_t = \chi_t + \xi_t \qquad (1.1)$$

or, if we look to the i-th variable

$$x_{i,t} = \chi_{i,t} + \xi_{i,t} \qquad (1.2)$$

In which $X_{i,t}$ is a large dataset that contains all the possible aspects of the economy, $\chi_{i,t}$ is the common component while $\xi_{i,t}$ is the idiosyncratic component. The common component is associated with a set of latent factors, $f_{i,t}$ from which we can capture the common comovements of all the variables. We can

write equation (1.2) as follows:

$$x_{i,t} = \lambda_{i,1}f_{1,t} + \cdots + \lambda_{i,r}f_{r,t} + \xi_{i,t} \tag{1.3}$$

where $f_{1,t}$, ..., $f_{r,t}$ are the common factors and $\lambda_{i,1}$, ..., $\lambda_{i,r}$ are the 'loading factors' that relate common factors with every variable of X_t.

To correctly identify the model we make the usual key assumptions: a high sample number N of time series ($N \gg r$), to represent all the characteristics of the economy and to estimate the space spanned by the factors (Stock and Watson 2002b); the idiosyncratic component $\xi_{i,t}$ must be uncorrelated to the common component $\chi_{i,t}$, and $\xi_{i,t}$ could be serially correlated and weakly cross-correlated. We estimate factors as 'diffusion indexes' (Stock and Watson 2002a), usually with the support of different criteria (Bai and Ng 2002 criteria, scree plot, etc.) to represent as well as possible the space spanned by pervasive component. Consequently we are able to identify the idiosyncratic part of each series[4].

After the factors estimation, we can write the VAR system equation as follows:

$$\begin{bmatrix} f_{1,t} \\ \dots \\ f_{r,t} \\ \Delta G_t \\ \Delta GDP_t \end{bmatrix} = \Phi(L) \begin{bmatrix} f_{1,t-1} \\ \dots \\ f_{r,t-1} \\ \Delta G_{t-1} \\ \Delta GDP_{t-1} \end{bmatrix} + \begin{bmatrix} v_{f_{1,t}} \\ \dots \\ v_{f_{r,t}} \\ v_{\Delta G_t} \\ v_{\Delta GDP_t} \end{bmatrix} \tag{1.4}$$

in which $\Phi(L)$ is a $q{\times}q$ matrix of VAR lag polynomials with $q = r + 2$. The model is composed by our variables of interest, the public consumption expenditure growth rate, ΔG, and the response variable ΔGdp

---

[4] See Baldini and Causi (2020) for a wide discussion on this point and for the procedures of estimation of common factors.

(Gdp growth rate)[5] and is 'augmented' with factors (FAVAR). Imposing $Y_t = [f_{1,t}, \dots, f_{r,t}, \Delta G_t, \Delta Gdp_t]$, we can write the previous equation as follows:

$$Y_t = \Phi(L)Y_{t-1} + v_t$$
$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_k Y_{t-k} + v_t \tag{1.5}$$

in which the term v_t represents unidentified shocks, and we can write it in the following way:

$$v_t = Ku_t \tag{1.6}$$

As usual in this econometric framework factors are the first element of the vector Y_t because they represent the pervasive comovements of all the sample variables; after them, there is our relation of interest that consist in the estimation of the impact of public expenditure on Gdp.

1.3.2 Identification strategy

To compute FMs we use the standard definition of multiplier at horizon i:

$$Multiplier_{h=i} = \frac{\Delta Gdp_{t+i}}{\Delta G_t} \tag{1.7}$$

This means that, starting from growth rate, we estimate the IRFs, we cumulate them and we consider the horizon i[6].

In a SVAR framework, the innovation of G in model (1.4) could be seen as a linear combination of different type of shocks: the automatic response of government spending to output, factors and

---

[5] They are expressed in logarithmic difference. For the manipulation of all the variables comprised in our dataset see Baldini and Causi (2020).

[6] As standard in this literature, the use of logarithms determines the need of transforming the estimated coefficients from elasticities to multipliers. We use the Auerbach and Gorodnichenko (2012) method and multiply elasticities by the conversion factor Gdp/G.

Gdp variation, the discretionary response of policy maker to some innovation and a random and discretionary shock to fiscal policy. The last one is what we properly call as 'structural shock', that is uncorrelated with other structural shocks of the economy. In synthesis we can write the equation of G innovation as follows:

$$v_{\Delta G_t} = \alpha_1 v_{f_{1,t}} + \alpha_2 v_{f_{2,t}} + \cdots + \alpha_k v_{f_{k,t}} + e_{\Delta G_t} \qquad (1.8)$$

in which $e_{\Delta G_t}$ is the structural innovation.

In the SVAR framework, researchers try to calibrate α_j parameters to estimate the right hand side of the equation (1.8) and then use them to compute the G 'cyclical adjusted' fiscal variable, that is the part of the fiscal shock not affected by cyclical component. We try to give a solution coherent with our model: using factors model we estimate proxies of spending variable not affected by cyclical component, and then we use these variables inside the FAVAR equation to estimate the effect of a fiscal shock on Gdp. So we are able to compute a cyclical adjusted spending variable that is independent from the macroeconomic influences, and to recover the part of G related to discretionary response of policy maker or to random shock. With the aim of obtaining an exogenous measure for G our strategy is to use different public consumption expenditure measures, after removing cyclical influence. We use different specification of the following equation:

$$\Delta G_t = \sum_{i=1}^{k} \lambda_{i,t}\,(L)f_{i,t} + \gamma_i(L)\Delta Gdp_t + \gamma(L)\Delta G_{j,t-1} + v_{G_t} \quad (1.9)$$

This is an ARDL model in which we consider different lag specification and different number of factor numbers to obtain different proxies of cyclical adjusted G. The estimated residuals of these models represent the G that is not influenced by the economic cycle (ΔG_t^{ca}), that can be considered as proxy of exogenous public expenditure.

For what concerns the specification problem, we order the variables according to the model (1.4). Factors come first, and they are extracted in sequence according to the variance explained; then we impose our fiscal variable, the G growth rate for Italy and, after this, the private Gdp growth rate, the variable on which we want to observe the impact. If we use a Cholesky decomposition with this variables sorting, we are imposing that fiscal spending could be seen as exogenous respect to Gdp, and that Gdp reacts to an increase of public spending within the same quarter.

1.4 Data and descriptive evidence

In the estimations we consider seasonally adjusted Gdp[7] and G, evaluating them at constant prices, so their growth rates are in real term. We follow the standard literature considering both G and Gdp per-capita growth rates.

A factors model requires a large panel of stationary time series that represents all the possible aspects of the economy at quarterly frequency. We use 100 time series of the Italian economy. All the series are taken from FRED database, except for the credit market series, for which the source is Bank of Italy. For the complete list and sources see Baldini and Causi (2020).

To extract factors we need to transform the dataset to obtain stationary variables. We follow the methodology based on the break in the mean literature. Dropping out the break in the mean, in particular those induced by the Great Recession, is useful to obtain stationary time series without over-differencing the data[8]. After this manipulation we can extract factors that will 'augment' our VAR model (FAVAR).

---

[7] We consider the private part of Real Gdp, obtained subtracting government final consumption expenditure from the overall Gdp.
[8] This point is discussed in Baldini and Causi (2020) and we adopt here the same methodology.

The Great Recession had a severe impact on the Italian economy (Fig. 1.1). From the first quarter of 2008 to the second quarter of 2009 real Gdp has fallen by 8,26%. After a weak recovery there was another fall in the first quarter of 2011, the so-called 'double dip'. At the end of 2014 real Gdp come back to the 2000 level, leaving to Italy a weak state of the economy.

Figure 1.1
Private Real Gdp vs government final consumption expenditure.
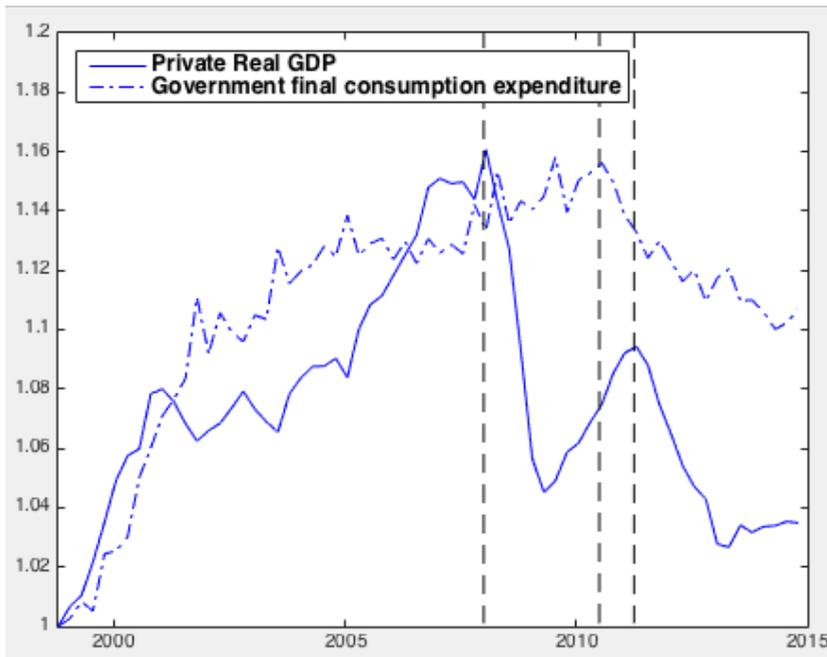
Normalized Gdp and G at constant price and quarterly adjusted.

Looking at public consumption, our policy variable of interest, we observe a level contraction after the mid of 2010 and more strongly with fiscal consolidation from 2012. From the mid of 2010 we observe a strong increase in government bond interest rate with a peak during 2011 (Fig. 1.2). The incumbent sovereign crisis determined from 2012 the starting of a phase of fiscal consolidation,

and in correspondence of this phase we see a gradual decrease of interest rates.

Figure 1.2
*Government final consumption expenditure
and 10 years government bonds rate of interest.*
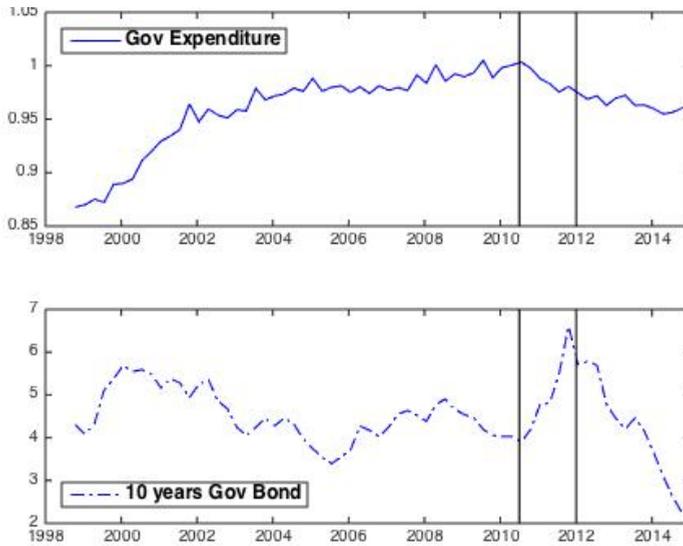
Government final consumption expenditure at constant price and quarterly adjusted.

Fig. 1.3 contains the variance explained by the factors extracted. The first factor is the more relevant and explains 21% of the total variance; if we take two factors we explain 38% of the total variance and with 3 factors 44%.

Figure 1.3
Factor variance, scree plot
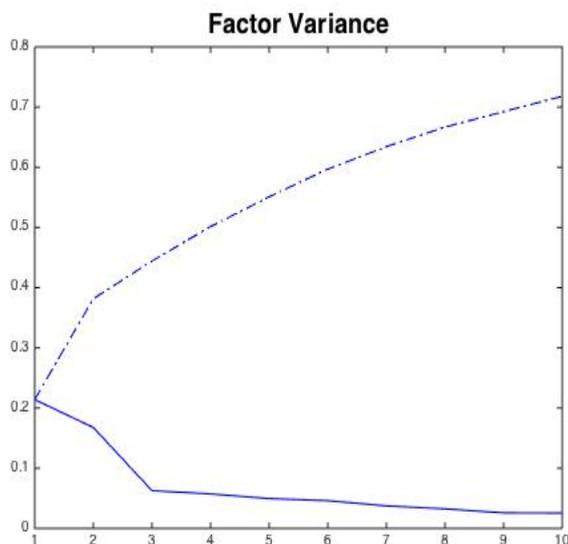
Which is the factors interpretation? In Fig. 1.4 we plot the first three factors versus the most correlated variables. We could say that the first factor dynamics is extremely related to real Gdp growth rate and the second factor is related to the HCPI growth rate. The third factor is correlated, though with lower intensity, with the Economic Policy Uncertainty index, that is an index that tries to capture the uncertainty about the economic policy action that will be undertaken. Also the Business Tendency Surveys seem to be adequately correlated. The third factor, in some sense, seems to explain the expectations on the state of the economy.

Figure 1.4

Factors description: factors vs the variables with higher correlation.
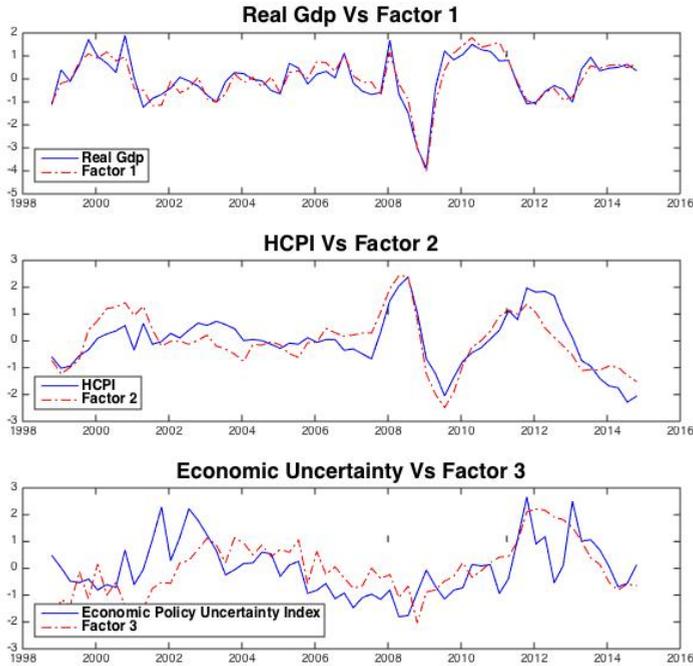
Factors and variables have unitary variance to be compared.

Notwithstanding this, the trial of factors identification could be misleading. Factors are used as interpreters of latent influences that can't be attributed to single variables. They are instruments for controlling the impact of common components.

1.5   Results

We can now estimate a FAVAR model in order to analyze IRF in a VAR context. Given the methodology in section 1.3, we estimate different kinds of cyclical adjusted G, specifying models with different lags. Basically we use equation (1.9) to estimate the ΔG_t^{ca}, changing the number of factors used and the number of lags. There

are different ways in which macroeconomic tendencies could have a structural effect on government spending, and different specifications help to obtain robust estimations. We estimate 9 models using 1, 2 and 3 factors, and changing the number of lags, from 2 to 4. So, we obtain 9 proxies of ΔG_t^{ca}, and we put them into the FAVAR model to estimate the Gdp response. In Fig. 1.5 we report three examples of ΔG_t^{ca}, represented by the residuals of equation (1.9).

Figure 1.5
Government final consumption expenditure: estimations and residuals
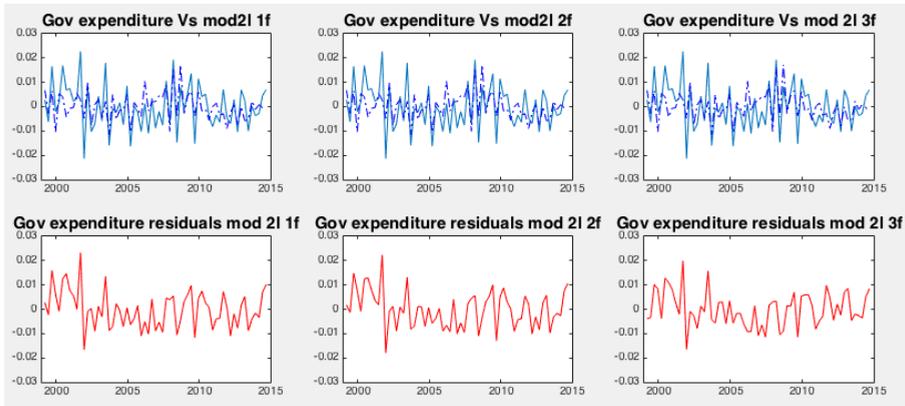
ΔG_t^{ca}(red lines) are the residuals of three models: the equation (1.9) estimated with 1 factor and with 2 lags, the equation (1.9) estimated with 2 factors and 2 lags and the equation (1.9) estimated 3 factors and 2 lags. Blu lines are the G original series and blu dotted lines are the G estimated.

After this step we proceed to the FMs estimation. Following the model (1.4), the equation can be specified as follows:

$$\Delta Gdp_t = \sum_{i=1}^{r} \lambda_k^* f_{k,t-1} + \gamma_j \Delta \hat{G}_{t-1}^i + \beta \Delta Gdp_{t-1} + v_t \qquad (1.10)$$

In which we can see that ΔGdp is function of the common factors, of its lag and of the lagged value of $\Delta \hat{G}_{t-1}^i$, where i is equal to 1, 2 or 3 depending on the different kind of ΔG that we have

estimated previously. This is exactly the last equation of the FAVAR model (1.4). Our variables are, essentially, growth rate, and for this reason we do not have problems of spurious correlation.

We specify the model imposing the parsimonious specification of 1 lag. In literature there are cases of specification procedures using 4 lags to capture all the year relation between variables. The risk, in a small sample framework, is to over-specify the model, and for this reason we preferred the most penalizing criteria in the lag selection. In our case BIC criteria suggest clearly to use only 1 lag, and ACF-PACF function suggests the same.

Results are shown in Tab. 1.1. We test models with different lag specifications, but the results confirm that the main relevant regressors are always the first lags of factors and G. Factors seem to be very influent in explaining the Gdp dynamics.

Table 1.1 – *FAVAR Gdp estimation*

VARIABLES	(1)	(2)	(3)
$f_{1,t-1}$	0.0193***	0.0191***	0.0187***
	(0.0034)	(0.0035)	(0.0035)
$f_{2,t-1}$	-0.0127***	-0.0125***	-0.0126***
	(0.0018)	(0.0019)	(0.0020)
$f_{3,t-1}$	0.0042	0.0037	-0.0003
	(0.0026)	(0.0027)	(0.0025)
$\Delta \hat{G}_{t-1}$	0.4756***	0.4576***	0.4743***
	(0.1177)	(0.1121)	(0.1114)
ΔGdp_{t-1}	-0.4135**	-0.4060**	-0.3861**
	(0.1723)	(0.1735)	(0.1761)
R-squared	0.6224	0.6126	0.6150
F(5,57)	18.7910	18.0278	18.2124

Sample goes from 1992:2 to 2014:4, and we estimate 3 models: in model (1) $\hat{G}$ is the residual of equation (1.10) estimated with 1 factor and 2 lags, in model (2) $\hat{G}$ is the residual of equation (1.10) estimated with 2 factors and 2 lags, and in model (3) $\hat{G}$ is the residual of equation (1.10) estimated with 3 factors and 2 lags. Errors are robust respect to autocorrelation (HAC errors).

G is significant and positive: a 1% increase of Government final consumption expenditure growth rate seem to cause a 0.47% increase of private real Gdp growth rate after 1 quarter; this coefficient is crucial to estimate fiscal multiplier. The significance of coefficients is very similar along the different model used, and the evidences are very similar for the other $\Delta \widehat{G}_{t-1}^{i}$ proxy. From these results we expect an increase in Gdp after 1 quarter and then a path that goes back to the steady state.

With these estimations we can compute IRFs for all the model considered, and then manipulate them (as explained in footnote 6) to obtain FMs: we want to observe the reaction of Gdp after a G shock specified in different ways. So, we have 9 G cyclically adjusted variables estimated by equation (1.9), and fiscal multipliers estimated with equation (1.10) using 2 and 3 factors. To sum up, we estimate 18 models IRFs along 20-horizon: in Fig. 1.6 we report the median, the 6-th and the 94-th percentile for quarterly and overall elasticities; in Fig. 1.7 the conversion of elasticities in multipliers.

Figure 1.6
Impulse response functions on Private Real Gdp.
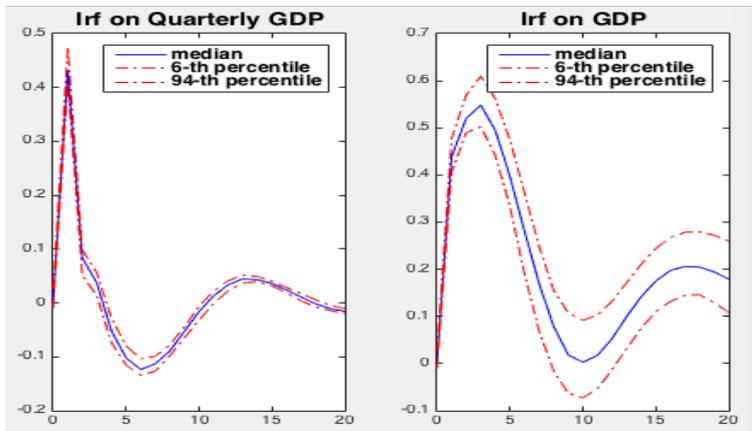
Impulse response functions: 18 models. We report the median, the 6-th and the 94-th quartile of the response of Private Real Gdp to a government final consumption expenditure shock.

Figure 1.7
Multipliers on Private Real Gdp.
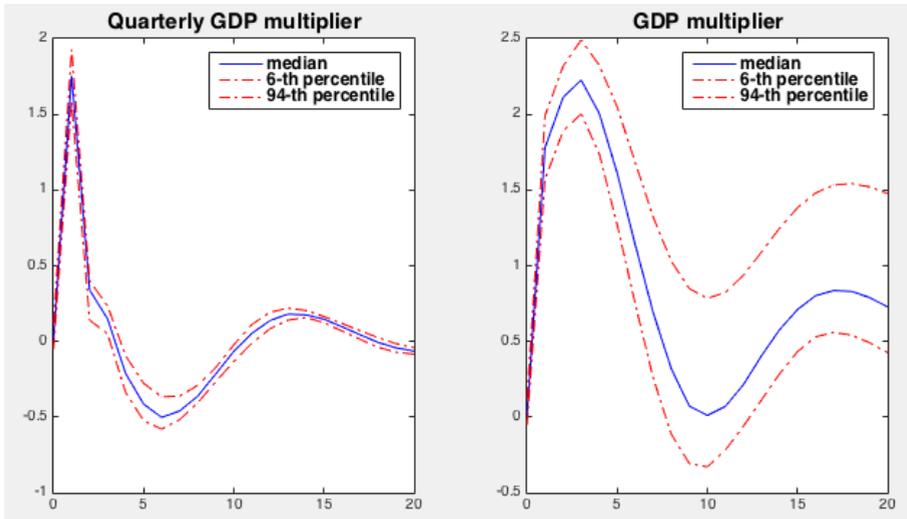
Multipliers: 18 models. We report the median, the 6-th and the 94-th quartile of the multiplier on Private Real Gdp from a government final consumption expenditure shock, obtained by the conversion of elasticities of Fig. 1.7.

A comparison of these results with Italian and international estimations bring at the conclusion that they are coherent with the existing empirical literature, though with some specificity. With respect to Italy, studies on the effects of fiscal shocks in a VAR environment are very few. The impact multipliers derived from our FAVAR model are higher than those suggested by Giordano *et al.* (2007) and Caprioli and Momigliano (2011, 2013). On the contrary, our long run multipliers are lower and FMs calculated at one year of distance are quite similar. Higher Italian spending multipliers in the short run than in the long run have been found also by Cimadomo and D'Agostino (2016), while Batini *et al.* (2012) found lower values of Italian FMs[9]. In a different methodological environment (a structural macro-econometric model)

---

[9] They consider however total government spending and not only government consumption.

the year multiplier of intermediate consum-ption is estimated at 1.8 by De Nardis and Pappalardo (2018).

To make the comparative exercise easier it is useful to calculate the cumulative multiplier, defined (following Ramey and Zoubairy 2018) as:

$$Cumulative\ Multiplier_{h=i} = \frac{\sum_{i=1}^{h} \Delta Gdp_{t+i}}{\sum_{i=1}^{h} \Delta G_{t+i}} \qquad (1.11)$$

The cumulative multiplier (CM) takes into account the persistence of fiscal shocks and is obtained dividing the cumulated Gdp response by the cumulated spending response.

In our model 1 year CMs range is 1.4-1.7 and the long run CM converges toward 0.8 (Fig. 1.8 and 1.9). In Giordano *et al.* (2007) the year cumulative multiplier has a range 1.4-2.2, while in Caprioli and Momigliano (2013) the range is 1.2-1.9. Long run CMs are instead higher than ours in all these models (between 1.8 and 2.5)[10]. These differences may be a consequence of our specification choice to use only one lag in order not to lose degrees of freedom in estimation, but it is worth remembering that studies that use four lags show increasing ranges of confidence and decreasing statistical significance when the responses are projected for long time periods.

---

[10] The same comparative evaluation could be suggested with the Italian estimations of Afonso *et al.* (2011) if their results are converted from elasticities to spending multipliers.

Figure 1.8
Cumulative multipliers on Private Real Gdp: 2 factors.
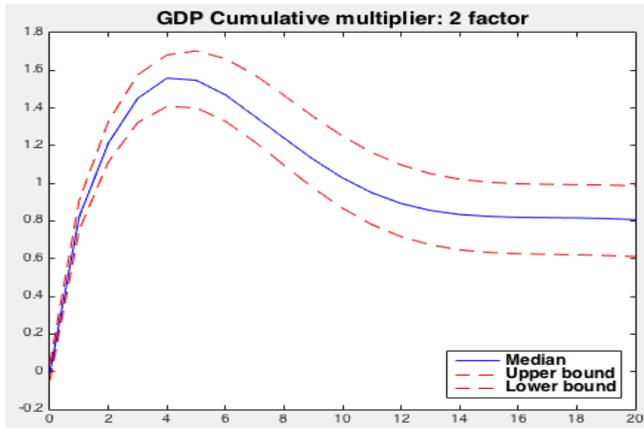
Cumulative multipliers: 9 models. We report the median, the upper bound and the lower bound of the cumulative multipliers as defined by (1.11), estimated through Impulse response functions with 2 factors.

Figure 1.9
Cumulative multiplier on Private Real Gdp: 3 factors.
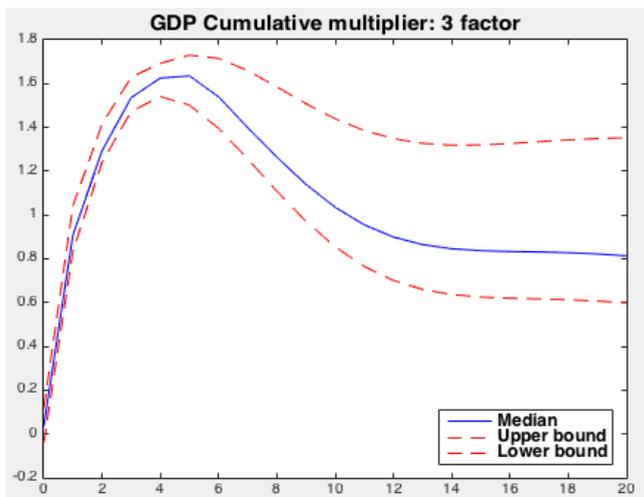
Cumulative multipliers: 9 models. We report the median, the upper bound and the lower bound of the cumulative multipliers as defined by (1.11), estimated through Impulse response functions with 3 factors.

If we look at the ranges of values proposed by the IMF for the assessment of the strength of fiscal policy effects in different countries (Batini *et al.* 2014), our results should suggest that Italian FMs lie in the upper category of 'high multipliers' in the short run and in the median category of 'medium multipliers' in the long run. Indeed, when FMs are estimated in research frameworks that allow comparisons between different countries, in particular in the Euro Area, Italian FMs are reported in a medium-high position: at a glance, above Germany but below France (Kilponen *et al.* 2015; Carreras *et al.* 2016); in any case below United States when public consumption and not the overall government expenditure is considered (Auerbach and Gorodnichenko 2012; Brinca *et al.* 2016).

In our sample there is a sub-period of deep recession and another sub-period where a second deep recession was combined with a sovereign crisis and was followed by a third period beginning with a fiscal consolidation. How the Italian FMs performed during these different states of the economy and does this behaviour fit with the suggestion of the international literature? Answering to this question means to estimate fiscal multipliers in different subsamples.

We select a starting window that goes from the last quarter of 1998 to the first of 2008, and then we compute the long run FMs with expanding windows from 2008 to the end of our sample[11]. We want to observe the effect of each quarterly observation added after the first quarter of 2008, and try to describe the path of long run FMs during the years 2008-2014. Even if we could incur in a loss of degrees of freedom, this is an interesting exercise to evaluate the evolution of FMs over time and over different contingent states of the economy; the parsimonious parametrization used helps in obtaining for the estimations a fairly good stability, the maximum possible in this kind of exercise.

Results are shown in Fig. 1.10 and 1.11. The sample is divided in four parts, looking at recession and expansion periods. We report the mean of long run FMs (the black line) estimated with 2 and 3 factors through 9 models, compared to the Gdp dynamics.

---

[11] With expanding windows we select a starting and an ending date for our initial sample, than we change the ending date to obtain the other samples.

Figure 1.10
Long run multipliers on Private Real Gdp: expanding window with 2 factors.
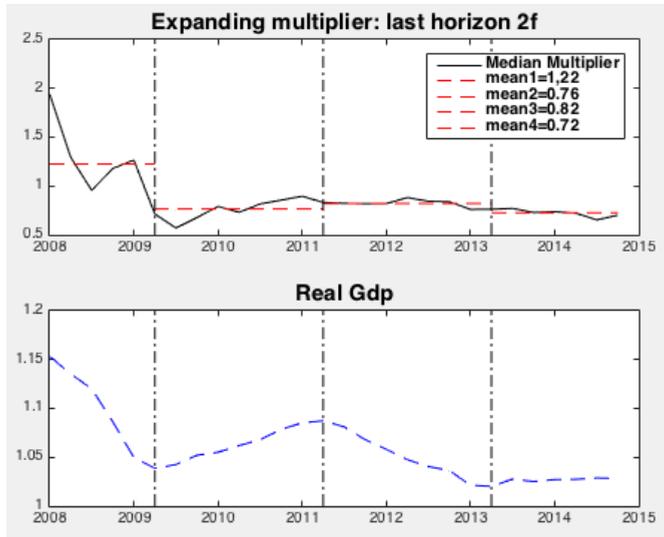
Long run multipliers on Private Real Gdp: 9 models mean.

Figure 1.11
Long run multipliers on Private Real Gdp: expanding window with 3 factors.
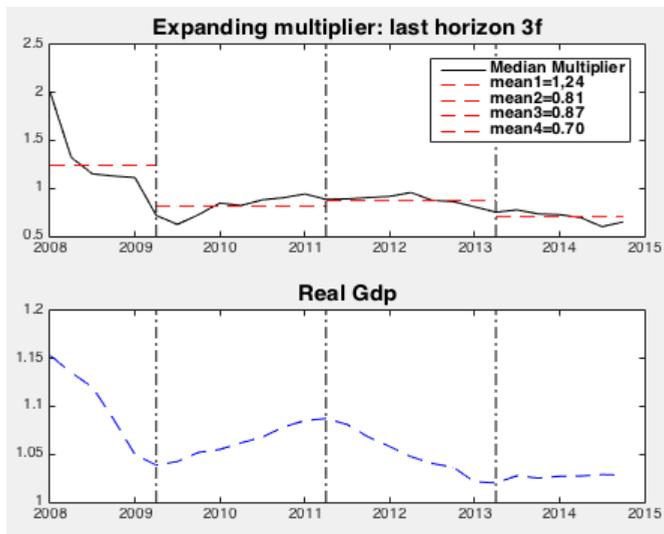
Long run multipliers on Private Real Gdp: 9 models mean.

At the beginning of 2008 long run spending multiplier was close to 2, and long run FMs in the first part of the graphs are, in mean, equal to 1.22 if computed with 2 factors, and 1.24 if computed with 3 factors. These values are substantially higher than the median obtained in the entire sample. This seems a confirmation in the Italian case of 2008-2009 that during a crisis we observe higher multipliers compared to those observed in phases of expansion. In the following period (2009-2011) multipliers fall in a range of 0.75-0.94, while in the third period, the double dip 2011-2012, multipliers become again higher (0.8-1). In the last period, after 2013, multipliers become stable and lower (0.73-0.84).

Last, we want to look at another kind of periodization, leaded by policy considerations. From the mid of 2010 the government final consumption expenditure started to fall, and beginning from 2012 the Italian economy entered in an episode of fiscal consolidation with the aim of stopping an impending sovereign crisis. We divide the sample in three parts: 2008-Q1 to 2010-Q2, 2010-Q2 to 2012-Q1 and 2012-Q1 to 2014-Q4 (Fig. 1.12 and 1.13). In the first period FMs are higher in mean and decreasing: going from a phase of crisis to a phase of short and moderate recovery FMs pass from 2.1 to a level well below 1. In the second phase (the 'double dip') FMs are stable or slightly increasing, while in the last phase the path becomes stable or decreasing. Therefore, during and after fiscal consolidation the recessive influence of spending multipliers lowered, becoming close to 0.7-0.8.

Figure 1.12
Long run multipliers on Private Real Gdp: expanding window with 2 factors.
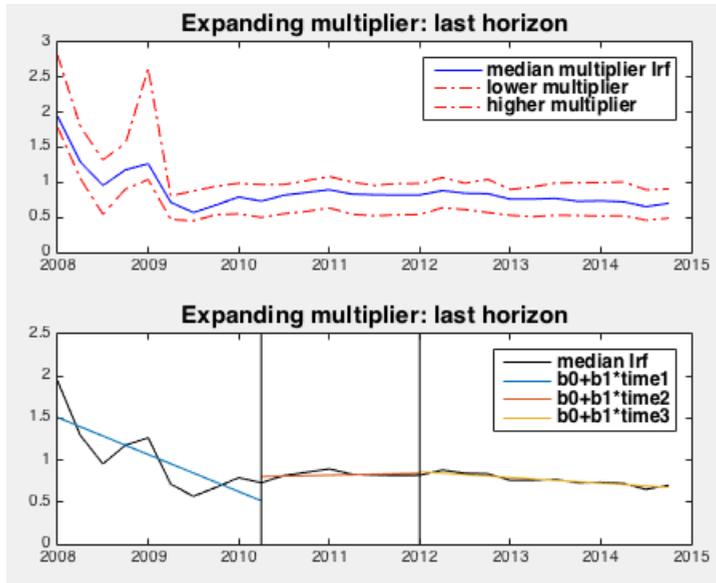
Long run multipliers on Private Real Gdp: 9 models mean.

When public spending grows a positive multiplier explains the size of the contribution of public spending to Gdp growth. When public spending decreases, a positive multiplier describes the contribution of public spending to Gdp contraction. This is what we can observe before and after the mid of 2010. In the 2009-2010 case it is possible to argue that the increase in direct public spending (not only the part deriving from automatic stabilizers, because our G variable is cleaned from cyclical components) played probably a role in mitigating the effects of the first peak of the Great Recession on Gdp and in sustaining the following period of moderate expansion.

Figure 1.13
Long run multipliers on Private Real Gdp: expanding window with 3 factors.
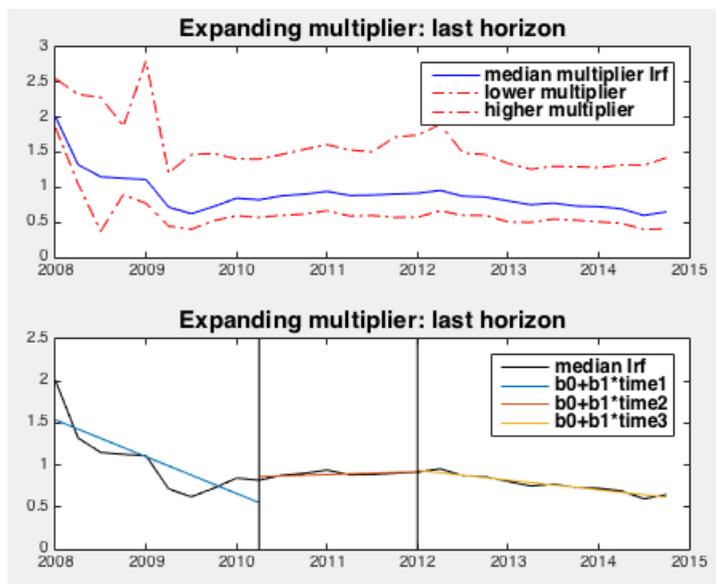
Long run multipliers on Private Real Gdp: 9 models mean.

The contraction of public consumption began after the mid of 2010 and our evidence allows to argue that the effects on Gdp were important, because amplified by an increase of FMs. Giving a look to Fig. 1.1 and 1.2 and reminding the state-contingencies of the economy that could affect FMs a possible interpretation could consider two elements. The first can be related to the arrival after two years of recession of difficulties in the banking system caused by the increase of non-performing loans, with the consequence of a credit crunch and a widening of the fraction of population and enterprises facing binding credit constraints. The second can originate from a non-cooperative behaviour by monetary policy: ECB left interest rates free to increase together with risk premium charged on Italy as well as on others Southern Eurozone countries during the turmoil caused by the France-Germany meeting in Deauville, held in October 2010.

The reduction of public consumption continued in the following years, with a peak linked to the 2012 fiscal consolidation. Just from this year, and with more clear evidence from 2013, our estimated FMs seem stabilizing at a lower level. They maintain a positive sign, so also after 2012 the reduction of public consumption contributed to the contraction of real Gdp, but with a lower multiplier effect.

Change in the size (not in the sign) of FMs can be explained by the nature of the fiscal consolidation, motivated by the objective of contrasting an incumbent financial sovereign crisis. Our results show that the 2012 policy has been successful in reducing the negative impact on Gdp of the fiscal contraction, thanks to gains obtained through the reduction of interest rates, the stabilization of the overall national climate after a turbulent period of uncertainty and instability and a positive change of confidence towards a future season of recovery and structural reforms.

1.6 Conclusions

We study Italian public spending multipliers for the period 1998-2014 and we estimate the effects of final government consumption shocks on Private Real Gdp in a quarterly framework. We are interested not only in the evaluation of the FMs size, but also in the investigation of the variability of FMs over time and across different states of the economy.

We propose a factor augmented VAR (FAVAR) model as a method that can help to solve common problems existing in the empirical estimation of the fiscal policy impact on the economy. With the FAVAR technique we are able to implement a parsimonious specification of the model taking into account the latent influence of all the macroeconomic variables of the economy. As a consequence it is possible, using expanding window regression, an analysis of the variation of FMs over small time intervals, and in particular over the different

macroeconomic contingencies that are present in the sample after 2008, like the Great Recession, the double dip and the Italian fiscal consolidation starting from 2012.

We find three main results. First of all, the short run (impact) FM computed for the Italian economy is equal to 1.7 in median and long run multiplier to 0.7, while our estimated range for the one-year cumulative multiplier is between 1.4 and 1.7 (0.8 in the long run). These results seem coherent with those existing in literature, though with some qualification, and suggest that the Italian economy is characterized by FMs that can be considered, as confronted with the international standards, of high size in the short run and of medium size in the long run.

The second result is that Italian long run FMs are not invariant through time and cycles. They show significant changes from the beginning of the Great Recession and through the following phases, with long run FMs going up and down from 2 to 0.5. This means that there is a big time inconstancy for the reactions to the fiscal stimulus, linked above all to the contingent state of the economy. We verify in both the Italian recessions following 2008 that multipliers seem showing higher values during 'bad times'.

Finally, we have evidence of a reduction of FMs during and after the 2012 Italian episode of sovereign crisis and fiscal consolidation. So, we confirm existing results that do not corroborate the hypothesis of expansionary fiscal consolidation. But the reduction and stabilization of multipliers during and after 2012 can be interpreted as evidence that the transmission of fiscal policy performs in slightly different ways when conditions of weak public finance and sovereign financial crisis prevail. The negative impact inflicted by the contraction of public expenditure is partially offset by the gains of financial stability: fiscal multipliers remain positive, but smaller.

The mix of 2012 Italian fiscal consolidation was 30% cuts of expenditures and 70% increases of taxes. Being our analysis limited to public consumption, we can't derive an overall precise evaluation. However, the evidence that the output effects of expenditure contractions were reduced through the positive effects of financial stabilization and improvement of confidence is in our opinion a signal of

effectiveness and success of the 2012 Italian fiscal consolidation.

The overall tendency in the analyzed sample, ignoring cycles, seems a reduction of the size of FMs. A possible structural explanatory factor could originate from the composition of Italian government consumption expenditure, mainly for two aspects. First, smaller shares have been allocated through time in activities bringing potentially the highest impact on output, as investments (declined during our sample period from 12.8% to 11.7% of total direct public spending). Second, a growing share of government consumption is allocated in the purchase of drugs and technologies destined to the national health system (36% of total public purchases in 2015 vs. 22% in 1995; Ufficio Parlamentare di Bilancio 2017), a category of goods with high import component and less impact on domestic Gdp.

References

Afonso, A., Baxa, J. and Slavik, M. (2011). Fiscal developments and financial stress. A threshold VAR analysis. *European Central Bank Working Paper,* 1319.

Alesina, A. and Perotti, R. (1995). Fiscal Expansions and Adjustments in OECD Economies. *Economic Policy*, 10(21), 207-248.

Arin, K.P., Koray, F. and Spagnolo, N. (2015). Fiscal multipliers in good times and bad times. *Journal of Macroeconomics*, 44, 303-311.

Auerbach, A.J., and Gorodnichenko, Y. (2012). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy*, 4(2),1-27.

Auerbach, A.J., and Gorodnichenko, Y. (2013). Fiscal Multipliers in Recession and Expansion. In Alesina, A. and Giavazzi, F. (eds.), *Fiscal Policy after the Financial Crisis*. Chicago: University of Chicago Press, 63-98.

Bai, J., and Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, 70(1), 191-221.

Baldini, A. and Causi, M. (2020). Restoring credit market stability conditions in Italy: evidences on Loan and Bad Loan dynamics. *The European Journal of Finance*, 26(7-8), 746-773.

Barrell, R., Holland, D. and Hurst, I. (2012). Fiscal consolidation: Part 2. Fiscal Multipliers and Fiscal Consolidations. *OECD Economics Department Working Paper*, 933.

Batini, N, Callegari, G. and Melina, G. (2012). Successful Austerity in the United States, Europe and Japan. *International Monetary Fund Working Paper*, 12/190.

Batini, N., Eyraud, L., Forni, L. and Weber, A. (2014). Fiscal Multipliers: Size, Determinants, and Use in Macroeconomic Projections. *International Monetary Fund Technical Notes and Manuals*, 14/04.

Baum, A., Poplawski-Ribeiro, M. and Weber, A. (2012). Fiscal Multipliers and the State of the Economy. *International Monetary Fund Working Paper*, 12/286.

Bilbiie, F.O., Meier, A. and Muller, G.J. (2008). What accounts for the changes in U.S. fiscal policy transmission? *Journal of Money, Credit and Banking*, 40(7), 1439-1469.

Blanchard, O.J., and Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics*, 117(4), 1329-1368.

Blanchard, O.J., and Leigh, D. (2013). Growth Forecast Errors and Fiscal Multipliers. *American Economic Review*, 103(3), 117-120.

Boitani, A. and Perdichizzi, S. (2018). Public Expenditure Multipliers in Recessions. Evidence from the Eurozone. *Università Cattolica del Sacro Cuore, Dipartimento di Economia e Finanza Working Paper*, 68.

Bonam, D., de Haan, J. and Soederhuizen, B. (2017). The effects of fiscal policy at the effective lower bound. *De Nederlandsche Bank Working Paper*, 565.

Born, B., Juessen, F. and Mueller, G.J. (2013). Exchange Rate Regimes and Fiscal Multipliers. *Journal of Economic Dynamics and Control*, 37(2), 446-65.

Brinca, P., Holter, H.A., Krusell, P. and Malafry, L. (2016). Fiscal multipliers in the 21st century. *Journal of Monetary Economics*, 77, 53-69.

Caggiano, G., Castelnuovo, E., Colombo, V. and Nodari, G. (2015).

Estimating fiscal multipliers: news from a non-linear world. *Economic Journal*, 125(584), 746-776.

Caprioli, F. and Momigliano, S. (2013). The Macroeconomic Effects of Expenditure Shocks During Good and Bad Times. *Banca d'Italia Workshops and Conferences, Fiscal Policy and Growth*, 79-110.

Carreras, O., Kirby, S., Liadze, I. and Piggott, R. (2016). Quantifying fiscal multipliers. *National Institute of Economic and Social Research Discussion Paper*, 469.

Cimadomo, J. and D'Agostino, A. (2016). Combining Time Variation and Mixed Frequencies: An Analysis of Government Spending Multipliers in Italy. *Journal of Applied Econometrics*, 31, 1276-1290.

Corsetti, G., Meier, A. and Müller, G.J. (2012a). What determines government spending multipliers? *Economic Policy*, 27(72), 521-565.

Corsetti, G., Meier, A. and Müller, G.J. (2012b). Fiscal Stimulus with Spending Reversal. *The Review of Economics and Statistics*, 94(4), 878-895.

DeLong, J.B. and Summers, L.H. (2012). Fiscal Policy in a Depressed Economy. *Brookings Papers on Economic Activity*, 233-297.

De Nardis, S. and Pappalardo, C. (2018). Fiscal Multipliers in Abnormal Times: The Case of a Model of the Italian Economy. *Ufficio Parlamentare di Bilancio, Nota di lavoro*, 1.

Dolls, M., Fuest, C. and Peichl, A. (2012). Automatic Stabilizers and Economic Crisis: US Vs Europe. *Journal of Public Economics*, 96(3-4), 279-294.

Fatás, A. and Summers, L.H. (2016). The permanent effects of fiscal consolidations. *National Bureau of Economic Research Working Paper*, 22374.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000). The Generalized Factor Model: Identification and Estimation. *The Review of Economics and Statistics*, 82(4), 540-554.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2005). The Generalized Dynamic Factor Model. One-Sided Estimation and Forecasting. *Journal of the American Statistical Association*, 100(471), 830-840.

Forni, L., Monteforte, L. and Sessa, L. (2009). The general equilibrium

effects of fiscal policy: Estimates for the Euro area. *Journal of Public Economics*, 93(3-4), 559-585.

Forni, L., Gerali, A. and Pisani, M. (2010). The macroeconomics of fiscal consolidations in a monetary union: the case of Italy. *Banca d'Italia Working Papers*, 747.

Gechert, S., and Will, H. (2012). Fiscal Multipliers: A Meta Regression Analysis. *Institut für Makroökonomie und Konjunkturforschung Working Paper*, 97.

Gechert, S. and Rannenberg, A. (2014). Are Fiscal Multipliers Regime-Dependent? A Meta Regression Analysis. *Institut für Makroökonomie und Konjunkturforschung Working Paper*, 139.

Giordano, R., Momigliano, S., Neri, S. and Perotti, R. (2007). The effects of fiscal policy in Italy: Evidence from a VAR model. *European Journal of Political Economy*, 23(3), 707-733.

Guajardo, J., Leigh, D. and Pescatori, A. (2014). Expansionary Austerity? International Evidence. *Journal of the European Economic Association*, 12(4), 949-968.

Hernandez de Cos, P. and Moral-Benito, E. (2016). Fiscal multipliers in turbulent times: the case of Spain. *Empirical Economics*, 50(4), 1589-1625.

Ilzetzki, E., Mendoza, E.G. and Végh, C.A. (2013). How Big (Small?) Are Fiscal Multipliers? *Journal of Monetary Economics*, 60(2), 239-54.

Jordà, O. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review*, 95(1), 161-182.

Jordà, O. and Taylor, A.M. (2016). The time for austerity: estimating the average treatment effect of fiscal policy. *Economic Journal*, 126(590), 219-255.

Karras, G. (2014). Is Fiscal Policy More Effective During Cyclical Downturns? *International Economic Journal*, 28(2), 255-271.

Kilponen, J., Pisani, M. Schmidt, S., Corbo, V., Hledik, T., Hollmayr, J., Hurtado, S., Jlio, P., Kulikov, D., Lemoine, M., Lozej, M., Lundvall, H., Maria, J.R., Micallef, B., Papageorgiou, D., Rysanek, J., Sideris, D., Thomas C. and De Walque, G. (2015). Comparing fiscal multipliers across models and countries in Europe. *European Central Bank Working Paper*, 1760.

Kirchner, M., Cimadomo, J. and Hauptmeier, S. (2010). Transmissions of Government Spending Shocks in the Euro Area. Time Variations and Driving Forces. *European Central Bank Working Paper*, 1219.

Leeper, E.M., Traum, N. and Walker, T.B. (2017). Clearing Up the Fiscal Multiplier Morass. *American Economic Review*, 107(8), 2409-2454.

Parker, J.A. (2011). On Measuring the Effects of Fiscal Policy in Recessions. *Journal of Economic Literature*, 49(3), 703-718.

Perotti, R. (2002). Estimating the Effects of Fiscal Policy in OECD Countries. *European Central Bank Working Paper*, 168.

Ramey, V.A. (2011). Can Government Purchases Stimulate the Economy? *Journal of Economic Literature*, 49(3), 673-685.

Ramey, V.A. and Zubairy, S. (2018). Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical data. *Journal of Political Economy*, 126(2), 850-901.

Riera-Crichton, D., Végh, C.A. and Vuletin, G. (2015). Procyclical and countercyclical fiscal multipliers: Evidence from OECD countries. *Journal of International Money and Finance*, 52, 15-31.

Sims, E. and Wolff, J. (2017). State-dependent fiscal multipliers: Calvo vs. Rotemberg. *Economic Letters*, 159, 190-194.

Stock, J.H. and Watson, M.W. (2002a). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics*, 20(2), 147-162.

Stock, J.H. and Watson, M.W. (2002b). Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association*, 97(460), 1167-1179.

Ufficio Parlamentare di Bilancio. (2017). I consumi intermedi delle Amministrazioni Pubbliche. *Focus tematico*, 3.

Wolff, G.B., Tenhofen, J. and Heppke-Falk, K.H. (2006). The macroeconomic effects of exogenous fiscal policy shocks in Germany: a disaggregated SVAR analysis. *Deutsche Bundesbank Discussion Paper Series*, 1: Economic Studies 41.

Woodford, M. (2011), Simple Analytics of the Government Expenditure Multiplier. *American Economic Journal: Macroeconomics*, 3(1), 1-35.

Chapter 2

Testing evenness multigroup residential segregation indexes using regular lattice data.
A first investigation on the major Italian Functional Urban Areas[1]

Federico Benassi and Alessia Naccarato

2.1 Introduction

Residential segregation of a minority group, whether it be an ethnic group or any subgroup of the population (a certain class of workers, a religious minority, etc.) consists of a spatial distribution that is not uniform with respect to the rest of the population. From a conceptual and operative point of view residential segregation of foreigners should be defined and faced as a multidimensional phenomenon. Massey and Denton (1988) were the first scholars to define segregation as a multidimensional concept, identifying the different dimensions that can be measured by different indexes: evenness, exposure, concentration, centralization and clustering.

Foreigners' residential segregation is a phenomenon widely studied. Restricting the perspective to studies related to Europe, reference can be made to Musterd (2005), Arbaci (2007), Marcińczak *et al.* (2016), Nielsen *et al.* (2017), Nieuwenhuis *et al.* (2017), Anderson *et al.* (2018a, 2018b). Numerous are also the studies dealing with particular territorial (sub)

---

partitions of Europe like the analysis of Arbaci (2008) referring to Southern Europe and the contributions about urban Europe and/or specific European cities: Musterd and Ostendorf (1998), Kesteloot and Cortie (1998), Bolt (2009), Musterd and Van Kempen (2009), Tammaru *et al.* (2016), Musterd *et al.* (2017), van Ham *et al.* (2018), Natale *et al.* (2018), Casacchia *et al.* (2019), Strozza *et al.* (2018). In Italy, studies on the territorial distribution of the foreign population began in the 1980s and then developed to the present day according to different lines of analysis. In particular, the studies conducted with reference to sub-national contexts, such as regions, provinces, metropolitan areas and Local Labour Markets Areas (Cristaldi 2002; Heins and Strozza 2008; Montanari and Paluzzi 2016; Benassi *et al.* 2015, 2018, 2019) have generally reported low level of residential segregation of foreigners who keep Italian cities – and the other territorial contexts – quite far from the experiences of marginalization and segregation found in North American cities. At the same time, although with some exceptions that only recently have been detected (Strozza *et al.* 2018), the models of poorly concentrated settlements identified in Italy are also peculiar within the European context, as they can be traced back to certain elements that are strongly conditioning: the scarcity of housing policies aimed at immigration, the high number of foreign communities, the work placement in segments that affect the way of living, the high social mix of the urban fabric (Natale 2006; Barbieri *et al.* 2019). Rather, differences in the degree of spatial grouping between the different geographical contexts analysed were found (Heins and Strozza 2008; Benassi *et al.* 2015). Generally in Central and Northern Italy the level of residential segregation is more limited than in Southern and Island contexts. In these results it seems that the links with the employment opportunities offered in the various local territories can be intuited, as well as a longer immigration history of the northern local contexts compared to those southern part of Italy. In contexts where the possibilities of employment are more varied and the number of nationalities involved is greater, the degree of spatial concentration for the foreign population is more limited. On the contrary, in contexts where employment opportunities are limited to certain sectors, a greater concentration of the foreign component of the population arise (Strozza *et al.* 2018).

Despite the studies that have appeared on the subject are numerous, with reference to the Italian case most of them make use of one-group or two-group segregation indexes; that is to say measures that do not take into account the complexity that accompanies a society racially different (Reardon and Firebaugh 2002) and therefore that should be considered as measures partially inadequate. Moreover, most of the existing studies do not allow a robust "spatial" comparison between different territories since the elementary data usually refer to territorial administrative contexts characterized by different geographies and therefore potentially affected by the MAUP effect (Gehlke and Biehl 1934; Openshaw 1978, 1984; Reardon and O'Sullivan 2004).

Based on these premises the contribution proposes a first investigation on the residential segregation of foreigners in the major metropolitan areas of Italy using evenness multigroup segregation measures computed on data referred to regular and comparable geographies. The contribution is in fact based on data referred to a geographical regular lattice: grids of 100 meters by 100 meters on which the 2011 census data have been reported. The data were made available by the European Commission - Joint Research Center (JRC)/Knowledge Centre on Migration and Demography (KMCD), within the international data challenge called D4I - *Data challenge on Integration of migrants in cities*[2].

The contribution tries to assess the sensitivity of the multigroup indexes to different geographical contexts and to provide a synthetic picture of the level of residential segregation across the major Italian metropolitan areas in terms of evenness, one of the most important dimension of residential segregation. The paper is organized as follows. Section 2.2 is devoted to the description of the geographical contexts of analysis, data used and measures applied. In section 2.3 results are presented and discussed. In section 2.4 some preliminary conclusions are drawn.

---

[2] Information about the D4I Data Challenge are available at this link: <https://bluehub.jrc.ec.europa.eu/datachallenge/>. The main findings from the Data Challenge are available in Tintori *et al.* (2018).

2.2 Geographical contexts, data and measures

2.2.1 Geographical contexts

The analysis is conducted on a sub set of Functional Urban Areas (FUAs hereafter) of Italy. The Functional Urban Areas (FUAs) are geographical contexts initially proposed by the Organization for Economic Cooperation and Development (OECD) on the base of a defined methodology essentially grounded in the gravitational models. FUAs are in fact composed by a city (or core) and a commuting zone that is functionally interconnected to the city (OECD 2012). FUAs are functional areas that overcome administrative boundaries and thus represent a very useful tool for investigating typically urban phenomena like foreigners' residential segregation.

Italy presents several FUAs scattered all over the national country's surface (Figure 2.1(a)). Our analysis concerns the major FUAs: Rome, Milan, Naples, Turin, Palermo, Genoa, Florence, Bari, Bologna, Catania and Venice (Figure 2.1 (b)). In 2011 more than 19 million people reside in the selected FUAs (32.1% of the total population resident in Italy).

Figure 2.1
Functional Urban Areas (FUAs) of Italy (a) and selected FUAs (b).
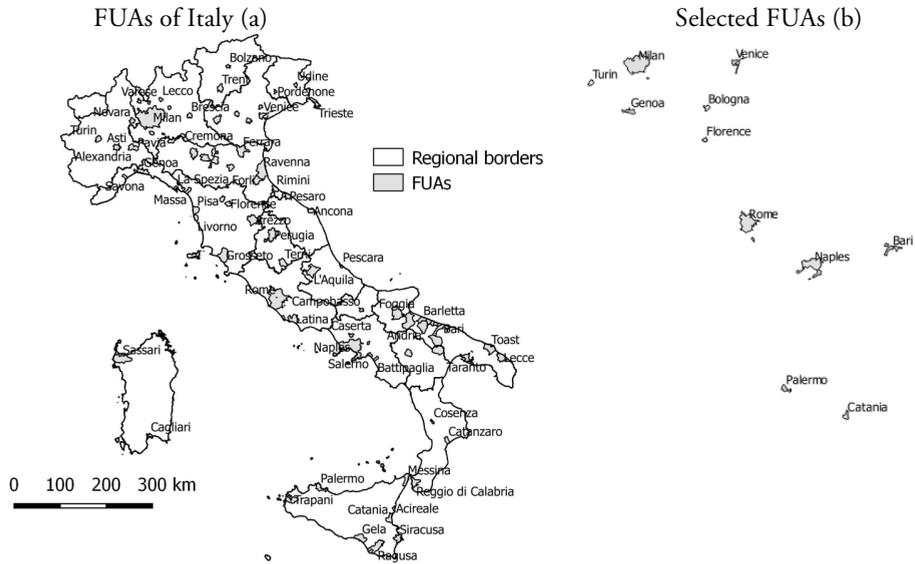
Sources: our own elaboration on OECD data

The selected FUAs are very different one each other's in terms of geographical location, demographic dimension, demographic density and urban inner structures (Table 2.1). Five of them (Milan, Turin, Genoa, Bologna and Venice) are located in the Northern part of the country, the most dynamic macro area of Italy in terms of economic performances and productivity and where the incidence of foreign population is bigger compared to others macro areas of the country (Benassi and Naccarato 2018). Rome and Florence are the only two metropolitan FUAs located in Central part of Italy, while the others four FUAs (Naples, Palermo, Bari and Catania) belong to South and Island division. In terms of demographic dimensions the three bigger FUAs – Rome, Milan and Naples – represent 1/5 of the total population resident in Italy in 2011 and slightly less than 2/3 of the metropolitan population of Italy here analyzed.

Table 2.1 – *Some demographic characteristics of the selected FUAs (2011).*

Functional Urban Areas	Population (absolute values)	Population density (per km^2)	% share of population in metropolitan area over the national value
Milan	4,858,686	1,263	8.2
Rome	4,040,243	656	6.8
Naples	3,345,651	2,230	5.6
Turin	1,727,103	1,015	2.9
Palermo	1,002,547	672	1.7
Florence	756,792	436	1.3
Bari [i]	735,413	652	1.2
Bologna	715,438	386	1.2
Genoa	715,379	598	1.2
Catania	622,959	1,021	1.1
Venice	551,069	833.5	0.9
Total	**19,071,280**	-	-

Source: OECD data base 'Region and Cities – Metropolitan areas' (<https://stats.oecd.org/Index.aspx?Datasetcode=CITIES#>). [i] Data refers to 2012. Data extracted on 15 September 2019.

It is useful to recall that urban and metropolitan areas of Italy are territorial contexts that more than others have gained by the intense growth recorded by the foreign population resident in Italy in the recent past (Strozza *et al.* 2016). For all of the aforementioned reasons we believe that the selected FUAs represent an interesting case study for testing the multigroup residential measures.

2.2.2 Data and Measures

Data used in the contribution are provided by the *Data Challenge on Integration of Migrants in Cities* (D4I). D4I is an initiative launched at the end of 2017 by Joint Research Center (JRC) - Knowledge Centre on Migration and Demography (KMCD) to disseminate to external researchers a data set showing the concentration of migrants in EU cities at high spatial resolution (grid cells of 100 by 100 meters).

This data set was produced on the basis of *ad hoc* extractions of the 2011 Census data provided by the National Statistical Institutes of 8 EU Member States (Tintori *et al.* 2018). We refer to the JRC Technical Report (Alessandrini *et al.* 2017) for more details about methods applied for the processing of the original census data and for other technicalities. In this contribution we use data of D4I referred to Italy at FUAs level. For Italy, data essentially refers to resident population enumerated in 2011 demographic census broken down by country of citizenship at different geographical scales (*i.e.* from the finest one, grid, to the biggest, the national level).

As for the measures we have computed two multigroup indexes of residential segregation. The first one (Dmulti) is the multi-group version of Duncan's dissimilarity index (D) (Morgan 1975; Sakoda 1981). The second one (Hmulti) is the multi-group version of Theil's entropy index (H) (Theil 1972; Theil and Finizza 1971). Both indexes refer to the evenness dimension of residential segregation. This dimension basically refers to the extent to which groups are evenly distributed among organizational units (Massey and Denton 1988) and is normally considered one of the most relevant dimensions in residential segregation.

For a very exhaustive overview on the mathematical and statistical characteristics and properties of each single indicator here mentioned we remind to Reardon and Firebaugh (2002) of which we adopt the same notation in formalizing the indicators, in particular: t denotes size and π denotes proportion. Subscripts i and j index territorial units; and subscripts m and n index group. Hence: t_j = number of cases in territorial unit j; T = total number of cases; π_m = proportion in group m; π_{jm} = proportion in group m, of those in unit j (Reardon and Firebaugh 2002).

The first index – the multi-group version of the Duncan's dissimilarity index – can be write as:

$$Dmulti = \sum_{m=1}^{M} \sum_{j=1}^{J} \frac{t_j}{2TI} \left| \pi_{jm} - \pi_m \right| \qquad (2.1)$$

In equation (2.1) *I* represents the Simpson's Interaction Index (Lieberson 1969; White 1986):

$$I = \sum_{m=1}^{M} \pi_m (1 - \pi_m) \qquad (2.1a)$$

The multi-group version of the Theil's entropy index can be write as:

$$Hmulti = \sum_{m=1}^{M} \sum_{j=1}^{J} \frac{t_j}{TE} \pi_{jm} ln \frac{\pi_{jm}}{\pi_m} \qquad (2.2)$$

In equation (2.2) *E* denotes Theil's Entropy Index (Theil 1972):

$$E = \sum_{m=1}^{M} \pi_m ln \left(\frac{1}{\pi_m} \right) \qquad (2.2a)$$

Both indexes vary from 0 to 1. Values equal to 0 means the absence of residential segregation. On the opposite, values equal to 1 means complete residential segregation. The indexes have been computed by using the R package 'OasisR' (Tivadar 2019).

2.3 Results, presentation and discussion

Palermo, Naples and Catania record, respectively, the highest level of residential segregation both in terms of dissimilarity and entropy. More in details, Palermo records the higher values for both indexes of evenness, 0.6300 and 0.3242 respectively. In relation to *Dmulti*, Naples is in the second place (0.5343), followed by Catania (0.5330) and Bari (0.5097). On the opposite, the lowest values of *Dmulti* are recorded by Rome (0.4455), Bologna (0.4226) and Florence (0.4166).

Table 2.2 – *Dmulti and Hmulti. Major FUAs of Italy. 2011*

Dmulti

Palermo	0.6300
Naples	0.5343
Catania	0.533
Bari	0.5097
Venice	0.4745
Milan	0.4659
Genova	0.4619
Turin	0.4462
Rome	0.4455
Bologna	0.4226
Florence	0.4166

Hmulti

Palermo	0.3242
Naples	0.2783
Catania	0.2685
Rome	0.2584
Bari	0.2515
Venice	0.2369
Milan	0.2345
Florence	0.2273
Genova	0.2245
Bologna	0.213
Turin	0.2044

Source: our elaboration on JRC D4I data
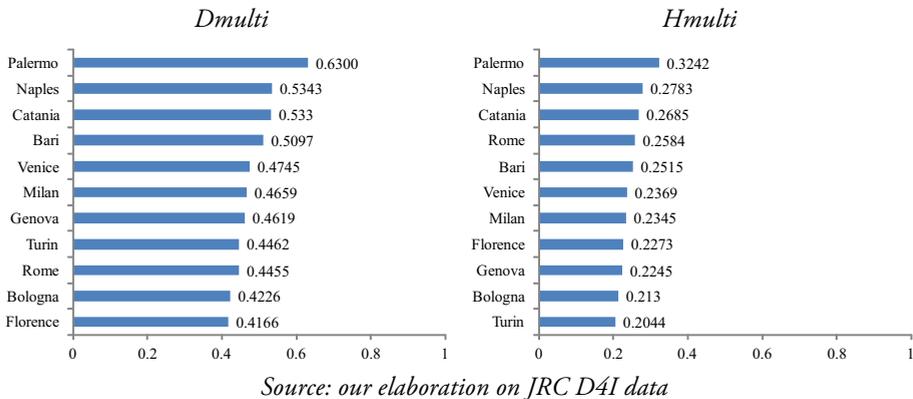
In terms of entropy, Naples follows Palermo in second place (0.2783) while in the third place we find, again, Catania (0.2685). The lowest values for indicator *Hmulti* are those of Genoa (0.2245), Bologna (0.213) and Turin (0.2044).

The results obtained seem to be consistent with those published in previous studies: with a certain degree of variability, the general levels of residential segregation that characterize the Italians' metropolitan areas are not particularly high.

Within this general framework, however, an important variability emerges between metropolitan areas of the Centre-North on the one hand and of the South and Islands on the other indicating the existence of an important South-North divide.

In the FUAs of North and Centre the levels of residential segregation are generally lower than those of the Southern areas and Islands. This could be depend by several demographic and socio-economic factors. In this perspective, an interesting dimension of analysis is how the level of residential segregation can be related to the local labor markets of the local contexts.

In a guise of conclusion we have studied the relationship between the level of multigroup segregation and the unemployment level in each FUAs. The last is a *proxy* indicator of the economic and labour market dynamism but also of attractiveness to the foreign population.

Figure 2.2 gives us an idea of what emerge. Again, the picture is dual: the statistical relationship is positive and 'strong' and this explain, in a certain measure, the reason behind the higher level of residential segregation detected in the FUAs of southern Italy.

Figure 2.2
Multigroup segregation indexes and unemployment rate as a percentage of labour force. Scatter plot. Major Italian FUAs. 2011
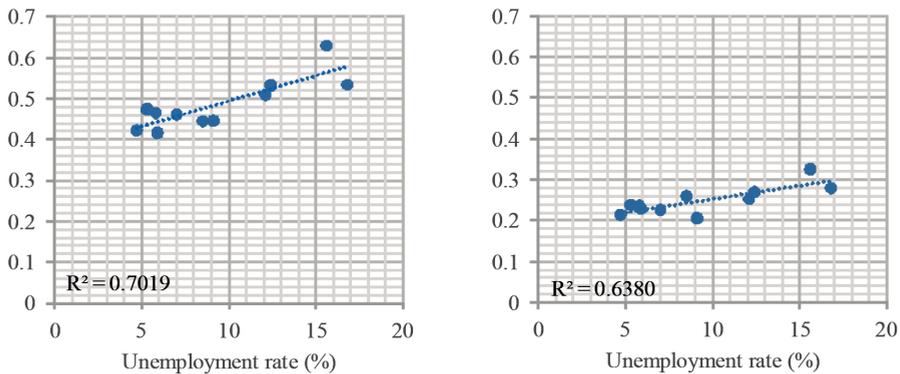
Source: our elaboration on JRC D4I data and OECD data base 'Region and Cities – Metropolitan areas' (<https://stats.oecd.org/Index.aspx?Datasetcode=CITIES#>). Data extracted on 17 September 2019.

2.4 Conclusion

Summarizing, the use of spatial comparable data on the one hand and multi-group segregation indexes on the other, has allowed us to obtain interesting and original results, perfectly comparable from a geographical point of view.

From what has emerged, it seems to be possible to say that less residential segregation and therefore greater territorial integration of foreigners passes through labour (at least in relation to the dimension of evenness). In this perspective a clear South-North distinction clearly emerges.

Higher employment rates improve the lives not only of foreigners but of the society as a whole (at least in macroeconomic terms). These

are obviously the first results that are worth reflecting on and that will be the subject of further analysis and study also through the replication of this analysis to other European FUAs waiting for the release of data from the second round of demographic census.

References

Alessandrini, A., Natale, F., Sermi, F. and Vespe, M. (2017). High resolution map of migrants in the *EU JRC Technical Reports,* EUR 28770 EN, doi: 10.2760/0199.

Andersson, E.K., Lyngstad, T.H. and Sleutjes, B. (2018a). Comparing Patterns of Segregation in North-Western Europe: A Multiscalar Approach. *European Journal of Population*, 34(2), 151-168.

Andersson, E.K., Malmberg, B., Costa, R., Sleutjes, B., Stonawski, M.J. and de Valk, H. (2018b). Comparative study of segregation patterns in Belgium, Denmark, the Netherlands and Sweden: Neighbourhood concentration and representation of non-European migrants. *European Journal of Population*, 34(2), 251-275.

Arbaci, S. (2007). Ethnic segregation, housing systems and welfare regimes in Europe. *European Journal of Housing Policy*, 7(4), 401-433.

Arbaci, S. (2008). (Re)Viewing Ethnic Residential Segregation in Southern European Cities: Housing and Urban Regimes as Mechanisms of Marginalisation. *Housing Studies*, 23(4), 589-613.

Barbieri, G.A., Benassi, F., Mantuano, M. and Prisco, M.R. (2019). In search of spatial justice. Towards a conceptual and operative framework for the analysis of inter – and intra-urban inequalities using a geo-demographic approach. The case of Italy. *Regional Science, Policy & Practice*, 11(1), 109-121.

Benassi, F., Ferrara, R. and Strozza, S. (2015). La reciente evolución de los patrones de asentamiento en las principales comunidades de inmigrantes en Italia. *Papeles de Población*, 21(86), 73-104.

Benassi, F., Heins, F., Lipizzi, F. and Paluzzi, E. (2018). Measuring Residential Segregation of Selected Foreign Groups with Aspatial

and Spatial Evenness Indexes. A Case Study. In Perna, C., Pratesi, M. and Ruiz-Gazen, A. (eds.), *Studies in Theoretical and Applied Statistics*, vol. 227. Springer: Springer Proceedings in Mathematics & Statistics, 189-199.

Benassi, F., Lipizzi, F. and Strozza, S. (2019). Detecting foreigners' residential patterns in urban contexts: two tales from Italy. *Applied Spatial Analysis and Policies*, 12(2), 301-319.

Benassi, F. and Naccarato, A. (2018). Foreign citizens working in Italy. Does space matter? *Spatial Demography*, 6, 1-16.

Bolt, G. (2009). Combating residential segregation of ethnic minorities in European cities. *Journal of Housing and the Built Environment*, 24(4), 397-405.

Casacchia, O., Natale, L. and Verdugo, G. (2019). Minority segregation processes in an urban context: a comparison between Paris and Rome. In Crocetta, C. (eds.), *Theoretical and Applied Statistics. In memory of Corrado Gini*, vol. 274. Springer: Springer Proceedings in Mathematics & Statistics, pp. 83-90.

Cristaldi, F. (2002). Multiethnic Rome: Toward residential segregation? *GeoJournal* 58, 81-90.

Gehlke, C.E. and Biehl, K. (1934). Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, 29(185A), 169-170, doi: 10.2307/2277827.

Heins, F. and Strozza, S. (2008). La geografia insediativa degli stranieri all'interno delle province italiane: differenze e determinanti. In Bonifazi, C., Ferruzza, A., Strozza, S. and Todisco, E. (eds.), *Immigrati stranieri al censimento del 2001*, Studi Emigrazione 171, 573-601.

Kesteloot, C. and Cortie, C. (1998). Housing Turks and Moroccans in Brussels and Amsterdam: The difference between private and public markets. *Urban Studies*, 35(10), 1835-1853.

Lieberson, S. (1969). Measuring Population Diversity. *American Sociological Review,* 34(6), 850-862.

Marcińczak, S., Musterd, S., van Ham, M. and Tammaru, T. (2016). Inequality and rising levels of socio-economic segregation. Lessons from a pan-European comparative study. In Tammaru,

T., van Ham, M., Marcińczak, S. and Musterd, S. (eds.), *Socio-economic segregation in European capital cities: East meets West.* Routledge: London, 358-382.

Massey, D.S. and Denton, N.A. (1988). The dimensions of residential segregation. *Social Forces*, 67(2), 281-315.

Montanari, A. and Paluzzi, E. (2016). Human mobility and settlement patterns from eight EU countries to the Italian region of Lombardy, Veneto, Tuscany, Lazio and Sicily. *Hungarian Geographical Bulletin*, 65(4), 331-344

Morgan, B.S. (1975). The segregation of socioeconomic groups in Urban Areas: A comparative analysis. *Urban Studies*, 12(1), 47-60.

Musterd, S., Marcińczak, S., van Ham, M. and Tammaru, T. (2017). Socioeconomic segregation in European capital cities. Increasing separation between poor and rich. *Urban Geography*, 38(7), 1062-1083.

Musterd, S. and van Kempen, R. (2009). Segregation and housing of minority ethnic groups in Western European cities. *Tijdschrift Voor Economische En Sociale Geografie*, 100(4), 559-566.

Musterd, S. (2005). Social and ethnic segregation in Europe: Levels, causes, and effects. *Journal of Urban Affairs*, 27(3), 331-348.

Musterd, S. and Ostendorf, W. (1998). *Urban segregation and the welfare state: Inequality and exclusion in western cities* (1st ed.). Routledge: London.

Natale, F., Scipioni, M. and Alessandrini A. (2018). Spatial segregation of migrants in EU cities. In OECD (eds.), *Divided Cities: Understanding Intra-urban Inequalities*. OECD Publishing: Paris, 75-109, retrieved from https://doi.org/10.1787/9789264300385-6-en.

Natale, L. (2006). Vicini l'uno all'altro: condividere lo spazio all'interno di Roma. In Sonnino, E. (eds.), *Roma e gli immigrati. La formazione di una popolazione multiculturale*. Milano: Franco Angeli, 165-194.

Nielsen, M.M., Haandrikman, K., Christiansen, H., Costa, R., Sleutjes, B., Rogne, A.F. and Stonawski, M. (2017). Residential segregation in 5 European countries: Technical report. *ResSegr Working Paper*, 2.

Nieuwenhuis, J., Tammaru, T., van Ham, M., Hedman, L. and Manley,

D. (2017). Does segregation reduce socio-spatial mobility? Evidence from four European countries with different inequality and segregation contexts. *IZA Discussion Paper Series*, 11123.

OECD. (2012). Redefining 'urban': a new way to measure metropolitan areas, *OECD Publishing, and Organisation for Economic Co-operation and Development Staff*, 2012.

Openshaw, S. (1984). *The Modifiable Areal Unit Problem*, Norwich: Geo Books.

Openshaw, S. (1978). An empirical study of some zone-design criteria. *Environment and planning A*, 10(7), 781-794.

Reardon, S.F. and O'Sullivan, D. (2004). Measures of Spatial Segregation. *Sociological Methodology*, 34(1), 121-162.

Reardon, S.F. and Firebaugh, G. (2002). Measures of multi-group segregation. *Sociological Methodology*, 32(1), 3-67.

Sakoda, J.N. (1981). A generalized index of dissimilarity. *Demography*, 18(2), 245-250.

Strozza, S., Benassi, F., Bonifazi, C., Heins, F. and Lipizzi, F. (2018). Comparing residential segregation: selected origins in selected EU metropolitan areas. In Tintori, G., Alessandrini, A. and Natale, F. (eds.), *Diversity, residential segregation, concentration of migrants: a comparison across EU cities. Findings from the Data Challenge on Integration of Migrants in Cities (D4I)*, EUR 29611 EN, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-98732-8, doi:10.2760/ 823648, JRC115159.

Strozza, S., Benassi, F., Ferrara, R. and Gallo, G. (2016). Recent demographic trends in the major Italian urban agglomerations: The role of foreigners. *Spatial Demography*, 4(1), 39-70.

Tammaru, T., van Ham, M., Marcińczak, S. and Musterd, S. (2016). *Socio-economic segregation in European capital cities: East meets West*. Routledge: London.

Theil, H. (1972). *Statistical decomposition analysis with application in the social and administrative* (1st ed.). Amsterdam: North-Holland.

Theil, H. and Finizza, A.J. (1971). A note on the measurement of racial integration of schools by means of information concepts. *Journal of Mathematical Sociology*, 1, 187-194.

Tintori, G., Alessandrini, A. and Natale F. (2018). *Diversity, residential segregation, concentration of migrants: a comparison across EU cities. Findings from the Data Challenge on Integration of Migrants in Cities* (D4I), EUR 29611 EN, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-98732-8, doi:10.2760/823648, JRC115159.

Tivadar, M. (2019). OasisR: An R package to bring some order to the world of segregation measurement. Journal of Statistical Software, 89(7), 1-39, doi: 10.18637/jss.v089.i07.

van Ham, M., Tammaru, T. and Jannsen, H.J. (2018). A multi-level model of vicious circles of socio-economic segregation. In OECD (eds.), *Divided Cities: Understanding Intra-urban Inequalities*. OECD Publishing: Paris, 135-153, available at https://doi.org/10.1787/9789264300385-8-en.

White, M. (1986). Segregation and Diversity Measures in Population Distribution. *Population Index,* 52(2), 198-211.

Chapter 3

On the social irrationality of the limits to public debt

Roberto Ciccone

I. Aggregate demand and lacking social rationality of public debt reduction policies

1. In the wake of Keynes's *General Theory*, a large body of works developed in the 1940s and '50s which predicated the use of fiscal policy, and particularly of public budget deficits, in order to raise the levels of global output and employment[1]. The premise of this literature, which had perhaps in Lerner's 'functional finance' (Lerner 1943) its most radical expression, was that private aggregate demand could not be sufficient to absorb full-employment output, and theorized the recourse to public deficit spending as an effective tool to achieve the aim.

In our view a return to those positions, for long quite alien to mainstream thought[2], though now recovering credit under the impact of recent economic downturns, is necessary and urgent. And all the more so inasmuch as that approach can be reinforced by placing it into a consistent theoretical setting altogether different from the traditional one, something which at the time was not at hand and today is allowed by the advances economic analysis has made along lines alternative to neoclassical theory – we particularly refer to the modern reappraisal of

---

[1] A representative group of these contributions is collected in Kaounides and Wood (1992), Vol. II, Part I.

[2] Though not majority, the functional finance perspective of fiscal policy has however continued to be shared hitherto: representative cases are the works by economists at the Levy Economics Institute in the US and by Philip Arestis and Malcom Sawyer in Europe.

the classical theory of prices and distribution started by Piero Sraffa (1960) and further developed by later scholars.

In this perspective the present paper adopts two fundamental theoretical premises, closely linked to each other. The first one is the adhesion to an explanation of income distribution in terms of social and institutional circumstances as found in the classical authors, and the consequent rejection of the theory of distribution based on forces of demand for, and supply of, productive 'factors' which is proper to neoclassical theory. The second premise lies in the idea that the long-run levels of output no less than their fluctuations are ruled by the size of aggregate demand, whose ultimate determinants are in turn conceived as independent of the potential output of the economic system. This view of the role of demand would evidently be inconsistent with the neoclassical explanation of distribution, in which output is constrained by the clearing condition in factor markets and demand can pull it below that level only temporarily or in case of price rigidities. This is why the concept of demand as independent from potential output finds in the first premise a necessary complement, a socially regulated division of the product being fully compatible with the systematic limitation of output by the extent of demand and with its corollary in the labour market, persistent involuntary unemployment[3].

Once the expenditure of the community is viewed as the *normal* determinant of the level of output, the outlay of the public sector, which concurs both directly and indirectly to form total expenditure, is consequently conceived as one of the elements systematically affecting aggregate production. If public expenditure exceeds the revenues of the public sector, the occurrence of budget deficits obviously engenders (leaving aside monetary financing) the accumulation of public debt. Now, admitting the influence of public spending on output is crucially relevant to the impact of public deficits and debt on the economic

---

[3] The work of Pierangelo Garegnani is the fundamental reference for the integration of the classical approach to the explanation of distribution, in its contemporary resumption, with the Keynesian conception of demand as the determinant of the levels of activity. Garegnani (1978) and Garegnani (1979) provided the starting point of that line of thought.

system. It should be pointed out in this regard that in the conceptual structure adopted here, and referred to above, there is no reason to assume that public deficits necessarily generate reductions in one or other component of private expenditure, as is often claimed by traditional economic analysis. The elasticity of output, and therefore of the formation of savings, excludes any scarcity of the latter, and therefore any need for increases in the interest rate and consequent negative influence on private investment (an influence which, moreover, is subject to be questioned, in its own general analytical basis, by the results of the capital theory debate[4]). The same elasticity of output and volume of savings would, on the other hand, make it unnecessary for the private sector to compress its consumption expenditure even if the hypotheses underlying the so-called 'Ricardian equivalence' were accepted[5]. Critical implications for those assumptions will emerge, however, precisely from what will be argued in the paper.

A central point in our perspective is therefore that, through their effect on the level of income, public deficits generate *additional* private savings, *i.e.* savings which would not materialize in the absence of those deficits; correspondingly, increases in the stock of public debt are mirrored by equivalent increases in the stock of private wealth. It follows that, as far as aggregate production is supposed to respond to aggregate demand, and referring for greater clarity and simplicity to a closed economy, no limits to public deficits and debt can be detected in the willingness to save of the private sector or in the size of private wealth, nor any trade-off with capital accumulation need arise – on the contrary, private investment could be propelled by the higher levels of aggregate demand fed by public spending.

The accumulation of public debt can thereby be the means by which the economy enjoys higher levels of income as well as of private wealth, in both the financial and real components of the latter. On the other hand, so far economic analysis has not identified any generally

---

[4] Cf. Garegnani (1978), paras. 5-7.

[5] See Ciccone (2013), par. 1.2, pp. 16-20 for a more detailed discussion of the issues just mentioned.

valid boundary for the dimension which public debt may reach, nor any definite threshold beyond which it would produce effects conceivable as damaging to society as a whole.

2. On account of the direct influence of public spending on the levels of output, as well as of public deficits on the formation of private wealth, one can question the social rationality of policies aimed at reducing the size of public debt, at least so long as the levels of private demand do not persistently stand so high relative to potential output as to make it advisable to run budget surpluses. To the extent deficit spending occupies a vacuum of demand, there would be no interest of society in diminishing it, even less reversing it into a surplus, unless the entailed losses of higher levels of income and wealth would be adequately counterbalanced. However, in the conceptual framework adopted here no obvious compensation of such losses appears to be warranted by a reduction of public deficit and/or debt, and we maintain that in the face of activity levels constrained by the size of demand, policies pursuing the diminution or even just the limitation of public debt would be irrational from the point of view of the community[6].

The degree of social rationality of the policies under discussion does not prove higher, and can even appear lower, if the deficit and debt targets are defined as ratios to domestic product. Owing precisely to their negative impact on demand, the efficacy of those policies in reducing deficit and debt ratios can be seriously endangered, not only temporarily, by their unfavorable influence on GDP. To the insanity concerning

---

[6] One might think that a positive effect for society as a whole of a reduction in public debt would result from lower interest rates. A low interest rate on public debt may indeed be in the overall concern of the community, interest payments representing 'contractual' disbursements of the Treasury – as such removed from political deliberation, hence lacking the social legitimacy of other categories of expenditure and, moreover, of reasonably much lesser impact on aggregate demand. However, whether public debt reduction is able to lower the interest rate is questionable, especially in the absence of undisputable theoretical bases for establishing such a connection. Rather, what the observation suggests to be extremely significant for the conditions under which public debt is placed is the action of the Central Bank, a fundamental task for monetary authority within the overall economic policy framework of the country.

income and wealth those policies would thereby add 'perverse' results in respect to the public finance objectives they would be designed to hit[7].

II. Lacking social rationality of the (Zero) Future Debt Condition applied to public debt

3. The defect of social rationality in the prescription of abstract boundaries for deficit spending and debt can be found even in what is, as far as we know, the only notion of 'sustainability' of public debt equipped with theoretical foundations provided by the literature. Although without specifying any clean-cut limit for the dimension of public debt, that concept deduces from the maximizing behaviour of individuals, hence from their rationality, the commitment for budget policy to shift over time from a deficit course to a primary (at least) surplus path. In the following part of the paper the validity of that notion of sustainability will be questioned by objecting precisely to the alleged rationality of the private behaviour put at its basis. Our main point is that with output levels limited by demand that behaviour would have nothing rational, and therefore cannot be legitimately presupposed. However, we will firstly note that on account of the specific nature of public debt as private asset, the rationality of such bearing of agents cannot be unconditionally acknowledged even disregarding any influence of aggregate demand on income – as applies to the neoclassical setting, which is the theoretical context that kind of behaviour properly belongs to.

4. The notion of sustainability increasingly adopted in the literature, especially at the theoretical level[8], states that the public debt, and more generally fiscal policy, moves along a sustainable path if the current amount of debt is equalized by the present value of the future primary

---

[7] Cf. Ciccone (2002) and (2013).

[8] See for instance Romer (2012), p. 586.

surpluses realized by the public sector[9]:

$$B_0 = \sum_{1}^{+\infty} (T_t - G_t)(1 + i)^{-t} \qquad (3.1)$$

where B_0 is the stock of public debt at the (end of) the current period, T_t and G_t are the flows of, respectively, revenues from taxation and public expenditure, and i is the (constant) interest rate paid on public debt.

This condition may appear quite intuitive, as at first sight it closely resembles a redemption requirement – but precisely such similarity should actually ring an alarm bell, since one would expect the notion of sustainability expressing something different from a mere refunding capacity. However, on close inspection one can realize that, taken literally, the above mentioned condition imposes a rather weak, if not ineffective, constraint on fiscal policy, by resolving itself in the requirement for the public sector to carry out a sequence of primary surpluses which, however, the infinite time horizon leaves unspecified both in the initial moment and in the size of the surpluses themselves. Even leaving aside that consideration, the question of whether a rational foundation exists for asking fiscal policy to satisfy that condition, is answered differently according to the theoretical framework one adopts. And, as will be argued, the need, or just the collective avail, to subjugate the Government budget to the constraint at issue are contingent to specific conditions even within the most favorable analytical context, which is the neoclassical one – whereas they are definitely missing in a different theoretical approach.

In order to put forward more in detail what has been just hinted about the sustainability condition, let us consider how it is obtained, using Barro (1989) as a reference. The starting point is the so called *intertemporal budget constraint* of the public sector, actually consisting in

---

[9] This condition is sometimes referred to as a 'solvency' requirement, distinguishing it from (various) 'sustainability' conditions that are less theoretically defined but more immediately applicable in practice. Cf. e.g. Wyplosz (2011), p. 4.

the accounting identity which, for any length of time, connects the initial stock of debt, the final stock and the succession of expenditure (of any kind, including transfers) and revenue flows of the sector. Assuming a constant interest rate, and abstracting from monetary financing of public deficits, in each period of the time interval 0-H the following equalities hold:

$$
\begin{aligned}
B_1 &= B_0(1 + i) + G_1 - T_1 \\
B_2 &= B_1(1 + i) + G_2 - T_2 \\
B_t &= B_{t-1}(1 + i) + G_t - T_t \\
&\ldots \ldots \ldots \ldots \ldots \ldots \ldots \\
B_H &= B_{H-1}(1 + i) + G_H - T_H
\end{aligned}
\tag{3.2}
$$

where $B.$, $G.$ and $T.$ stand, respectively for the stock of public debt, public expenditure and public revenues.

By means of obvious substitutions, the set of these equalities reduces to the single equality:

$$
B_H = B_0(1 + i)^H + \sum_{1}^{H} (G_t - T_t)(1 + i)^{(H-t)}
\tag{3.3}
$$

and, after dividing by $(1 + i)^H$ and reorganizing the terms, that same identity is expressed as equality between the present value of the debt stock and the discounted values (with a discount rate equal to the interest rate) of the sum of future budget surpluses and the final debt stock:

$$
B_0 = \sum_{1}^{H} (T_t - G_t)(1 + i)^{-t} + B_H(1 + i)^{-H}
\tag{3.4}
$$

Finally, by extending to infinity the time interval, the relation becomes:

$$B_0 = \sum_{1}^{\infty}(T_t - G_t)(1 + i)^{-t} + \lim_{H \to \infty} B_H(1 + i)^{-H} \qquad (3.5)$$

which shows the sum of the discounted value of total future budget surpluses and the discounted limit value of future debt equalizing the value of the present stock of debt[10].

The latter equality expresses the budget constraint of the public sector in the form commonly used in the literature. Neither the presence of discounted magnitudes, nor the infinity of the time interval it ideally covers should lead us astray from the fact that we are still in the presence of a mere accounting identity, likewise the equalities it is obtained from. (This may be even clearer if the extension to an infinite time interval is applied to current, rather than discounted, values, in which case the equality takes the form:

$$\lim_{H \to \infty} B_H = \lim_{H \to \infty} B_0(1 + i)^H + \lim_{H \to \infty} \sum_{1}^{H}(G_t - T_t)(1 + i)^{H-t} \qquad (3.6)$$

transparently showing that the two sides of the equality are just different ways to express the limit value, whether finite or infinite, of the future stock of debt.)

The identity nature of the equality entails that it is bound to be verified in any case, so that the term 'constraint' appears to be a misnomer – no coercion on the Government budget being imposed by the equality other than that arising from arithmetic. The equality ceases to be an identity and acquires the nature of a restriction only when the condition is imposed that:

$$\lim_{H \to \infty} B_T(1 + i)^{-H} = 0 \qquad (3.7)$$

with the equality consequently turning into:

---

[10] Equality (3.5) above corresponds to equality (5.22) in Barro (1989), p. 203.

$$B_0 = \sum_1^\infty (T_t - G_t)(1 + i)^{-t} \qquad (3.8)$$

which is the sustainability condition mentioned at the beginning.

The requirement that the sum of the present values of future budget surpluses equals the present stock of debt therefore ensues from stating that, as established by (3.7), the present value of the future stock of debt tends, in the limit, to be nil. Hence it is the latter obligation, which hereafter will be referred to as FDC (Future Debt Condition) which constitutes the actual content of the notion of sustainability we are discussing, and it is its premises and implications which we need to look at in order to assess the relevance of the deriving concept of sustainability.

5. The meaning of the FDC is less obvious than might appear at a glance. In the first place this condition does not set any upper boundary to the size of public debt, for it is consistent with an unlimited growth of the latter, provided the debt grows at a rate lower than the interest rate. Assuming, for the sake of simplicity, that the debt grows at the constant rate γ and that the interest rate is in turn constant, the present value of future debt over an infinite time horizon would be:

$$\lim_{t \to \infty} \frac{B_0(1 + \gamma)^t}{(1 + i)^t} = B_0 \lim_{t \to \infty} \left[\frac{1 + \gamma}{1 + i}\right]^t \qquad (3.9)$$

For $\gamma < i$ the fraction in brackets would be lesser than unity, and the value of the limit would be zero.

A growth of public debt at a rate lower than the interest rate implies that primary surpluses are obtained in the budget of the public sector, and that therefore taxes cover the whole public expenditure and at least part of interest payments. This requisite, which the constraint we are discussing reduces to, is not particularly cogent, especially in consideration of the indefiniteness of the time from which it should be satisfied. Moreover, the infinite time interval along which the condition is spread sets no lower limit to the size of the required primary surpluses, which

could be however small, provided their sign is positive.

We can thus see that imposing the tendency to zero of the present value of future debt does not offer a concept of sustainability which appears immediately relevant, and this precisely on account of the infinite time horizon which that condition is set into. One might think that referring to an infinite time interval is a purely abstract aspect, inessential to the content and relevance of the condition. But little reflection suffices to see that if future debt were to refer to a finite date, rather than an indefinitely distant one, the cancellation of the *present* value of the future debt would entail a zero value of its *current* value at that date. Thus, it appears that the notion of sustainability we are dealing with does not escape the alternative between a hardly specifiable content and the drastic (and even trivial) coincidence with the total repayment of the debt by an arbitrarily definite time.

If it is doubtful that the FDC may be taken as a requirement for the sustainability of public debt, a view not uncommonly shared in the literature[11], what seems to be even less glimpsed is the logical weakness of the premises (not always made explicit) from which that condition is taken to descend.

As will be better detailed in the following paragraph, the FDC is put forward as an implication of the maximizing behaviour of the 'representative agent' – representative, in the present context, of the whole private sector – that plans consumption and savings over an infinite time horizon. It is maintained that a positive present value of the net assets held in the infinitely distant future would not be optimal, because it would be tantamount to indefinitely foregoing potential additional consumption (hence utility). So her planning will be such as to reduce to zero, to the limit, the present value of her future assets, including public bonds, which correspondingly imposes the zero present value for the net future liabilities of the public sector.

The FDC would thus derive from the analysis of consumer behaviour, typical of the neoclassical theoretical setting. Before arguing

---

[11] Cf. for instance Wyplosz (2011), Bagnai (2004), Chalk e Hemming (2000), Spaventa (1988).

that in the different approach here adopted there would be no reason for the private sector to want a zero present value of its net wealth, we can notice that even within the neoclassical framework that behaviour loses its generality, if not its consistency, when it is taken to apply to the share of private assets consisting of public debt.

6. Barro (1989) offers a clear exposition of the theoretical foundation of the FDC[12].

Abstracting, for a first step, from the existence of public debt, the intertemporal budget constraint of the 'representative agent' for a time interval ending at the finite date H can be expressed in this way[13]:

$$K_0 + \sum_{1}^{H} (1 + i)^{-t} W_t = \sum_{1}^{H} (1 + i)^{-t} C_t + (1 + i)^{-H} K_H \qquad (3.10)$$

where K_0 e K_H represent the capital stock owned by the agent at, respectively, the initial date 0 and the final date H, W_t is the labour income (wage) produced during the generic period t (with the quantity of labour employed in each period taken as constant and equal to 1), C_t represents the agent's consumption in period t, and i is the constant interest rate, assumed to be equal to the rate of return on capital. In (3.10) the present value of the resources available to the agent along the time interval 0-H, amounting to the initial capital stock and the sum of the present values of income labour, equals the present value of the uses of those resources, namely consumption flows and capital existing at the final date.

---

[12] Barro's contribution, as well as other works we will also refer to, sets out relations between dated aggregate values with no explicit consideration of price variables, neither for products nor for assets, with the sole exception of the interest rate. The neoclassical approach of these elaborations suggests their broad interpretation as belonging to time sequences of general equilibrium systems with their implied price sets, although it could not be specified which species of equilibria – if temporary or intertemporal. In any case, on account of the critical purpose of the present work, we will in turn abstract from any issues relating to product and asset prices.

[13] Cf. the corresponding equality (5.4) in Barro (2009), p. 180.

Then, under the usual assumption of a utility function increasing in the levels of consumption C_t,

«the maximization of utility entails setting $[K_H] = 0$ (which is a transversality condition for this problem)[14]. Otherwise people would leave behind some resources that could have been used to raise consumption and hence utility at an earlier date» (Barro 1989, p. 180).

By imposing the condition $K_H = 0$[15] the agent's budget constraint takes the form:

$$K_0 + \sum_{1}^{H} (1+i)^{-t} W_t = \sum_{1}^{H} (1+i)^{-t} C_t \qquad (3.11)$$

Note that underlying the cancellation of the present value of the capital stock owned at the final date there is the faculty of the agent to 'consume' capital. Under the full-employment conditions proper to the neoclassical theoretical setting the private sector could in principle do that by reducing gross savings, hence investment, below depreciation, with corresponding decrease in the capital stock and corresponding increase in the share of product allotted to consumption.

In Barro's view an infinite time horizon – interpretable «as reflecting altruistic linkages from parents to children to grandchildren, and so on» (Barro 1989, p. 179) – can be treated in analogy to the case of a finite terminal date:

«The same kind of result applies when the horizon is infinite. In this case the transversality condition is that the term $[(1+i)^{-H} K_H]$ approaches zero as H approaches infinity. That is, the family does not leave over asymptotically any resources that have a finite, positive present value» (Barro 1989, pp. 180-1).

---

[14] Transversality conditions guarantee the optimality of the solution in dynamic economic models. Cf. Becker (2008) and Kamihigashi (2008). See also Bellino (2013).

[15] In the context of a finite time horizon the zero present value of $(1+i)^{-H} K_H$ obviously requires K_H to be zero.

According to Barro the condition of utility maximization over an infinite time interval is therefore:

$$\lim_{H \to +\infty} (1 + i)^{-H} K_H = 0^{16} \qquad (3.12)$$

and condition (3.11) keeps on holding, with the only difference that the sums figuring in it consist of infinite numbers of terms[17,18].

---

[16] A negative present value of the household's future assets (namely, the household being indefinitely in a net debtor position), the so called 'Ponzi game', is excluded by assumption in Barro (2009), p. 180. In Barro and Sala-i-Martin (2004) that possibility is ruled out with the argument that it would presuppose the willingness of another household to be a perpetual net creditor, in conflict with the maximisation of the latter's utility (p. 92). In Blanchard and Fisher (1989) the 'no-Ponzi-Game' condition is formulated in terms of constraint on the household's debt growth rate, to be kept lower than the interest rate. In this way the zeroing of the present value of future liabilities is guaranteed by the same condition of transversality as for the present value of future assets – namely, the latter value, whether positive or negative, must grow at a rate lower than the interest rate (Blanchard and Fisher 2009, pp. 49-50; see Kamihigashi 2008, p. 386 for a standpoint adverse to identifying 'no-Ponzi-game' and transversality conditions).
It would seem, however, that arguments such as those mentioned above are entirely unnecessary in the present context. The aggregation of the private sector into a 'representative agent', together with abstraction from dealings with outside economies, is sufficient to preclude the existence of private debt as such even in each single period – at least as long as the private sector is prevented from being debtor, rather than creditor, towards the public sector.

[17] Equation (3.11) above, extended to an infinite number of periods, corresponds to the equation (5.5) which Barro (1989, p. 181) writes in terms of per capita values. In Barro's analysis the finite value of the sums that appear there (providing, respectively, the present values of wage and consumption flows) is in fact assured by the assumption that in the steady state to which the system converges per capita variables are constant, together with the positivity of the interest rate, the equilibrium value of which equals the positive rate of temporal preference (cf. Barro 1989, pp. 179 and 181). In terms of absolute magnitudes, this is tantamount to the general condition that along the steady state the interest rate exceeds the growth rate. As shown in Barro and Sala-i-Martin (2004, p. 92), the latter requirement is fulfilled by the transversality condition adopted in the present context (see also fn. 18 below).

[18] The zero limit for the present value of future capital as a condition for utility maximization, and its interpretation in analogy with the case of finite horizon, are also present e. g. in Barro and Sala-i-Martin (2004) and Blanchard and Fisher (1989),

7. Let us now introduce the activity of the public sector and its effects on the intertemporal budget constraint of the private sector, figured by the infinitely living 'representative agent'. Be B_0 and B_H the values of the public debt at respectively the initial and final date, T_t the taxes levied and G_t the public expenditure in the generic period t. Referring to an

---

though in slightly more general forms and with some more caution than in Barro (1989). Barro and Sala-i-Martin (2004) set the condition at issue in the form (p. 90):

$$\lim_{t \to \infty} v(t) \cdot a(t) = 0$$

where $v(t)$ is «the value of the increment of income received at time t in units of utils at time 0» (p. 89), and $a(t)$ represents the household's net assets «measured in real terms, that is, in units of consumables» (p. 88). The zero value for the limit of the product $v(t) \cdot a(t)$ therefore imposes the zero value of the loss of utility deriving from indefinitely leaving income in the form of assets, rather than consuming it – which, for a however positive marginal utility $v(t)$, entails the tendency to a null value of the amount of household's assets $a(t)$. The two authors also propose the correspondence with the finite horizon case:

> «If we think of infinity loosely as the end of the planning horizon, the intuition is that optimizing agents do not want to have any valuable assets left over at the end. Utility would increase if the assets, which are effectively being wasted, were used instead to raise consumption at some dates in finite time» (Barro and Sala-i-Martin 2004, p. 92).

The authors admit that the literature has identified circumstances in which the transversality condition on infinite horizon they refer to is not necessary, or is not the one necessary, for optimisation, and to which, therefore, the extension to the limit of the finite horizon condition is not applicable (*ibid.*, p. 92, fn. 8, and pp. 613-615). Nevertheless, they end up assuming that the transversality condition they are considering «is a necessary condition for optimization in our infinite-horizon problems» (p. 615).

In Blanchard and Fisher (1989) the transversality condition in infinite horizon (p. 40, eq. 8) is substantially the same as that adopted by Barro and Sala-i-Martin, and in turn would be:

> «best understood by considering the same maximization problem with the infinite horizon replaced by a finite horizon T. […] The infinite horizon transversality condition… can be thought of as the limit of [the finite horizon one] as T becomes large» (Blanchard and Fisher 1989, p. 43).

However, even Blanchard and Fisher do not fail to mention in footnote that «as intuitive as this argument for the transversality condition is, there are infinite horizon problems in which the transversality condition is not necessary for the optimal path» (p. 82, fn. 8).

infinite time horizon, condition (3.11) (in which, remember, the limit of the present value of future capital is set to zero) turns into:

$$K_0 + B_0 + \sum_1^{+\infty}(1+i)^{-t}W_t = \sum_1^{+\infty}(1+i)^{-t}C_t +$$

$$+ \sum_1^{+\infty}(1+i)^{-t}T_t + \lim_{H \to +\infty}(1+i)^{-H}B_H \qquad (3.13)$$

The meaning of the latter equality should be evident: at the left-hand side the present value of the resources available to the private sector now includes the initial stock of the public debt, which of course constitutes an asset of the private sector, while at the right-hand side the present value of the uses now includes the sum of future taxes and the future stock of public debt held by the household.

Correspondingly, the following equality[19] must hold, as we already know, for the budget of the public sector:

$$B_0 = \sum_1^{+\infty}(1+i)^{-t}T_t - \sum_1^{+\infty}(1+i)^{-t}G_t +$$

$$+ \lim_{H \to +\infty}(1+i)^{-H}B_H \qquad (3.14)$$

The step subsequently made by Barro is to impose in the household's budget constraint (3.13) the tendency to zero for the present value of the future public debt, on the same grounds as for setting to zero the limit of the present value of future capital:

«[the] condition for an individual's optimization problem, $[\lim_{H \to +\infty}(1+i)^{-H}K_H]$, ensured that people did not leave over any resources that asymptotically had positive present value. But the public debt is held by individuals as part of their assets; so it follows from the same argument that $[\lim_{H \to +\infty}(1+i)^{-H}B_H]$ must approach zero asymptotically» (Barro 1989, p. 203).

The unwillingness of the private sector to hold a positive present value of future public debt, which implies the tendency to limit to a

---

[19] Analogously to equality (5.22) in Barro (1989, p. 203).

sufficient extent subscriptions of newly issued bonds[20], entails the cancellation of the corresponding item in equality (3.5) above, so that the Government's intertemporal budget constraint becomes equality (3.8), which is the sustainability condition here under discussion[21].

However, the extension to the public debt of the condition which cancels the present value of future capital relies on a *false* analogy, which does not take into account the different nature of the two kinds of assets. As previously noted, it is true that reducing to zero the present value of future capital the private sector would be able, in principle, to correspondingly increase the present value of future consumption flows. The same does not necessarily hold, even within the same theoretical context adopted by Barro, for the credit of the private sector towards the public sector. The switch from a positive present value of the future stock of public debt to a null one requires an equivalent increase in the present value of future primary surpluses of the public sector – hence either an increase in the present value of future taxes or a reduction in the present value of future flows of public expenditure, or a combination of the two.

Turning to the budget constraint of the private sector, and in particular to the r. h. s. of (3.13), in the first case the cancellation of the present value of future public debt ($\lim_{H \to +\infty} (1 + i)^{-H} B_H$) would be entirely compensated by an increase in the present value of taxes ($\sum_1^{+\infty} (1 + i)^{-t} T_t$), with no increase in the present value of consumption levels: indeed, the private sector would be paying the additional taxes by forgiving its credit towards the public sector – the same result arising, for the private sector as a whole, in the event of the mere repudiation of the public debt. It follows that, in the case under consideration, for the 'representative agent' the reduction to zero of the present value of future wealth consisting of public bonds would not allow her to obtain any increase of utility. Hence there would be no

---

[20] Specifically, the subscription of new issues against the debt tranches coming due should be such that the growth rate of the stock of securities held by the private sector is kept below the interest rate.

[21] Cf. Barro (1989), p. 204, eq. (5.23); Blanchard and Fisher (1989), p. 55, eq. (24).

reason for refusing the indefinite position of net creditor of the public sector, which would entail absurdly preferring part of her assets be just cancelled out rather than owned perpetually.

　　Implications somewhat more complex would arise if the present value of future debt were brought to zero through a corresponding reduction in the levels of public expenditure, the present value of taxes remaining unchanged. In this case the budget constraint of the 'representative agent' would reckon either a diminution, on the l. h. s. of (3.13), of the present value of labour incomes ($\sum_1^{+\infty}(1+i)^{-t}W_t$), or, on the r. h. s., an increase of the present value of the levels of consumption ($\sum_1^{+\infty}(1+i)^{-t}C_t$). The full-employment condition proper to the neoclassical context entails the second type of adjustment, since it implies that the resources left free by lesser public uses would be transferred to private uses, which on account of the zero present value of final capital could only be allotted to the production of consumption items. Thus, the present value of (private) consumption levels would be raised in counterpart of the cancellation of the present value of future public debt, similarly, it might seem, to what obtains by reducing to zero the present value of final capital. At a closer inspection, however, even such analogy proves to be unwarranted. Indeed, the private sector would produce and consume more goods, but this would take place at the expense of the production and consumption of goods otherwise provided by the public sector. It follows that the utility enjoyed by the 'representative agent' could be either increased or decreased by such replacement, according to the net effect occasioned by the change in the composition of her consumption basket. Differently from what the agent would be able to get by letting the (present) value of future capital moving to nil, she would not necessarily obtain a gain of utility by reducing to zero the (present) value of her future portfolio of public bonds.

8. Thus, even without questioning the theoretical framework in which the argument is encircled, no character of generality can be attached to the alleged interest of the 'representative agent' in eventually nullifying her credit towards the public sector. As we have seen, that interest would be unjustified, should the cancellation of the present value of public debt

be obtained by an increase in the present value of taxes, and it *could* exist in the case of a reduction in the present value of public expenses, depending on the algebraic sign of the change of total utility attainable by substituting private consumption for public consumption. Hence, even assuming, for the sake of the argument, that change be positive, the preference of the agent for a null present value of future public debt would decisively depend on the actual policy (tax increase or expenditure reduction) by which that null value would be obtained.

We can conclude that for a null present value of public debt to be viewed as optimal by the 'representative agent', two conditions should be satisfied:

i) the agent must be confident that the Government will pursue the result by reducing public expenditure rather than raising taxes;

ii) the agent strictly prefers the (private) consumption goods whose production will be increased, to the (public) consumption goods whose production will be decreased.

There is no reason to regard as granted the fulfilment of either condition. True, one can conceive of arguments which may be relevant in those respects: as to the first condition, we might presume that, through institutional instruments (such as choosing its representatives in political organisms), the private sector would be able to address the choice of the fiscal policy to be adopted; with regard to the second condition, the *cliché* is at hand that for goods and services produced by, or on behalf of, the public sector the quality/cost ratio is lower than for those produced by the private sector. What in any case drops is the alleged *necessity* for the private sector, as pure reflection of its rational conduct, of escaping from being indefinitely in the position of net creditor of the public sector. The two requirements mentioned above entail that the inconvenience of such position would be contingent on the occurrence of specific factual circumstances, in defect of which the rational behaviour of the private sector would be consistent with, or even ask for, the perpetual maintenance of public bonds in its portfolio.

9. Taking the standpoint of neoclassical theory, one might wonder that any conflict with maximization possibly deriving from the FDC would

be prevented by what the theory regards as the correct functioning of the price system. In consideration of the character of the two parties involved, reasonably the adjustment to a price-quantity configuration satisfying the FDC would lie on the behaviour of the private sector alone, for Government's actions of fiscal policy are hardly reducible to price functions, and at any rate there can be no guarantee that the responsiveness to prices would be one which helps the FDC to be fulfilled consistently with private maximization.

A paradoxical result which such asymmetry could produce becomes evident in the context of what is the 'purest' form of modern neoclassical theory, the general intertemporal equilibrium with complete markets. As is well known, in that analytical setting agents take at time 0 their maximizing decisions to buy and sell for the entire sequence of future dates. If, for the reasons previously offered, at whatever set of prices the attainment of the FDC would bring about no increase, or even a decrease, of private utility, maintaining the FDC among the maximizing conditions could entail an optimum decision plan of the private sector which includes *no* subscription of public bonds at any date. In this case the nil present value of future public debt would be accomplished by preventing even the coming into existence of public debt. In the face of a problem of inconsistency between the private maximizing role of the FDC on the one hand, and the public budget choices of the Government on the other, the solution would thereby be that of cancelling the very source of the problem – as well as, in fact, any *raison d'être* for the FDC itself[22].

---

[22] Taking for granted assumptions and analytical structure of the general intertemporal equilibrium, the conditions we are referring to would imply zero values for B_0, B_1, B_2 and so on, that is to say that since an indefinitely remote time anterior to the arbitrary initial date 0 fixed by the analysis, the intertemporal plan of the representative agent did not contemplate the subscription of public debt bonds at any future date.

III. Further on the Future Debt Condition

10. If not unquestionable within the dominant theoretical framework, the logic of the FDC would be totally absent in a different analytical context. More in particular, that condition is inconsistent with the case that the levels of activity of the economic system are limited by the level of aggregate demand. As we are going to argue, even taking for granted the validity of the household's intertemporal budget constraint, the admitted influence of aggregate demand upon the levels of income of the private sector would on its own eliminate the trade-off between levels of consumption and future stocks of assets, including public bonds, which is the necessary premise of the FDC.

Referring, for the sake of intuitiveness, to a finite time horizon, let us start by expressing the private intertemporal budget constraint in terms of values taken as net of capital depreciation and capitalized at the final date H, assuming that the rate of return on capital does not differ from the rate of interest on public debt:

$$K_0(1 + i)^H + B_0(1 + i)^H + \sum_1^H (1 + i)^{H-t} W_t =$$

$$= \sum_1^H (1 + i)^{H-t} C_t + \sum_1^H (1 + i)^{H-t} T_t + K_H + B_H \qquad (3.15)$$

Using into the latter equality the budget constraints which hold in each single period, the generic form of which is:

$$K_{t-1}(1 + i) + B_{t-1}(1 + i) + W_t = C_t + T_t + K_t + B_t \qquad (3.16)$$

we obtain:

$$K_0 + B_0 + i \sum_1^H K_{t-1} + i \sum_1^H B_{t-1} + W_t =$$

$$= \sum_{1}^{H} C_t + \sum_{1}^{H} T_t + K_H + B_H \,{}^{23} \tag{3.17}$$

Equality (3.17) shows the parity between: at the l. h. s., the sum of the *current* values of the initial wealth (capital assets plus public bonds) of the private sector and of the total income flows the latter earns from capital, labour and as interest on public debt; at the r. h. s., the *current* values of consumption flows and tax payments of the private sector over the whole time interval, summed to the total wealth the sector owns at the final date.

---

[23] Considering, for instance, a two-period time span, equality (3.15) would be:

$$K_0(1+i)^2 + B_0(1+i)^2 + \sum_{1}^{2} (1+i)^{2-t} W_t =$$

$$\sum_{1}^{2} (1+i)^{2-t} C_t + \sum_{1}^{2} (1+i)^{2-t} T_t + K_2 + B_2$$

which, isolating the final total wealth at the r. h. s. and ordering in a different way the terms at the l. h. s., becomes:

$(1+i)[K_0 + iK_0 + B_0 + iB_0 + (W_1 - C_1 - T_1)] + (W_2 - C_2 - T_2) = K_2 + B_2$
and then:

$K_0(1+i) + B_0(1+i) + i[K_0(1+i) + B_0(1+i) + ((W_1 - C_1 - T_1)] +$
$(W_1 - C_1 - T_1) + (W_2 - C_2 - T_2) = K_2 + B_2 \tag{3.15'}$

On the other hand, in the single periods the following budget constraints must separately hold:

$$K_0(1+i) + B_0(1+i) + (W_1 - C_1 - T_1) = K_1 + B_1$$
$$K_1(1+i) + B_1(1+i) + (W_2 - C_2 - T_2) = K_2 + B_2$$

Using the first of the latter two equalities into (3.15') we can write:

$K_0 + B_0 + (iK_0 + W_1 + iB_0) + (iK_1 + W_2 + iB_1) - C_1 - C_2 - T_1 - T_2 = K_2 + B_2$

Finally, positing $Y_t = iK_{t-1} + W_t$ for total income produced in period t and $R_t = iB_{t-1}$ for interest on public debt also paid in period t:

$$K_0 + B_0 + \sum_{1}^{2} Y_t + \sum_{1}^{2} R_t - \sum_{1}^{2} C_t - \sum_{1}^{2} T_t = K_2 + B_2$$

which is the equality of the kind of (3.18) in the text.

Further, substituting total produced income (output), which we call Y_t, for the sum of capital and labour incomes earned in period t by the private sector and posit R_t as the interest on public debt yielded in period t, we get:

$$K_0 + B_0 + \sum_1^H Y_t + \sum_1^H R_t = \sum_1^H C_t + \sum_1^H T_t + K_H + B_H \qquad (3.18)$$

What has been done so far are just transformations of the intertemporal budget constraint of the private sector. Let us now impose the condition that the output produced in each period is determined by the level of aggregate demand, in particular through a direct relation between output and public deficit. Starting from the equality between total output and aggregate expenditure:

$$Y_t = C_t + I_t + G_t \qquad (3.19)$$

and assuming:

$$C_t = c(Y_t + R_t - T_t) \qquad (3.20)$$

with c representing the given share of consumption expenditure out of private disposable income, it follows that:

$$Y_t = cY_t + c(R_t - T_t) + I_t + G_t \qquad (3.21)$$

Defining public deficit as:
$$D_t = G_t - T_t + R_t \qquad (3.22)$$

where for the sake of simplicity taxes are supposed to be fixed independently of income, we obtain:

$$Y_t = cY_t + c(R_t - T_t) + I_t + D_t - (R_t - T_t) \qquad (3.23)$$

and hence:

$$Y_t = \frac{1}{1-c} I_t + (T_t - R_t) + \frac{1}{1-c} D_t \qquad (3.24)$$

We can then replace the public deficit by the increment in the stock of public debt (abstracting from other forms of deficit financing):

$$Y_t = \frac{1}{1-c} I_t + (T_t - R_t) + \frac{1}{1-c} (B_t - B_{t-1}) \qquad (3.25)$$

and substituting for Y_t in (3.18):

$$K_0 + B_0 + \sum_1^H \left[\frac{1}{1-c} I_t + (T_t - R_t) \right] + \frac{1}{1-c} \sum_1^H (B_t - B_{t-1})$$

$$= \sum_1^H C_t + \sum_1^H (T_t - R_t) + K_H + B_H \qquad (3.26)$$

In the last sum at the l. h. s. each stock of public debt at a date other than 0 or H appears twice, respectively with opposite sign, and therefore can be cancelled out. In turn, the sum of the differences between tax and interest payments, which appear both at left and right hands of the equality and can also be dropped, so that the latter equality reduces to:

$$K_0 + B_0 + \frac{1}{1-c} \sum_1^H I_t + \frac{1}{1-c} (B_H - B_0) = \sum_1^H C_t + K_H + B_H \quad (3.27)$$

The sum of the flows of investment (remember we abstract from capital depreciation) amount to the difference between final and initial

capital stocks:

$$\sum_{1}^{H} I_t = K_H - K_0 \qquad (3.28)$$

Substituting in (3.27) for the latter difference and isolating the sum of consumption flows, we get:

$$\frac{c}{1-c}[(K_H - K_0) + (B_H - B_0)] = \sum_{1}^{H} C_t \qquad (3.29)$$

We have thereby arrived at a formulation of the intertemporal budget constraint showing that, given the initial stocks of capital and public debt, the consumption flows of the private sector over the sequence of periods depend *positively* on the magnitudes of the public debt and capital stock at the final date. The underlying logic of this direct relation is clearly that of the influence of demand on the levels of income, hence consumption, of the agents. For a given initial stock of public debt, the size of the final stock will be larger, the higher the flows of public deficits over the sequence of periods; and, given the initial stock of capital, the final stock will in turn be larger, the higher the flows of investment realized during the time interval. Higher levels of public deficits and private investment generate higher levels of aggregate demand and income, and therefore higher levels of private consumption.

The result just reached leads us to understand that even over a finite time interval, the 'representative agent', i.e. the private sector, would have no interest in a reduction, let alone the cancellation, of the final stock of public debt, and the same for the final stock of capital. As we have pointed out, the flows of consumption available to the private sector would on the contrary increase with the growth of the two kinds of private wealth. The idea is therefore reversed that the owning of net positive assets at the final date would imply the gratuitous sacrifice of potential consumption by the private sector: and the reason for that

ultimately lies in that, with a demand-side determination of outputs, consumption is not the alternative, but is rather the 'joint product' of the accumulation of wealth by the community.

11. While on the one hand the FDC is taken to represent a theoretically founded condition for the sustainability of public debt albeit devoid, as has been earlier emphasized, of definite applicative implications, on the other hand the actual proposals and debates of economic policy often refer to the sustainability of public debt as the mere stabilization of its ratio to domestic output GDP (whatever the value of the ratio), a quite pragmatic notion generally put forward without the assistance of any theoretical argument and just in contrast with an otherwise unsustainable tendency of the debt ratio to 'explode'. Some authors have however attempted at connecting the two notions, arguing that the stability of the debt ratio would allow the FDC to be fulfilled, hence would have precisely in the latter its theoretical basis[24].

Indeed, a stable ratio of public debt to GDP entails that the FDC is satisfied if the rate of growth of output g keeps lower than the (real) interest rate r, since in that case the debt would in turn grow at the rate g and the limit of the present value of its future magnitude would therefore be nil:

$$\lim_{H \to +\infty} B_0 \left(\frac{1+g}{1+r} \right)^H = 0 \qquad (3.30)$$

The same would obviously not hold if the rate of growth exceeded the interest rate, for with a stable ratio to GDP the debt would grow at the rate $g > r$, and the present value of future debt would tend to infinity. The confinement of the argument to the $g < r$ case is sometimes justified either on the theoretical grounds that the opposite case implies conditions of 'dynamic inefficiency' (i.e. of failing maximization of per capita consumption), from which agents would successfully depart by lowering their saving rate; or merely by appealing to what can be observed in certain historical periods, although

---

[24] Blanchard *et al.* (1990), pp. 11-2; Gros (2010), p. 17; Ley (2010), pp. 5-6.

admittedly not in others.

Even independently from the conditions under which it would hold, the theoretical support the stabilization of the debt ratio receives by the FDC is at any rate undermined by our criticism of the latter as part of the rational behaviour of the private sector. The attempt thereby fails to supply even with that analytical basis the so commonly adopted policy target of a stable ratio of public debt to GDP.

References

Bagnai, A. (2004). Keynesian and Neoclassical Fiscal Sustainability Indicators, with Applications to EMU Member Countries, *Dipartimento di Economia Pubblica Working Paper*, 75, Università degli Studi di Roma 'La Sapienza'.

Barro, R.J. (1989). The Neoclassical Approach to Budget Deficits. In Barro, R.J. (eds.), *Modern Business Cycle Theory*. Oxford: Basil Blackwell, 178-235; repr. in Kaounides and Wood (1992), Vol. III, 418-30.

Barro, R.J. and Sala-i-Martin, X. (2004). *Economic Growth*. Cambridge, MA and London: The MIT Press.

Becker, R. (2008). Transversality condition. In Durlauf, S.N. and Blume, L.E. (eds.), *The New Palgrave: a dictionary of economics*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan, 382-4.

Bellino, E. (2013). On the Stability of the Ramsey Accumulation Path. In Levrero, E.S., Palumbo, A. and Stirati, A. (eds.), *Sraffa and the Reconstruction of Economic Theory*, Vol. 1. London: Palgrave Macmillan, 70-14.

Blanchard, O.J. and Fisher, S. (1989). *Lectures on Macroeconomics*. Cambridge, MA and London: The MIT Press.

Blanchard, O.J., Chouraqui, J.C., Hageman, R.P. and Sartor, N. (1990). The Sustainability of Fiscal Policy: New Answers to an Old Question, *OECD Economic Studies*, 15.

Chalk, N. and Hemming, R. (2000). Assessing Fiscal Sustainability in Theory and Practice, *International Monetary Fund, IMF Working Paper*, 00/81.

Ciccone, R. (2002). *Debito pubblico, domanda aggregata e accumulazione*, Roma: Aracne.

Ciccone, R. (2013). Public Debt and Aggregate Demand: Some Unconventional Analytics. In Levrero, E.S., Palumbo, A. and Stirati, A. (eds.), *Sraffa and the Reconstruction of Economic Theory*, Vol. 2. Houndmills, Basingstoke, U.K.: Palgrave MacMillan, 15-43.

Garegnani, P. (1978). Notes on Consumption, Investment and Effective Demand: I, *Cambridge Journal of Economics*, 2(4), 335-53.

Garegnani, P. (1979), Notes on Consumption, Investment and Effective Demand: II, *Cambridge Journal of Economics*, 3(1), 63-82.

Gros, D. (2010). Adjustment Difficulties in the GIPSY Club, *CEPS Working Document*, 326.

Kamihigashi, T. (2008). Transversality conditions and dynamic economic behaviour. In Durlauf, S.N. and Blume, L.E. (eds.), *The New Palgrave: a dictionary of economics*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan, 384-7.

Kaounides, L.C. and Wood, G.E. (1992). *Debt and Deficits*, Vol. I, II and III. Aldershot (GB) and Brookfield (Vermont, USA): Edward Elgar.

Lerner, A.P. (1943). Functional Finance and the Federal Debt, *Social Research*, Vol. 10, 38-51; repr. in Kaounides e Wood (1992), Vol. II, 97-110.

Ley, E. (2010). Fiscal (and External) Sustainability, *MPRA Paper*, 23956, University Library of Munich.

Romer, D. (2012). *Advanced Macroeconomics*. New York: McGraw-Hill.

Spaventa, L. (1988). Introduction: is there a public debt problem in Italy? In Giavazzi, F. and Spaventa, L. (eds.), *High Public Debt: The Italian Experience*. Cambridge: Cambridge University Press, 1-24.

Sraffa, P. (1960). *Production of Commodities by Means of Commodities*. Cambridge: Cambridge University Press.

Wyplosz, C. (2011). Debt Sustainability Assessment: Mission Impossible, *Review of Economics and Institutions*, 2(3), 1-37, retrieved from http://www.rei.unipg.it/article/view/42.

Chapter 4

Incomes from capital in alternative economic theories

Saverio M. Fratini[*]

4.1 Introduction

Empirical studies claim that, in the most industrialized countries, the share of capital in the national incomes is in the interval 25-30% (see Piketty 2014, p. 222, figure 6.5)[1]. The nature of these incomes is quite evident in practice: they are equity dividends, capital gains, interest on loans, and managerial incentives. Much more complex is their analysis from the point of view of economic theory. In particular, as is clear, this analysis is closely linked to the notion of capital adopted.

As is well known, as far as income distribution is concerned, there are two alternative views in economic theory (see, in particular, Garegnani 1984, pp. 291-292). The first is the classical/Marxian approach, according to which income distribution is a social phenomenon: it depends on the conflict among social classes. The second is the neo-classical theory, which understands income distribution as a market phenomenon regulated by the equilibrium between supply and demand[2].

[1] According to Rognlie (2016), the decadal average of the net capital share of private domestic value added is above 30% in Canada and Italy (cf. Rognlie 2016, p. 16, table 2).

[2] As is well-known, Veblen introduced the term 'neo-classical' with reference to Marshall's theory. With the aim of avoiding possible confusions, we can stress that here we consider as 'neo-classical' a certain representation of the working of the economic system. According to the neo-classical view, there are two groups of agents in the

Here we shall try to address the theoretical explanation of incomes from capital from both standpoints. We shall start, in section 4.2, with the classical/Marxian approach, in which – as we shall see – capital is understood as the amount of value that allows producers to pay costs in advance, before revenues are obtained, and profit is a residual – a surplus value. As for the neo-classical approach, we shall consider two different versions. In section 4.3, we shall deal with the traditional marginalist theory, in which capital is seen as a 'factor of production' and the rate of interest as the price firms pay for its use. Then, in section 4.4, we shall discuss the Arrow-Debreu model, in which capital is neither a 'factor of production', nor an amount of value that allows firms to pay costs in advance. Some conclusions are drawn in section 4.5.

4.2. Capital and profit in the classical/Marxian approach

In order to introduce the conception of capital within the classical/Marxian approach, we need to start with a quick reconstruction of the basic features of the capitalist mode of production. In particular, we can list the following important characteristics of this social organization:

---

economy: households and firms. Households want to maximize their utility; firms their profit. Supply and demand functions arise from these maximizations. Prices and produced quantities are set so as to balance supply and demand on each market. For sure, both the traditional marginalist theory and the Arrow-Debreu theory refer to this view and, accordingly, they can be considered as two versions of the neo-classical approach.

The neo-classical description of the working of the economic system is decidedly different from the one provided by the classical economists and Marx. In the latter, the key element is the organization of the society in classes with opposed interests. Economic activities mainly concern the conflict between different social classes and the competition among the members of each class. Although there is a 'core' of regularities whose study can be handled with purely logical (or even mathematical) tools, the working of an economic system cannot be understood without reference to its social and institutional characteristics.

i)   Products are commodities. Production is intended for the market and not (directly) for consumption.

ii)   Labour is performed by wage-earning workers. Workers are forced to sell their labour-power due to their separation from the means of production. Labour-power is the commodity workers sell in exchange for wages.

iii)   Natural resources (lands, mines, oilfields, etc.) are of private property. Every natural resource has an owner, and the class of landowners is distinct from the class of capitalists[3].

iv)   Production processes are organized (directly or indirectly) by capitalists. They buy the inputs and sell the outputs.

The inputs of the production processes are: commodities (raw materials, tools, machines, etc.); labour-power performed by workers with different skills; and the productive uses of natural resources. Points i-iii above imply that these inputs must be purchased on the market. Assuming, for the sake of simplicity, that they must be paid *ex-ante*, at the beginning of the production process, capital is necessary in order to finance the expenses of production, and this is the reason why capitalists are the only possible organizers of the production processes (point iv).

In fact, since production takes time, in every single process, the employment of inputs must precede the production of outputs. Accordingly, inputs are generally purchased and employed before outputs are sold. Therefore, costs and revenues of the same process are not simultaneous, as the former generally precede the latter. As a result, the costs of a certain process cannot be financed by the revenues of the

---

[3] In the classical political economy, the distinction between capital and natural resources, as well as between capitalists and landowners, plays an important role. In Marx's analysis, in particular, landed property is seen as a 'monopoly' of the class of landowners (Marx, 1909, pp. 723 and 732-733). This monopoly, in Marx's view, is at the origin of ground-rent, which, accordingly, is not an income from capital. Since land may be sold like any other article of commerce, to the buyer, the rent may appear merely as an income due to the capital invested paying the price of land. However, as Marx remarks, this way of reasoning means justifying the existence of rent with its existence, since the price of land is nothing else than the present value of future rents (Marx 1909, p. 901).

same process. Hence, capital is the amount of purchasing power that is required, for each process, to finance the costs.

The amount of capital invested is then recovered out of revenues when the outputs are sold. Moreover, revenues leave, normally, a surplus over and above costs. This surplus value is profit[4]. The profit per unit of capital invested (for a single process) is the rate of profit.

4.2.1 An example

Let us assume there are N commodities, A different kinds of labour services, and B different sorts of natural resources. The production process of a generic commodity n, with $n = 1, 2, \ldots, N$, which starts in a certain period t, employs a vector of commodities $\mathbf{X}_t^n \in \mathbb{R}_+^N$, a vector of labour services $\mathbf{L}_t^n \in \mathbb{R}_+^A$, and a vector of productive uses of natural resources $\mathbf{\Lambda}_t^n \in \mathbb{R}_+^B$ in order to get an output C_{t+1}^n in the period $t + 1$.

Referring to the economy as a whole, the employment of inputs in period t is $\mathbf{X}_t = \sum_n \mathbf{X}_t^n$, $\mathbf{L}_t = \sum_n \mathbf{L}_t^n$ and $\mathbf{\Lambda}_t = \sum_n \mathbf{\Lambda}_t^n$ and the output of period $t + 1$ is a vector $\mathbf{C}_{t+1} = [C_{t+1}^1, C_{t+1}^2, \ldots, C_{t+1}^N]$. Let $\mathbf{p}_t \in R_+^N$, $\mathbf{w}_t \in \mathbb{R}_+^A$ and $\mathbf{\rho}_t \in \mathbb{R}_+^B$ be the (row) vectors of commodity prices, wage rates, and rent rates in period t, if wages and rents are paid *ex-ante* – namely, at the beginning of the process, then the total investment of capital in period t is:

$$K_t = \mathbf{p}_t \cdot \mathbf{X}_t + \mathbf{w}_t \cdot \mathbf{L}_t + \mathbf{\rho}_t \cdot \mathbf{\Lambda}_t \qquad (4.1)$$

The capital advanced is then recovered by revenues in period t +1. However, under normal conditions, $\mathbf{p}_{t+1} \cdot \mathbf{C}_{t+1} > K_t$. This means that the revenues leave a profit over and above the costs of production:

$$\Pi_{t+1} = \mathbf{p}_{t+1} \cdot \mathbf{C}_{t+1} - K_t \qquad (4.2)$$

---

[4] As is known, referring to the economic system as a whole, the existence of this profit is grounded on the fact that the economy is technically able to produce more than the amount of commodities that enables the repetition of the production processes, namely commodities employed as means of production and subsistence for the workers. See: Garegnani (1984, pp. 292-294).

The rate of profit is the amount of profit obtained for each unit of capital invested. Accordingly:

$$\pi_{t,t+1} = \frac{\Pi_{t+1}}{K_t} \tag{4.3}$$

From this example, it should be clear that:

i) The amount of capital K_t is not the quantity of an input. The quantities of inputs employed are $\mathbf{X}_t$, $\mathbf{L}_t$ and $\mathbf{\Lambda}_t$.

ii) The rate of profit is not the price of capital. The amount of profit Π_{t+1} is a residuum and not the result of a price-times-quantity multiplication.

4.2.2 The ordinary rate of profit

Assuming free competition among producers, this approach focusses the attention on a theoretical (persistent) position characterized by the uniformity of the rate of profit on the capital invested in different sectors of economic activity. Following Smith, we can call this uniform rate the 'ordinary rate of profit' and the relative prices associated with it the 'natural prices'[5].

Thanks to Sraffa's theory (1960), we know that commodities' natural prices and the ordinary rate of profit corresponding to a given wage rate must be determined simultaneously as the solution of a system of equations. In particular, starting from the same example considered in the previous sub-section, for the sake of simplicity, we can set aside natural resources and assume that there is just one kind of labour service[6]. In this case, if capital must get the same rate of profit π

---

[5] In Marx's analysis, the same concepts correspond to the 'general rate of profit' and the 'prices of production'.

[6] As Garegnani argues, within the classical approach, the possibility of referring to homogeneous labour and a single wage rate rests on the hypothesis of a given structure of wage rates for workers with different skills (for details, see Garegnani 1984, p. 293, footnote 5).

independently of the sector in which it is invested, then the following conditions must hold[7]:

$$p^n C_{t+1}^n = (\mathbf{p} \cdot \mathbf{X}_t^n + w L_t^n)(1 + \pi) \quad \forall\, n = 1, 2, \dots, N \qquad (4.4)$$

Adopting the national net output as numéraire commodity – as Sraffa does – then:

$$\mathbf{p} \cdot (\mathbf{C}_{t+1} - \mathbf{X}_t) = 1 \qquad (4.5)$$

For a given level of the wage rate w, equations (4.4) and (4.5) are able to determine the vector of natural prices $\mathbf{p}$ and the ordinary rate of profit π.

Finally, equations (4.4) allow us to define the price vector as a function of the distribution variables: $\mathbf{p} = \mathbf{p}(w, \pi)$. Substituting this function within equation (4.5), we get:

$$\mathbf{p}(w, \pi) \cdot (\mathbf{C}_{t+1} - \mathbf{X}_t) = 1 \qquad (4.6)$$

Equation (4.6) is what Garegnani (1984) calls the 'surplus equation'. It is particularly useful for a comparative statics exercise. It shows how the rate of profit that solves the system of equations (4.4)-(4.5) varies as the wage rate level changes. In particular, equations (4.6) tells us that there is an inverse relation between w and π. Therefore – under normal conditions and *ceteris paribus* – incomes from capital tend to be high when the wage rate level is low. This fact provides the theoretical ground for the conflict between social classes that can be observed in the real world.

---

[7] Since we are here referring to a persistent position of the economy, the time index of the quantities is only used to stress that the inputs are used before obtaining the outputs.

4.3. Capital as a factor of production

Within the classical/Marxian approach, the organization of the society in three classes – workers, landowners, and capitalists – entails the division of the national income into three parts: wages, rents, and profit. Within the marginalist theory, these three different incomes are understood as what firms pay to households for the employment of three 'factors of production': labour, land and capital[8].

Once this standpoint is adopted, the existence of the social classes becomes inessential. The working of the system depends on the decisions taken by two different sorts of economic agents: households and firms.

Firms organize the production processes. They demand factors of production and supply commodities. Factor prices – wage rate, rent rate, and interest rate – are included into the price vectors. Given a price vector, each firm decides its production plan in order to maximize its (net) profit, namely, the difference between revenues and the costs for the employment of the production factors.

Households supply production factors – which are available in given quantities[9] – and demand the commodities produced[10]. Given a price vector, each household decides its consumption plan in order to maximize its utility subject to the budget constraint.

---

[8] Actually, the idea of distribution variables as the prices paid for the use of the factors of production preceded the development of the marginalist theory. The embryonic form of this idea was already present in the 'vulgar economists' towards whom Marx addresses numerous criticisms. In particular, according to Marx, these factors of production are the result of the transformation of social classes into things. See Marx (1909, p. 966).

[9] As is clear, although households are endowed with given quantities of factors of production, this does not mean that they are inelastically supplied. It may well be assumed that factor services are also desired by households for consumption. In this case, the quantity supplied to firms may change as the price system varies.

[10] In most versions of the theory, households demand consumption goods and services. Nonetheless, there are also models in which households demand capital goods too. Specifically, in Walras's theory, households demand capital goods with their saving in order to sell their productive services to firms. See Walras ([1926] 1977), p. 267.

In this framework, workers, landowners, and capitalists – provided that they can be distinguished by some special assumption – are on the same side: that of households. They are providers of factors of production. Then, wages, rents, and interest paid by firms to households have exactly the same nature. They are incomes for households and costs for firms.

4.3.1 The equilibrium rate of interest

Once the distribution variables are understood as factor prices, they must be determined at the same time and by the same mechanism as all the other prices, namely, by a system of general equilibrium conditions.

Focussing on the capital market, the rate of interest is thought to fall whenever the demand for capital by firms is smaller than the quantity of this factor made available by households, and to rise in the opposite case. The variation of the rate of interest with respect to the other factor prices should entail a change in the methods of production in use. In particular, a fall in the rate of interest should bring about the adoption of more capital-intensive methods so that, *ceteris paribus*, the demand for capital increases, leading to a reduction of the initial excess supply.

The fundamental idea behind this mechanism of substitutability between factors – or better, between their productive services – is well known. The first-order conditions for firms' (net) profit maximization impose the equality between factor prices (expressed in terms of produced commodity) and their respective marginal productivity. Hence, the principle of diminishing marginal productivity implies that, if the rate of interest falls and the other factor prices are unchanged, the adoption of a more capital-intensive production plan is needed in order to restore the equality between the marginal product of capital and the rate of interest.

Therefore, the equilibrium rate of interest is typically conceived as the rate that makes demand for capital by firms equal to the supply provided by households – which in turn depends on their present and past saving decisions[11]. In Marshall's words:

---

[11] As for the supply of capital in the marginalist framework, we refer the reader to Fratini (2019b).

interest, being the price paid for the use of capital in any market, tends towards an equilibrium level such that the aggregate demand for capital in that market, at that rate of interest, is equal to the aggregate stock forthcoming there at that rate (Marshall 1920, p. 534).

4.3.2 Interest and profit

In the marginalist approach, incomes from capital are what firms pay to households for the use of a factor of production. They are understood as the result of a price-times-quantity multiplication and enter into firms' costs. Hence, incomes from capital lose their residual nature. They are not the difference between revenues and costs.

Let us denote by L^n, Λ^n K^n the quantities of the factors labour, land, and capital employed in the production of a certain commodity n, with n = 1, 2, … N. Given the technical conditions of production, the amount of commodity n obtained is expressed as a certain (differentiable) function of the quantities of the factors of production used:

$$C^n = F^n(L^n, \Lambda^n, K^n) \tag{4.7}$$

Accordingly, let w, ρ, and i be the wage rate, the rent rate, and the interest rate, respectively, the total costs of production of a final output C^n are: $wL^n + \rho\Lambda^n + (1+i)K^n$.

For a given price p^n of commodity n, the amount of profit earned from the production of a quantity C^n is:

$$\Pi^n = p^n C^n - [wL^n + \rho\Lambda^n + (1+i)K^n] \tag{4.8}$$

Within this theory, the amount of profit Π^n is not part of the incomes from capital. Here, the net income earned because of the employment of capital is the amount of interest iK^n that is included in the costs. The profit Π^n is understood as income earned by firms. In fact, as noted at the beginning of this section, in the marginalist theory, firms decide on their production plans in order to maximize their profit, namely, the difference between revenues and the costs of the employment of the production factors.

Once it is clear that they are not understood as incomes from capital, the nature of firm profits is not very clear. It is not clear what their counterpart in the real world can be. However, this is not a serious problem, because these profits disappear with their maximization.

In fact, substituting equation (4.7) into equation (4.8), the first-order conditions for the maximization of the amount of profit Π^n are:

$$\begin{cases} p^n F_L^n(L^n, \Lambda^n, K^n) - w = 0 \\ p^n F_\Lambda^n(L^n, \Lambda^n, K^n) - \rho = 0 \\ p^n F_K^n(L^n, \Lambda^n, K^n) - (1 + i) = 0 \end{cases} \qquad (4.9)$$

where $F_J^n(\cdot)$ is the partial derivative of $F^n(\cdot)$ with respect to factor J, with $J = L, \Lambda, K$. Therefore, if the production function $F^n(\cdot)$ exhibits constant returns to scale – i.e. it is homogeneous of degree one – and firms are using the factors in optimal quantities, then $\Pi^n = 0$[12,13].

4.3.3 Capital and the value of capital goods
It should be clear that the idea of the interest rate as a price represents the other side of the coin with respect to the idea of capital as a factor of production – that is, as something that firms materially employ in production, together with labour and land. Capital must, in fact, be

---

[12] According to Euler's formula for homogeneous functions, if the production function $F^n(\cdot)$ is homogeneous of degree one, then:
$C^n = F_l^n(L^n, \Lambda^n, K^n) \cdot L^n + F_\Lambda^n(L^n, \Lambda^n, K^n) \cdot \Lambda^n + F_K^n(L^n, \Lambda^n, K^n) \cdot K^n$. Therefore, equation (4.8) can be written as follows: $\Pi^n = [p^n F_l^n(L^n, \Lambda^n, K^n) - w]L^n + [p^n F_\Lambda^n(L^n, \Lambda^n, K^n) - \rho]\Lambda^n + [p^n F_K^n(L^n, \Lambda^n, K^n) - (1 + i)]K^n$. It is now clear that the first-order conditions (4.9) entail $\Pi^n = 0$.

[13] Wicksteed – the 'purist' of the marginalist theory – writes that the existence of a profit or surplus above the costs is nothing more than a fantasy. He maintains that anyone who understands the facts cannot claim that, once the remuneration of factors corresponds to their marginal product, there remains any residuum to be divided or appropriated. He concludes – clearly as a criticism of the classical/Marxian approach – that 'the vague and fervid visions of this unappropriated reserve, ruling upward as we recede from the marginal distribution, must be banished for ever to the limbo of ghostly fancies' (Wicksteed, 1910, p. 573).

substitutable for other factors and therefore must play the same role and must satisfy the same need.

The conception of capital as a factor of production represents, no doubt, the biggest difficulty met by the marginalist theory of distribution. The idea of capital as something that is materially used in the production process is associated with its identification with the set of capital goods (tools, raw material, machines, etc.) employed. Hence, the theory tries to present capital as an amalgam of capital goods[14]. That is the reason why the problems encountered by this theory with reference to capital as a factor of production are often erroneously believed to be problems of 'aggregation' – i.e. the transformation of a vector of quantities of heterogeneous commodities into a single magnitude. Actually, no real problem exists for the aggregation of the capital goods into the corresponding amount of value by means of their prices. The problems arise if we want to regard this amount of value as the quantity of a factor of production and the rate of interest as the price of its use.

At the cost of a drastic simplification, we can say that the basic idea is that of the existence of a special substance, a sort of 'jelly', of which all capital goods are made and from which their productivity derives. If this jelly existed, the form of the various types of capital goods that it actually takes would be a secondary aspect. In fact, first, to every possible set of heterogeneous capital goods, one could associate the corresponding quantity of jelly. Secondly, sets of capital goods containing a greater quantity of jelly would make it possible to obtain a larger product, all other things being equal.

Despite several attempts, the last of which was Samuelson's (1962), a substance or jelly with these extraordinary properties has never been found. Thus, in its absence, the employment of capital was

---

[14] The double presence of capital and capital goods generates a sort of schizophrenia in the marginalist theory. Taking Solow as the personification of this theory, Samuelson wrote that '[o]ne might almost say that there are two Solows': (a) 'the orthodox priest of the MIT school', who regards capital as 'a great variety of heterogeneous physical capital goods'; and (b) 'the busman on a holiday who operates brilliantly and without inhibitions in the rough-and-ready realm of empirical heuristics' and makes use of a 'Clark-like concept of aggregate "capital"' (Samuelson, 1962, p. 193).

generally identified with the value of capital goods used in production. It was therefore tried to create a hybrid between the (right) idea of capital as a value fund that advances costs and the (wrong) idea of capital as an amalgam of capital goods. However, as was already pointed out by Wicksell (1934, p. 149), the value of capital goods employed cannot be considered as the quantity of a factor of production[15]. Doing that can in fact lead to paradoxical results.

First, as Samuelson (1966, p. 582) wrote, there is, in general, no unambiguous way to say that a process of production is more capital-intensive than another – namely, that it entails a greater employment of capital per unit of labour. Since the prices of capital goods change as the interest rate varies, if we take two methods of production of the same commodity, it is possible that the method that is more capital-intensive for a certain level of interest rate becomes the method that is less capital-intensive for a different level. The ranking of methods on the basis of capital intensity cannot be done independently of the level of interest rate.

Second, given a certain interest rate, the method that employs more capital per unit of labour is not necessarily the one that enables the

---

[15] In the marginalist theory, the amount of output obtained depends on the productive services provided by the factors of production in use. Therefore, the quantity of each factor must be representative of the quantity of productive services that that factor is able to provide. It must be expressed in technical units. To give an example, we could try to measure the employment of labour in meters (adding up workers' heights) instead of hours (i.e. adding up the hours worked by the labourers). In this case, since taller workers do not generally provide more productive services than shorter workers, an increase in the amount of work would not lead with certainty to the production of a greater output. The relationship between the quantity of labour employed, measured in meters, and the quantity produced would become ambiguous, so that we could no longer represent it by a mathematical function of general validity. Hence, labour must be measured in technical units, that is, in hours of work.

However, as Wicksell stresses, value is not a technical unit of measure of capital goods since '[t]he productive contribution of a piece of technical capital, such as a steam engine, is determined not by its cost but by the horse-power which it develops' (Wicksell, 1934, p. 149). Therefore, although there is no problem in aggregating capital goods in terms of value, this amount of value cannot be understood as representative of the technical contribution of the set of capital goods. It cannot be understood as the quantity of a factor of production.

highest output per unit of labour to be obtained. This point is essentially an implication of the previous one. Suppose that, for a certain level of the rate of interest, method α is more capital-intensive than method β and also provides a greater output per unit of labour. If, for a different level of the interest rate, method β has become the most capital-intensive, then, for the same amount of labour employed, the method that employs the greatest capital is the one that results in the lowest output.

Third, when an increase in the interest rate entails a change in the production method in use – i.e. the one that minimizes the unit cost of production – the method that comes into use does not necessarily employ less capital per unit of labour than the method previously used. In other words, an increase in the interest rate may not make (relatively) more expensive those methods that require more capital. This result clearly casts into doubt the idea that the interest rate can be seen as the price that firms pay for the use of the factor 'capital'.

The point is that, since there is no convincing way to conceive the productive factor 'capital', the rate of interest cannot be thought of as its price. This emerged during the Cambridge capital theory debate[16] and was, in fact, explicitly stated by Bliss in his authoritative book on capital theory:

The value which accrues from a sale is the product of price and quantity sold. Hence if the rate of interest is the price of capital, the quantity of capital must be the wealth on which an interest yield is calculated. It will be shown shortly why this view is incorrect, but to cut a long story short, the conclusion may be announced at once. The rate of interest is not the price of capital (Bliss 1975, pp. 6-7).

4.4 Interest and profit in the Arrow-Debreu theory

As shown in section 4.2, in the classical/Marxian approach, incomes from capital are a profit, namely, the difference between revenues and the costs of production – which were totally or partially advanced by capital. By contrast, in the marginalist theory, they are the interest firms pay on the

---

[16] A number of surveys of the Cambridge capital theory debate are available. In particular, the readers may refer to Harcourt (1969 and 1972).

capital borrowed from households, and hence they are included into the costs. In this framework, the difference between revenues and costs represents the profit of firms and vanishes in equilibrium.

We shall now consider the Arrow-Debreu theory of value, in which prices are determined by a market-clearing equilibrium, as in the marginalist approach. However, in contrast to the latter, production processes are not assumed to employ 'factors of production', but instead, are assumed to employ Arrow-Debreu commodities, namely, goods and services with a specific place and date of delivery.

Actually, the Arrow-Debreu model is a very peculiar one. On the one hand, no form of capital seems present in this framework: neither the classical (an amount of value that allows the payment of costs at the beginning of the process), nor the marginalist (a factor of production). On the other hand, profit and interest – at least in name – are still there, but they cannot correspond to the incomes from capital discussed in sections 4.2 and 4.3. Therefore, as will be clear at the end of this section, Arrow-Debreu theory provides no support for the marginalist explanation of incomes from capital[17].

4.4.1 Own-rates of interest

In the Arrow-Debreu framework, as previously mentioned, production processes do not employ factors of production. Accordingly, the rate of interest is not understood as the price firms pay for the use of capital.

In fact, in the Arrow-Debreu models, instead of one rate of interest there are many 'own-rates of interest'[18]. Let p_t^n and p_{t+1}^n be the

---

[17] Some scholars seem to believe that Arrow-Debreu theory is the rigorous, but formally complicated, version of the old marginalist theory. Actually, it is not so. Notwithstanding these theories are both neo-classical as far as the essential features of the representation of the economic system are concerned (see footnote 2), their views about income distribution are utterly different. The marginalist idea according to which production processes employ services provided by factors of production, so that distributive variable are the prices of these services, is completely absent in Arrow-Debreu theory. Specifically, advocates of the Arrow-Debreu approach are opposed to the interest rate understood as the price for the use of capital (see Koopmans, 1957, pp. 113-115, and Bliss, 1975, pp. 10 and 346).

[18] On the notion of own rate of interest, see in particular: Debreu (1959, pp. 33-34)

prices of commodity n (with n = 1, 2, …, N) delivered in period t and $t+1$ respectively, the own-rate of interest of commodity n between the two periods $r_{t,t+1}^n$ is defined by the equation:

$$\frac{p_t^n}{p_{t+1}^n} \equiv 1 + r_{t,t+1}^n \qquad (4.10)$$

It is clear from equation (4.10), that this rate (or factor) of interest in neither the price of a factor of production, nor a source of income. It is just a relative price: $(1 + r_{t,t+1}^n)$ is the quantity of commodity n delivered in period $t+1$ that an agent must pay in order to have a unit of commodity n delivered in period t. Besides, since this quantity can be less than 1, the own rate of interest $r_{t,t+1}^n$ can be negative (but not smaller than -1). Hence, as is clear, the own-rates of interest are essentially useless for the explanation of incomes from capital.

4.4.2 Profits in a private ownership economy

As far as profit is concerned, the Arrow-Debreu theory takes up the marginalist idea that the difference between revenues and costs forms the profit of firms. Hence, in this theory, in contrast to the classical/Marxian approach, profit has no linkage with the investment of capital.

If there are N different goods and services with T possible dates of delivery, putting aside the possibility of different places of delivery[19], there are $L = N \times T$ Arrow-Debreu commodities. Let $\mathbf{p} \in \mathbb{R}_+^L$ be a price vector and $\mathbf{y}^f \in \mathbb{R}^L$ the production plan of firm f – that is, a list of quantities of inputs, with negative sign, and outputs, with positive sign – then $\pi^f = \mathbf{p} \cdot \mathbf{y}^f$ is the firm's profit[20]. In a 'private ownership economy'

---

and Bliss (1975, pp. 51-55).

[19] The existence of just one possible place of delivery and one stream of events can be assumed here for simplicity.

[20] In other words, let $\mathbf{y}^f = [y_1^f, y_2^f, …, y_L^f] \in \mathbb{R}^L$ be the production plan of a firm f, and then $\mathbf{y}^f$ is a vector of net supplies of commodities. This means that if $y_n^f < 0$, then it (taken in terms of absolute value) is the quantity of commodity n employed as input by

(Debreu 1959, pp. 78-80), this profit – which can be gains or losses – is divided amongst households, and entered into their budget constraints, in accordance with some exogenously given shares.

These shares cannot reflect the investment of capital (savings) households made for the very simple reason that saving and investment are inconceivable within this framework. As is known, in the Arrow-Debreu model, the L commodities, or rather, the promises of their delivery, are traded simultaneously in a single instant, the initial moment of the first period. This assumption has a number of relevant implications, one of which will be focused on here, namely, the impossibility of finding a role for the investment of capital[21].

On the producers' side, if all the markets are open for one single instant only, then every firm can trade both inputs and outputs simultaneously, in the single instant in which markets are open. In other words, revenues and costs, in this model, are necessarily simultaneous, and this makes the investment of capital impossible. In fact, as we have seen, capital is invested in order to finance the costs in advance, before revenues are obtained. For instance, as is known, if wages are paid in the same moment that output is sold, no capital is involved in the payment. In the Arrow-Debreu framework, this applies not only to wages, but also to expenditure on all the inputs (including capital goods), which must necessarily take place in the same instant as the outputs obtained with them are sold.

On the consumers' side, households cannot and do not need to move their purchasing capacity across time. Households' wealth arises and is entirely spent in the one instant in which the markets are open. Saving in order to transfer purchasing power to some future date would thus actually be impossible in the Arrow–Debreu model, as no further trade can take place after the initial instant of the first period[22]. As Currie

---

firm f. If instead $y_n^f > 0$, then it is the quantity of commodity n obtained as output by that firm. As a result, $\mathbf{p} \cdot \mathbf{y}^f$ directly expresses the difference between revenues and costs.

[21] For a survey of the literature on saving and investment in intertemporal equilibrium models, the reader is referred to Fratini (2019a).

[22] Some authors claim that the assumption that markets do not reopen after the initial instant is innocuous because even 'if markets were reopened at later dates for the same

and Steedman (1990, p. 147) have pointed out, the idea of transferring wealth over time has no real meaning in this framework[23]. It is clear that capital cannot exist in this theory, and that, therefore, no form of income from capital, be it profit or interest, can exist either[24].

Finally, we can stress that, independently of any consideration about its nature and the principle adopted for its distribution, the profit maximized by firms is not a real source of households' wealth – at least if constant returns to scale are assumed. In fact, let $\mathbf{p}^*$ and $\mathbf{y}^*$ be equilibrium price vector and aggregate production plan, it can be easily proved[25] that

Arrow-Debreu commodities, then no additional trade would take place anyway' (Geanakoplos 1987, p. 122). This is not so. The question is not whether further trade will take place when the markets reopen, but rather, whether agents will behave differently in the initial instant in the knowledge that markets will open again. In particular, the assumption that markets will reopen gives rise to all the problems connected with expectations and speculative trade, which are completely avoided in the Arrow-Debreu framework.

[23] In each period there can be – and typically there are – existing stocks of commodities (produced or inherited from the past) that exceed the quantities consumed in that period. Is this saving? The point is complex, and an answer cannot be given here. However, we want to warn the readers about a possible confusion between consumption understood as the activity of consuming and consumption as an expenditure. It is this latter concept that is generally related to saving and capital accumulation. On the point, see also Fratini (2020).

[24] With the aim of avoiding possible misunderstandings, it should be stressed that this statement does not refer to the neo-Walrasian theory in general, but to the Arrow-Debreu equilibrium theory. As already pointed out, the latter is a very peculiar version of the neo-Walrasian approach in which a number of complications are ruled out thanks to *ad hoc* assumptions. By contrast, the phenomena related to the investment of capital can very well be found in neo-Walrasian models in which trades take place sequentially.

[25] Let $\mathbf{z} \colon \mathbb{R}_+^L \to \mathbb{R}^L$ be the market excess-demand function from households. Because of the market clearing condition, we know that $\mathbf{z}(\mathbf{p}^*) = \mathbf{y}^*$. Since the Walras's Law entails $\mathbf{p}^* \cdot \mathbf{z}(\mathbf{p}^*) = 0$, then $\mathbf{p}^* \cdot \mathbf{y}^* = 0$. See the proof of proposition 17.F.1 in Mas-Colell, Whinston and Green (1995), p. 607.

This means that there can be no profit in aggregate, but it seems possible, at first glance, that one firm can have strictly positive profit if another has losses. However, if the set of feasible production plans includes the possibility of inactivity (the null vector), then no firm, in equilibrium, will abopt a production plan that entails losses. As a result, in equilibrium, each firm must have zero profit.

$\pi^* = \mathbf{p}^* \cdot \mathbf{y}^* = 0$. This means that, in equilibrium, households' purchasing capacity depends on the value of their endowments of commodities only, as there is no income coming from firms' profit.

4.5 Conclusions

Three different approaches are considered in this paper. In the classical/Marxian approach (section 4.2), incomes from capital are understood as a surplus whose amount depends on the class conflict. At the beginning of the process of production, a certain amount of capital M is invested in order to advance the costs for the inputs employed. The inputs are transformed into a bundle of commodities C. When these commodities are sold on the market, an amount of revenues M' is obtained. It is the capitalist circuit M-C-M' and capitalists' incomes arise from the difference $M' - M$. This difference or surplus value is the profit, that is, the residual part of revenues over and above the costs of production.

By contrast, in the marginalist theory (section 4.3), the distribution variables – i.e. wage rate, rent rate, and interest rate – are understood as the prices of three factors of production: 'labour', 'land', and 'capital'. According to this view, incomes from capital are what firms pay to households for the employment of the factor of production called 'capital'. In particular, the rate of interest – seen as the price for the use of capital – is determined by means of a supply-and-demand equilibrium.

Thanks to the capital theory debates of the 1960s, it became clear that the idea of capital as a factor of production and the rate of interest as the price of its use is untenable. This was explicitly admitted by authoritative neoclassical economists, such as Samuelson (1966), Hahn (1982), and Bliss (1975). However, at the same time, these scholars maintained that the modern general equilibrium theory is not affected by those problems, since it does not rest on the idea that production processes employ factors of production, but instead, employ Arrow-Debreu commodities.

As a matter of fact, in the Arrow-Debreu general equilibrium

model (section 4.4), there is neither capital understood as the amount of value invested at the beginning of each process, nor capital as a factor of production. The problem is that, in this model, there is no idea or role for capital. As a result, incomes from capital can hardly find an explanation within the Arrow-Debreu theory.

As capital, in the Arrow-Debreu model, is not understood as a factor of production, the traditional conception of the rate of interest is no longer there. In its place, we find many commodity own-rates of interest, but they are essentially useless in the explanation of income distribution. Profit – intended as what firms maximize – is still there, but: i) it is distributed across households on the basis of shares arbitrarily assigned; and ii) its equilibrium amount must be zero.

Therefore, on the basis of the analysis developed in this paper, we can conclude that neither the marginalist theory, nor the Arrow-Debreu theory, can provide us with a convincing theoretical explanation of incomes from capital[26]. Thus, the old theory of the classical economists – submerged and forgotten because of the advent of the neoclassical approach – seems to be the only possible way to proceed.

References

Bliss, C.J. (1975). *Capital Theory and Distribution of Income*. Amsterdam/New York: North-Holland/Elsevier.
Debreu, G. (1959). *Theory of Value*. New York: Wiley.

---

[26] The result achieved here is much more general than that proposed by Eatwell in a recent paper, in which he claims that 'there is no neo-classical theory of the rate of profit' (Eatwell 2019, p. 10). More precisely, Eatwell maintains that there is no way to use Debreu's theory in order to solve Adam Smith's problem concerning the determination of 'natural prices.' Consequently, it cannot arrive at the determination of the level of the ordinary/normal rate of profit (in this respect, see section 4.2.2 here). Although Eatwell's conclusion is certainly agreeable, it seems to be just a corollary of the point already raised by Garegnani (1976 and 2012) about the neo-Walrasian approach and its abandonment of the 'normal position' method.

Eatwell, J. (2019). 'Cost of production' and the theory of the rate of profit. *Contributions to Political Economy*, 38(1), 31-58.

Fratini, S.M. (2019a). On the second stage of the Cambridge capital controversy. *Journal of Economic Surveys*, 33(4), 1073-1093.

Fratini, S.M. (2019b). Neoclassical theories of stationary relative prices and the supply of capital. *Metroeconomica*, 70(4), 723-737.

Fratini, S.M. (2020). Interest, profit and saving in Arrow-Debreu equilibrium models. *Bulletin of Political Economy*, forthcoming.

Garegnani, P. (1976). On a change in the notion of equilibrium in recent work on value: a comment on Samuelson. In Brown, M., Sato, K. and Zarembka, P. (eds.), *Essays in Modern Capital Theory*. Amsterdam: North-Holland.

Garegnani, P. (1984). Value and Distribution in the Classical Economists and Marx. *Oxford Economic Papers*, 36(2), 291-325.

Garegnani, P. (2012). On the present state of the capital controversy. *Cambridge Journal of Economics* 36(6), 1417-1432.

Geanakoplos, J. (1987). Arrow-Debreu model of general equilibrium. In Eatwell, J., Milgate, M. and Newman, P. (eds.), *The New Palgrave Dictionary of Economics*, Vol. I. London: Macmillan, 116-124.

Hahn, F.H. (1982). The neo-Ricardians. *Cambridge Journal of Economics*, 6(4), 353-362.

Harcourt, G.C. (1969). Some Cambridge controversies in the theory of capital. *Journal of Economic Literature*, 7(2), 369-405.

Harcourt, G.C. (1972). *Cambridge Controversies in the Theory of Capital*. Cambridge: Cambridge University Press.

Koopmans, T.C. (1957). *Three Essays on the State of Economic Science*. New York/London/Toronto: McGraw-Hill.

Marx, K. (1909). *Capital*, vol. III. Chicago: C.H. Kerr & Co.

Mas-Colell, A., Whinston, M.D. and Green, J.R. (1995). *Microeconomic Theory*. Oxford: Oxford University Press.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Cambridge (Mass.) and London: Harvard University Press.

Rognlie, M. (2016). Deciphering the fall and rise in the net capital share: Accumulation or scarcity? *Brookings Papers on Economic Activity*, 2015(1), 1-69.

Samuelson, P.A. (1962). Parable and realism in capital theory: the surrogate production function. *Review of Economic Studies* 29(3), 193-206.

Samuelson, P.A. (1966). A summing up. *Quarterly Journal of Economics*, 80(4), 568-583.

Walras, L. ([1926] 1977). *Elements of Pure Economics or the Theory of Social Wealth*, Jaffé, W. (Trans.). New York: Augustus M. Kelley Publishers.

Wicksell, K. ([1911] 1934). *Lectures on Political Economy*, vol. I, Classen, E. (Trans.). London: Routledge.

Wicksteed, P.H. (1910). *The Common Sense of Political Economy*. London: Macmillan.

Chapter 5

'Ecological contract' and green economy

Francesco Longobucco

The Private Law has a privileged eye on the Enviromental Law, as the Private Law is a whole of old instruments which must be bent to create new functions and therefore must evolve (this phenomenon is so-called '*heterogenesis of the purposes*' of the rules contained in our Civil Code). Paradigmatic is the case of illegal discharges (cf. art. 844 C.C.).

Thus, I would underline the importance of creating a framework of method to be followed in the analysis of the institutes, for those who want to cooperate in the development of Enviromental Law also on the civil point of view.

Under this point of view, for example, the Enviromental Law impacts today's theory of property. Is the environment a classical good? It is a particular good without an owner, so that the traditional notion of property in the classical sense is no more available, while a potential ownership emerges [that, for example, of the future generations (*in incertas personas*)], the physicality of the traditional goods is outdated in our case, validating the idea that today goods must be considered more as *interests* than as *res* just as the ancients meant.

The Enviromental Law also impacts with the same notion of the modern contract and the negotiating activity. It becomes, for scholars, the test bench on which to sample what, by my cultural conviction, is the current crisis of the classical dichotomy between private interest and public interest, between proprietary interest and non-pecuniary interests (cf. art. 1174 C.C.). If we assume that the freedom of contract is no more a *dogma* and that it is instead today increasingly a synthesis of the autonomy of the parties and the heteronomy of the legislator (as autonomy of parties is no more an uncontrolled and uncontrollable

power), then, the fundamental right to a healthy environment, to be realized in a pro-active way, becomes one of the limits (internal or external is a little matter) to the traditional freedom of contract. Perhaps handbooks of Italian Private Law should be updated when we talk about the traditional limits to the autonomy of parties, as, in the Legal Public Italian and European order, the interest in a healthy environment, according to the ultimate protection of the person, must be certainly inserted. The interest in a healthy environment characterizes the inner '*causa*' of the legal acts between private individuals and between private and public administrations [see, for example, all the interesting matter of the '*appalti verdi*' (green public contracts) or the 'CAM clauses' to be entered in the public contracts].

Here comes the paradigm of the '*ecologically conformed contract*' in the new scenario of the Green Economy. This paradigm finds its base in Art. 3 *quater* of the Environmental Consolidated Law (the Italian *'Testo Unico Ambientale'),* according to which every legally relevant human activity (under the Code) must comply with the principle of sustainable development, in order to ensure that satisfaction of the needs of current generations cannot compromise the quality of life and the possibilities of future generations. Also the activity of the public administration must be aimed at allowing the best possible implementation of the principle of sustainable development, for which, in the context of the discretional comparative choice between the public and the private interests, those aimed to protect the environment and the cultural heritage must be object of a priority consideration.

And what about the Italian Constitution? Already Art. 9 of the Italian Constitution (about the landscape protection) indirectly protects the interest to a healthy environment, according to the best public scholarship. Then, it comes into evidence Art. 41, para 2, of the Italian Constitution (with its limits to the autonomy of parties regarding safety, freedom, public utility). Thus, it is not true that the protection of the environment does not exist in our Constitution, rather it exists – albeit indirectly – also considering that the Constitutional provisions are directly applicable to the activity of the private parties. All of this with the implication on the level of applicable civil remedies: it is not

surprising if the violation of the principle of sustainable development, as a principle of legal public order, could lead, as some scholars have claimed, to the nullity of the contract according to Art. 1421 C.C. which can be lodged *ex officio*.

How then does the traditional freedom of contract could be ecologically conformed? Certainly it could be conformed through the Italian and European general fundamental Principles: let's think, for example, to the principles of prevention, precaution, sustainability, energy efficiency, 'horizontal' subsidiarity according to Art. 118 of the Italian Constitution.

Then, the ordinary rules and the subsidiary rules come to evidence with the possibility of a wide expansion of a Regional Private Law, specific for the territory, that could also conforms the contractual activity.

Yet, let's think to the hetero-introduced rules in the contractual activity (Art. 1374 C.C.): let's consider, for example, the various rules of the Italian Regulatory Authority for Energy and Gas which, in the context of the general phenomenon of the regulatory 'hetero-integration' of the contract, create a whole of secondary rules (those on the price, for example) in order to conform the autonomy of parties. Let's think also to the 'autopoietic rules' (a sort of best practices that the companies themselves have to take) which also conform the contractual regulation.

In this way the traditional freedom of contract becomes more consistent with the environmental interest. And it doesn't matter if it's the fifth or the sixth model of contract: the first contract of the Civil Code, then we have the labor contract, then the consumer contract, then the contract between companies, now the ecologically conformed contract. A sixth contract.

And what about the function of this new model of contract, that is the ecologically oriented contract? We can surely think to a preventive function, to a compensation function, to a punitive function, but also to a 'heuristic and promotional' function, as suggested by some scholars, which is to say a stimulus to develop the interpretative and applicative attitude of the Italian jurists in setting point of a paradigm (the model of the ecologically oriented contract) which, although not specifically regulated, can certainly be legitimized in the Italian and European legal system.

Chapter 6

Integration of third country nationals and the European Union: an opportunity not to be missed[*]

Valeria Piergigli

6.1 Introduction: integrate or select immigrants?

There is nothing new about migrations. Individuals and peoples have always been on the move, in pursuit of better living conditions or mere survival, attempting to escape from wars, persecution, famine and various other adversities.

Over the last few decades the European continent has, for a number of reasons, been progressively an area of immigration, and it seems highly likely that the trend will be borne out in the years to come. According to surveys updated as to 1 January 2019, 21.8 million citizens from third countries are living in the European Union, accounting for 4.9% of the total population of the 27 States in the Union[1]. The increasingly intense migratory inflows from third countries, especially as from the beginning of the 21st century, has made a burningly relevant issue of the need to forge adequate tools at the various levels of government – local, national and supranational – to regulate realities

---

[*] The present contribution is an abridged and updated version of the essay *L'integrazione degli immigrati da paesi terzi nel diritto sovranazionale: limiti e potenzialità dell'Unione europea,* published in *Rivista AIC,* 3/2018, 1-26 and in G. Cerrina Feroni, V. Federico (eds.), *Strumenti, percorsi e strategie dell'integrazione nelle società multiculturali*, Naples, 2018, 209-241.

[1] The data are reported by Eurostat: <https://ec.europa.eu/eurostat/statistics-explained/index.php/Migration_and_migrant_population_statistics>.

that appear to have gone through quantitative and qualitative change. In fact, not only has a considerable increase in the number of persons choosing to migrate to the European continent been registered, but also the typology of migrants has changed in comparison with those of the last century. Apart from the particular – and often tragic – conditions of refugees, asylum-seekers and, in general, persons venturing into our territories in search of humanitarian protection[2], there is an increasingly generalised tendency of nationals from outside the EU arriving in Europe for economic reasons and for work not to return to their countries of origin, but to settle down in the places of immigration, and possibly send for the members of their families to join them. The powerful impact – demographic, social, cultural, religious, economic and political – resulting from these developments has driven the States of the European Union to rethink, or at any rate question, the models long tried in the approach to migration from third countries[3].

Another factor, subsequent to the tragic events of September 2001, prompting the national legislators to reform the regulations at present in force lies in the threat of international terrorism in the name of Islam and the associated need to safeguard the borders and maintain social cohesion. This has led to the adoption of revised and stricter migration policies, for which the EU simply outlines the common principles and approaches, leaving to the discretionary powers and sovereignty of the Member States regulation of the flows and concrete definition of the requisites for entry, residence, integration and – should it prove the case – naturalisation of the foreigners legally settled in the respective territories. The action taken in Italy to limit immigration and integration in the last decade in the name of public safety and protection of the national borders has been widely publicised[4].

---

[2] On these aspects, in Italian discussion, cf., for example, Woelk *et al.* (2016).

[3] Among the best-known, tried and tested models are to be found the assimilationist and multicultural approaches, on which see the various essays published in the volume edited by Cerrina Feroni and Federico (2018) and in Cerrina Feroni and Federico (2017).

[4] Consider, for example, the modifications introduced a decade ago in the Consolidated Law on immigration (legislative decree 286/1998) with the so-called 'security package'

In particular, the issue of integration of immigrants from outside the EU – linguistic and civic, to begin with – has for a good many years featured recurrently in the European political agenda, and has taken on central importance in the area of national policies. It is no longer being approached, as in the past, solely for the purposes of granting the status of citizenship, but also as requisite (or prerequisite) for entry, residence and access to certain social benefits in the host country.

But what exactly does 'integration' mean? In legal language there is no universally accepted definition of the term, which thus remains an indeterminate concept. In 2009, however, for the purposes of application of the Consolidated Law on immigration, Italian legislators established that integration is the process serving «to promote the coexistence of Italian and foreign citizens in respect of the values incorporated in the Italian Constitution, with reciprocal commitment to participating in the economic, social and cultural life of society» (art. 4-*bis* legislative decree 286/1998, introduced by art. 1, clause 25, Law 94/2009).

Effectively, the concept of integration – which, apart from matters of law, more broadly has to do with the human and social sciences – should, as long called for by the European organisations, lead to a two-way process involving both immigrants and host countries in a

---

(in particular: decree law 92/2008, converted into l. 125/2008, e l. 94/2009), to which, in the 18th legislature, the changes brought in with the so-called 'security decree' (decree law 113/2018, converted with modifications into l. 132/2018, Disposizioni urgenti in materia di protezione internazionale e immigrazione, sicurezza pubblica, nonché misure per la funzionalità del Ministero dell'interno e l'organizzazione e il funziona-mento dell'Agenzia nazionale per l'amministrazione e la destinazione dei beni seque-strati e confiscati alla criminalità organizzata), and the 'security decree *bis*' (d.l. 53/2019, converted with modifications into l. 77/2019, Disposizioni urgenti in materia di ordine e sicurezza pubblica. On the occasion of promulgation of the conversion law, the President of the Italian Republic pointed out to the President of Parliament certain unconstitutional aspects). Also worth mentioning is decree law 4/2019, converted with modifications into l. 26/2019 (*Disposizioni urgenti in materia di reddito di cittadinanza e pensioni*) which made supply of the basic income, designed to promote labour market entry and social inclusion conditional upon the requisite of at least 10 years residence in Italy, thus clearly favouring certain categories of beneficiaries.

combined effort[5]. True integration means rejecting the idea of assimilation, acculturation or artificial standardisation of the other, the foreigner. Rather, it requires the authorities of the countries of immigration to take measures to ease his/her inclusion. Naturally, the immigrant is expected to play an active and responsible part in the process of integration in the host country. At the same time, it is essential to favour respect and valorisation of the cultural identity of the immigrant populations which, as we well know, are ascribed to the category of 'new minorities', to distinguish them from the autochthonnous linguistic minorities with a long history in the territory of a great many European countries.

Nevertheless, it seems to be an undeniable fact that

«to some extent the process of assimilation by the majority component of the population is inevitable in that the legal order does however require the immigrant and the community to which he or she belongs to adapt to what constitutes the *values* characterising the host system – values customarily taken to be positive: liberty, equality, non-discrimination…»[6].

The fact is that for some time we have been witnessing something that probably goes beyond the requirement for the third country nationals (TCNs) to adapt to the basic principles of the host State. Suffice it to consider the progressive tightening of national regulations regarding demonstration of a certain level of knowledge of the language (and, often, also of the values and institutions) of countries which TCNs seek to enter for the purposes of family reunification or finding work and living. These are certainly legitimate requisites, and apparently serve for social inclusion, but they are structured and implemented in such a way in some cases as, in practice, to select the immigrants or would-be immigrants, excluding or repelling those unable to conform with the prescriptions or at any rate making the integration

---

[5] In this respect see 'The Common Basic Principles for Immigrant Integration Policy in the EU, unanimously adopted by the Justice and Home Affairs Council of the European Union (The Hague 2004).

[6] Translation from: de Vergottini (1995, p. 23), author's italics.

process excessively difficult for them.

In such cases integration does not seem to be so much a goal to achieve through a series of activities towards which the host country should contribute as, rather, an obligation (in terms of means or result) to be observed, or at any rate a burden in some cases placed entirely on the shoulders of the immigrant. Entry or residence in the territory of a State are made conditional upon having achieved a certain degree of integration, almost as if recognition of the legal condition of immigrant and a series of rights were a reward that had to be deserved, reserved for those who can demonstrate they have (already) become 'perfect citizens'[7]. This generates some tension, yet to be resolved, between the aspiration universally (and officially) proclaimed to respect of the pluralism of cultures and the principles of the liberal, democratic tradition, on the one hand, and the implementation of practices directed towards control of the diversity and assimilation in the majority culture on the other, in a climate that all too often appears contradictory and likely to exacerbate social conflict. Thus the situations of irregularity, illegality and marginalisation that efforts are ostensibly being made to eliminate or at least reduce are in fact aggravated.

Actually, in an increasingly cosmopolitan and globalised society characterised by the supersedence of economic and cultural barriers, decision-making on the management of national borders and the status of the resident population – through choices in the realm of immigration, integration and citizenship – remains strictly within the sphere of national sovereignty. The European institutions, for their part, endeavour – as far as possible, with both hard-law and, above all, soft-law instruments – to circumscribe the discretionary powers of the States and establish certain common, albeit minimum, standards for the integration of TCNs who apply for European long-residence permits or seek to reunite their families. Integration represents a sphere within which the EU inevitably comes up against certain limits, as explicitly emerges from the provision that the European Parliament and Council can bring in measures to support and incentivise Member States in

---

[7] Paraphrasing Carrera (2009).

promoting the integration of regularly resident extra-EU immigrants, but excludes «any harmonisation of the laws and regulations of the Member States» (TFEU art. 79.4).

Given all these circumstances, our aim in this paper is to focus attention above all on the potential available at the supranational level. In this respect, the contribution made by the Court of Justice – specially petitioned in the course of preliminary referral by the national judges – is showing great potential in settling some of the major issues involved in interpretation of the notion of integration, apodictically encapsulated in the texts of directives 109 and 86 of 2003. It is also proving valuable in dispelling the doubts that some national regulations raise with reference to the effectiveness, reasonableness, proportionality and non-discrimination of the measures adopted (and formally justified) in the name of socio-cultural inclusion of foreigners in the national host community.

The movement of people from one country to another is an irreversible tendency which needs to be faced up to with mature awareness, no longer to be addressed as a problem to be approached solely in terms of security. The need is, rather, to take it as an opportunity to improve not only the management of migratory flows, but also the policies for the integration of legally resident TCNs. It is a challenge that, if taken up in a spirit of solidarity and forward-looking, could open the way to a strong reaction against the tarnishing of the principles of legality and democracy, as well as bringing some restraint to bear on the xenophobic impulses and nationalistic backsliding variously emerging in diverse parts of Europe (and not only there) in recent times. Furthermore, it is a challenge that could mark a turning point in the direction of effective integration amongst the States and populations prepared to place their hopes on a newfound sense of unity, today more essential than ever in the old continent. In this perspective, more effective integration of extra-EU citizens could prove to be a factor in achieving greater European integration, as well as reinforcing the inalienable values upon which the Union rests and which have inspired the liberal-democratic tradition of the Member States.

6.2 Directives 2003/86/EC and 2003/109/EC: between *measures* and *conditions* of integration

When the Treaty of Lisbon came into force the aim to integrate TCNs legally resident in the States of the Union found an official place in the European political agenda, albeit with the limitations mentioned above. And yet before this, on the basis of article 63 of the Treaty of Amsterdam, the supranational institutions had already taken steps in the direction of common regulations which also brought in reference to the issue of integration. Directives 2003/86/EC and 2003/109/EC, respectively dealing with regulation of family reunification and the status of TCNs who are long-term residents were designed to regulate and circumscribe the decisional autonomy of the States, which however retained the right to self-determination in the choice and graduation of integration procedures[8].

It is worth taking a look at the protracted and troubled progression towards adoption of the directives in question. The decision-making process involved was the one enshrined in the Treaty of Amsterdam, hinging on the logic of intergovernmental cooperation and the rule of the unanimous vote in the Council. Arriving at the final conclusion took several years of negotiations, during which certain – less than transparent – procedures carried out within the Council and the resistance set up by certain States considerably watered down the initial proposals with which the Commission had intended to implement the Tampere programme objectives of the European Council dating back to 1999. Essentially, these objectives consisted in bringing national regulations to correspond more closely, so as to allow for equal treatment of regularly resident extra-EU citizens and promote more effective integration, acknowledging for them a series of rights and duties comparable with those of the EU citizens.

---

[8] The directives analysed in the text are, of course, not the only ones that refer to the integration of foreign nationals: such reference is also to be found, for example, in directive 2004/114/EC on the admission of citizens from third countries in EU States for reasons of study, or directive 2009/50/EC on the conditions of entry and residence of TCNs for the purpose of highly qualified employment.

In the final version of the two directives, the approach to integration turned out to be radically different from the original projects, mainly due to clauses for derogation from or at any rate limitation of recognition of the immigrants' enjoyment of rights, called for by Germany, Austria and the Netherlands. The undeclared but clearly understood intention behind the counterproposals was to control and reduce immigration from third countries. The formulation of restrictive provisions found justification in the need to conform the European regulations about to be adopted to the – already fairly selective – provisions in force or pending approval in those countries. In other words, what the German, Austrian and Dutch representatives really wanted was to maintain the status quo and, if possible, receive some sort of legitimation from above, leaving uncompromised the faculty of the national legislators to adjust their integration policies in the future. A consequence of this approach was the need, beginning at the stage of incorporation of the European regulations into the various countries' systems, to tighten up the regulations, so minimal and equivocal was the standard set at the supranational level. This eventually had, among the various results, that of triggering a downward domino effect, as it were, the Member States following one another in a spirit of reciprocal emulation, taking a distance from the Tampere programme and bringing in increasingly demanding requisites integration, beginning at the linguistic level, precisely for the sake of harmonising regulations and developing common practices[9].

A glance at the contents of the two directives can help to make these dynamics clearer.

Directive 2003/86/EC mentions the word 'integration' eight times and calls on the Member States to encourage family reunification since «It helps to create sociocultural stability facilitating the integration of third country nationals in the Member State» (4th whereas). However, the directive also recognises the *faculty* of the Member State to require of

---

[9] On the troubled road to adoption of the two directives and subsequent (restrictive) national regulations through incorporation, see among others Carrera (2014, especially pp. 171-173), Block and Bonjour (2013).

TCNs seeking family reunification, among the other requisites[10], «to comply with integration *measures*, in accordance with national law» (art. 7. clause 2) (emphasis added). Moreover, when deciding whether to authorise reunification of minor children over the age of twelve years who arrive independently of the rest of the family, «*may* verify […] whether he or she meets a condition for integration provided for by its existing legislation on the date of implementation of this Directive» (art. 4, clause 1) (emphasis added).

For its part, directive 2003/109/EC, modified in 2011 to extend the status to refugees and other beneficiaries of international protection (dir. 2011/51), mentions the term 'integration' five times and opens by stating that the integration of TCNs who are long-term residents in the Member States «is a key element in promoting economic and social cohesion, a fundamental objective of the Community stated in the Treaty» (4th whereas). Having stated so much, the text makes provision that, for the purpose of acquiring the status of long-term resident, the States *may*, in addition to other requisites[11], determine whether TCNs «comply with integration *conditions*, in accordance with national law» (art. 5, clause 2) (emphasis added). Should the TCN who has obtained from the authorities of a Member State recognition of the status of long-term resident intend to reside in the territory of another Member State the latter *may* require compliance with «the integration *measures* in accordance with national law», unless the integration *conditions* as indicated in the above-mentioned art. 5, clause 2 (art. 15, clause 3) have already been satisfied in the first Member State (emphasis added). In this

---

[10] According to art. 7, clause 1, the Member State may require reunification applicants to demonstrate use of appropriate accommodation, health insurance and resources sufficient to support self and family members. In the case of refugees and their family members, the above integration measures can be applied only after reunification has taken place (art. 7, clause 2), thereby implying that these measures may be imposed on other categories of migrants *before* they enter the host country.

[11] According to art. 5, clause 1, to acquire the status of long-term residents, the Member States require TCNs to prove that they have sufficient resources to support themselves and their dependent family members and insurance against illness. A further requisite is uninterrupted legal residence in the territory of the country for five years (art. 4, clause 1).

case it is assumed that the TCN already integrated in the society and lifestyle of a Member State is a person automatically integrated or readily integrable in the societies of the other States, with no need for further formalities[12]. However, given also a variety of languages in Europe, it is possible that the persons concerned may be required to attend language courses in the second Member State (art. 15, clause 3). Recognition of the status of long-term resident is permanent, with the exception of loss or revocation in the cases expressly indicated (art. 9), and entitles persons to equal treatment with the national citizens as regards access to employment, education, goods and services, social security and social assistance, tax benefits and freedom of association (art. 11)[13].

Thus, with a certain ambiguity, the European regulations alternate references to measures and conditions of integration without providing definitions and in any case leaving to the laws and practices of each country concrete specification of the measures and conditions. Suffice it to observe that in the Dutch version of directive 2003/86/EC, the word 'measures' is represented with the term 'conditions' of integration[14].

It is generally held that while the former make the immigrant responsible for compliance with obligations regarding means (e.g. attending a course to learn the language of the host country) and may also consist in requisites to be complied with before entry into the Member State (so-called integration from abroad), the latter entail obligations at the level of results (e.g. passing a language test)[15].

---

[12] See Carrera (2014, p. 157).

[13] Also to be noted is the 12th whereas of the same directive, which states that: «In order to constitute a genuine instrument for the integration of long-term residents into society in which they live, long-term residents *should enjoy* equality of treatment with citizens of the Member State in a *wide range* of economic and social matters, under the relevant conditions defined by this Directive» (emphasis added). So it appears that the equality of treatment invoked should apply to many but not all the economic and social sectors, as is then specified in art. 11.

[14] As indicated in the Conclusions by the advocate general J. Kokott, presented on 19 March 2015, in Case C-153/14 (§21).

[15] Groenendijk (2006, p. 224). The provision in art. 5, clause 2, is the 'Achilles heel' of directive 2003/109/EC according to the findings of Boelaert-Suominen (2005).

According to another point of view, the distinction lies in the respectively optional or mandatory nature of the measures and conditions for integration. In the case of the latter, obligation also extends, in the case of non-compliance, to the application of penalties, ranging from fines to non-renewal of temporary residence permits and expulsion from the territory of the State[16].

Noting the multifarious and incorrect application of directive 2003/86/EC by the Member States, the European Commission pointed out in a note issued in 2014 that, although the competence of the Union for integration takes second place to that of the States, the powers enjoyed by the latter are not unlimited. In particular, the *measures* mentioned in art. 5 must be proportionate and applied with the flexibility necessary to avoid undermining the useful effect of the directive, which consists in promoting, and not obstructing, family reunification. To this end, the Member States should take into account any particular individual circumstances (cognitive capacity, vulnerability, lack of access to facilities for language learning or other disadvantageous situations) and, should it be the case, provide for derogation or postponement for compliance with the measures[17].

---

[16] On the distinction between integration measures and conditions as used in the directives referred to in the text, see also Conclusions by the advocate general M. Szpunar, presented on 28 January 2015, Case C-579/13 (especially §§55, 85, 86 e 97), according to which the integration measures are to be considered less demanding than the integration conditions and cannot, therefore, include the obligation to pass an integration test, nor serve as a means to select the immigrants or control immigration. Along the same lines, see also Conclusions by the advocate general P. Mengozzi, presented on 30 April 2014, Case C-138/13. On the meaning of the concept of integration in the two directives, see furthermore at the level of interpretation: Hailbronner and Klarmann (2016) and Thym (2016a).

[17] See the *Communication from the Commission to the European Parliament and the Council on guidance for application of Directive 2003/86/EC on the right to family reunification*, Brussels 3 April 2014, where we read: «In other words, the integration measures that a Member State may require cannot result in a performance obligation that is in fact a measure that limits the possibility of family reunification. The measures must, on the contrary, contribute to the success of family reunification» (p. 17), adding that automatic refusal of reunification subsequent to failure in an integration test could constitute a violation of arts. 5.5, 8 and 17 ECHR (p. 17, note 55).

The fact is that the regulations adopted and progressively revised in various European countries, in some cases precisely on the occasion of incorporating the above-mentioned directives[18], have more generally adopted integration 'pre-requisites' or 'pre-conditions' – in the first place in terms of language – so devised as to translate into barriers to entry or obstacles to residence in the territory of the Member State of the TCNs submitting, respectively, application for family reunification or recognition of the status of long-term residents. An interesting corpus of case law by the Court of Justice on the legitimacy of these interventions is taking shape, as we will see below.

6.3 The contribution of the Court of Justice to interpretation of the national regulations

As we have seen, the Treaty of Amsterdam and, even more, the Treaty of Lisbon have led to the Europeanisation of immigration law and, to a lesser extent, of integration policies. This has entailed an extension of the tasks assigned to the Commission, and above all of the role of the Court of Justice which, with the Treaty of Lisbon, has acquired full jurisdiction over the measures adopted in accordance with TFEU art. 79, including the possibility of pronouncing on a preliminary referral without the limits set by the Treaty of Amsterdam[19]. Thus, if expectations are to see a process unfolding towards the implementation of a common model for integration of immigrants shared by the Member States, a certain degree

---

[18] Denmark, Ireland and the United Kingdom are not bound by directives 2003/86/EC and 2003/109/EC having made use of the opt-out clause. For further considerations on the regulations adopted in implementation of the above-mentioned directives on the part of some Member States, let me refer readers to Piergigli (2013).

[19] On the basis of art. 68, clause 1. Treaty of Amsterdam (TCE), preliminary referral to the Court of Justice in the immigration sector was admitted only for national courts of last instance, i.e. the courts against whose decisions no further appeal can be made according to the national law. For the extension of the competencies of the European Commission and the Court of Justice with the Treaty of Lisbon, see, at: Wiesbrock (2010), Block and Bonjour (2013) and Carrera (2014).

of optimism now seems to be justified.

However, if TFEU art. 79, clause 4 rules out EU competence to adopt binding integration policies and the directives of 2003 leave the matter to the discretionary powers of single Member States, what are the parameters that the Court of Justice can invoke?

The question began to call for concrete answers on the introduction and rapid propagation amongst the Member States of obligatory integration mechanisms for entitlement to the status of long-term resident and for family reunification. Here the tests to verify a certain knowledge of the language and institutions of the host country raised delicate issues of the compatibility of national regulations both with the objectives of the 2003 directives and with the general unwritten principles of EU law[20], as well as, more broadly speaking, the provisions included for various reasons in international documents regarding the protection of fundamental rights in conditions of equality. Moreover, even the countries that are not required to apply the directives on immigration have to respect the EU *acquis* and the international obligations deriving from ratification of the treaties, observing at least the principle of non-discrimination in the regulation of language tests for entry and residence in their territories.

In its reports on the implementation of directives 2003/86/EC and 2003/109/EC, the European Commission invites Member States to implement the objectives respectively extended to them and to bear in mind the general principles of EU law, including those regarding the effectiveness and proportionality. In the case of the integration measures and conditions concerning language, the Commission advises that a series of indicators can be used to assess conformity with the supranational regulations. Significant, therefore, will be, for example, the nature and level of language knowledge prescribed, ease of access to integration programmes, didactic material and tests, the costs of courses and tests, whether or not there are procedural guarantees in cases of decisions against entry or issue a long-period residence permits, and

---

[20] For a summary of the contents of these general principles, see Acosta Arcarazo (2011).

comparison with the integration requisites set for European citizens, for whom more rigorous standards would be expected[21].

These caveats have been corroborated by the Court of Justice on the occasion of judgements regarding interpretation of the 2003 directives[22]. Making reference to the various bodies and institutions including the ECHR, the European Social Charter, the Convention on the Rights of the Child, the EU Charter of Fundamental Rights and the ongoing jurisprudence of the Strasbourg judges on the right to private and family life, the Court of Justice has made it quite clear that the discretionary powers granted to national authorities in formulating requisites to be satisfied by the applicant family members for reunification must not lead to denial of such a fundamental right as family life, nor failure to take into account the overriding interests of the children of minor age (Case C-504/03)[23]. Basically, the fact that the concept of integration (and promotion of it) lacks definition both in the text of directive 2003/86/EC and in that of directive 2003/109/EC[24] cannot – according to the Luxembourg judges – be interpreted as giving Member States carte blanche to use the concept in such a way as to clash with the purposes of the regulations laid down at the supranational level which consist, respectively, in guaranteeing family unity and integration of TCNs settled as long-term residents in the territory of the Member

---

[21] See the reports of the European Commission on implementation of directives 2003/86/EC and 2003/109/EC, respectively of 8 October 2008 and 28 September 2011. See also the European Commission Green Paper on the right to family reunification for third country nationals living in the European Union issued on 15 November 2011.

[22] Besides the citations which will be made later on commenting on the individual decisions, on the jurisprudence of the Court of Justice of relevance here, see: Carrera (2014), Block and Bonjour (2013), and Groenendijk (2014).

[23] Court of Justice, 27 June 2006 Case C-504/03 (*European Parliament and EU Council*).

[24] As pointed out in Murphy (2010), the Court of Strasbourg is beginning to develop case-law on integration (also at the level of language and culture) as a criterion to be taken into consideration in cases of expulsion of immigrants; this approach could have fallen out on the interpretation of ECHR art. 8 and on national policies in the area of immigration and integration.

States. Therefore, making the granting of a long-term residence permit conditional on payment of an excessive charge (Cases C-508/10 and C-309/14)[25] or admission of the family members conditional on demonstration of a higher level of income than ordinarily prescribed on the basis of the reunification directive (Case C-578/08)[26] means – according to the Court – contravening the objectives pursued with the directives as well as the principles of proportionality, effectiveness and respect of fundamental rights.

Again, the Court has made it clear that the faculty of the Member States to decide on the granting of subsidies for accommodation for the lower – national and extra-EU – income brackets is to be interpreted restrictively, in such a way as not to compromise the useful effect of directive 2003/109/EC on application of the principle of equal treatment between long-term residents and national citizens laid down in art. 11 of the same directive (Case C-571/10)[27].

An interesting case, albeit concluded with a decision of no need to adjudicate by the Court of Justice, concerned, for the first time in 2011, the legitimacy of an integration requisite to be satisfied by an Afghan national in the country of origin prior to entry in the Member State for the purpose of family reunification. In a preliminary ruling, the Hague court asked the Court of Justice whether the obligation imposed by the Dutch regulations did not constitute an excessively severe interpretation of art. 7, clause 2. dir. 2003/86/EC, and whether «it is

---

[25] Court of Justice, 26 April 2012 Case C-508/10 (*European Commission v. Kingdom of the Netherlands*) and Court of Justice, 2 September 2015 Case C-309/14 (*Confederazione generale italiana del lavoro (Italian General Labour Confederation – CGIL), Istituto nazionale confederale assistenza (National Confederal Assistance Institute – INCA) and Prime Minister's Office, Ministry of the Interior, Ministry of Economy and Finance*).

[26] Court of Justice, 4 March 2010, Case C-578/08 (*R. Chakroun c. Minister van Buitenlandse Zaken*).

[27] Court of Justice, 24 April 2012, and case C-571/10 (*Servet Kamberaj and. Istituto per l'Edilizia sociale della Provincia autonoma di Bolzano (IPES), Giunta della Provincia autonoma di Bolzano, Provincia autonoma di Bolzano*).

relevant that citizens of some other third countries are exempted from the obligation to pass the civic integration test abroad solely by dint of their citizenship» (Case C-155/11)[28]. Effectively, the nationality factor can contribute to implementing actual selection of the quality (and not only of the quantity) of immigration, with consequent violation of the non-discrimination principle which is provided for in international conventions for the protection of human rights, in TFEU (arts. 18 and 19) and in directives 2003/86/EC and 2003/109/EC. Various European jurisdictions, including that of the Netherlands, expressly exonerate entire categories of extra-EU nationals from language and/or civic integration tests solely on the consideration that the provenance from Western countries, or at any rate countries with well-established democracy, suffices in itself to guarantee their integration.

In recent years the Court of Justice has continued to be consulted with increasing frequency for preliminary ruling on the conformity of the civic integration obligations laid down by Dutch law with directives 109 and 86 of 2003. With somewhat perfunctory arguments, the supranational judge confirmed his jurisprudence and arrived at a compromise solution between EU law and the faculty of the Member States to decide on the requisites for integration.

In two judgements on different occasions but one soon after the other, the Court reaffirmed the principle that Member States do not have total control over matters of immigration, for exercise of their powers, filtered with the proportionality test, must not violate the principle of non-discrimination, nor compromise the objectives of useful effect of the European regulations[29]. The Luxembourg judge, favouring a pragmatic approach glossing over the distinction between integration

---

[28] Subsequent to acceptance of the complaint by the Dutch government, which granted provisional residence permit to an Afghan woman, the Court of Justice declared that there was no longer the need to adjudicate on the application for preliminary ruling: Court of Justice, 10 June 2011, Case C-155/11 (*Bibi Mohammad Imran c. Minister van Buitenlandse Zaken*).

[29] For comment on the decisions cited below in the text, see Jesse (2016), Thym (2016b), Strazzari (2015) and Strazzari (2016, p. 447).

measures and conditions[30], ruled that the obligation to pass the civic integration test does not clash with art. 5, clause 2. dir. 2003/1009/EC, even if submitted to immigrants who have already achieved long-term resident status provided that the application procedures do not hinder achievement of the objectives pursued with the directive itself. The procedures are to be appraised by the referring court and can legitimately also include a fine for failure in the test, but any penalty system must be so structured as not to deprive the directive of its useful effect (Case C-579/2013)[31].

Similarly, according to the Court of Justice the faculty of Member States to bring some obligation in terms of integration to bear on the applicant for family reunification from abroad, in accordance with art. 7, clause 2. dir. 2003/86/EC, does not in theory imply that the Member States cannot require TCNs to pass an elementary test on knowledge of the language and society of the Member State concerned before authorising entry or residence in its territory. In practice, however, the obligation must not be such as to make exercise of the right to family reunification impossible or excessively difficult through high costs or failure to take into consideration the individual circumstances of the applicant, such as age, level of education, economic and health conditions (Case C-153/14)[32]. In fact, «The integration measures referred to in the first subparagraph of Article 7(2) of Directive 2003/86 must be aimed not at filtering those persons who will be able to exercise their right to family reunification, but at facilitating the integration of such persons within the Member States» (§ 57).

Reference to a principle of 'personalised proportionality'[33] and

---

[30] On the other hand, the Conclusions of the advocate general M. Szpunar concentrated on this decision (see above, note 15).

[31] Court of Justice, 4 June 2015, Case C-579/13 (*P e S c. Commissie Sociale Zekerheid Breda, College van Burgemeester en Wethouders van de gemeente Amstelveen*).

[32] Court of Justice, 9 July 2015, Case C-153/14 (*Minister van Buitenlandse Zaken c. K e A*).

[33] '*Personalized proportionality assessment*', in the recent case-law of the Court of Justice is discussed by Acosta Arcarazo in *The Security of the Status of Long-Term Non-EU Residents in the EU: Some Thoughts on Case C-636/16 López Pastuzano,* in <http://eulawanalysis.blogspot.it/2017/12/the-security-of-status-of-long-term-non.html>.

the desirability of making assessments case-by-case recently found confirmation in the decision with which the Luxembourg judge, consulted for preliminary ruling by a Spanish administrative court, stated that the expulsion of a long-term resident from the territory of a Member State is legitimate «solely where he or she constitutes an actual and sufficiently serious threat to public policy or public security» (Case C-636/16, § 25)[34]. The expulsion injunction, whether an administrative penalty or the consequence of a criminal conviction, can be adopted by the Member State only after due consideration of a series of elements including – the Court of Justice points out, referring to art. 12, clause 3. dir. 2003/109/EC – links with the country of residence or absence of links with the country of origin (§ 26). In other words, verification of the accomplished integration of the immigrant in the host country constitutes a form of enhanced protection against expulsion (§§ 23-24), much like the EU provision that has for some time been in force in favour of worker citizens of the Union[35].

This is the point arrived at in the interpretation offered by the Court of Justice to the directives of 2003 and certain national regulations relating particularly to matters of integration. The fact that the interventions of the Court – above all since the Treaty of Lisbon came into force – were prompted by preliminary referrals is a (positive) sign of increased cooperation between the national and European judges, and between Member States and the European Union. The greater readiness shown by the judges in the individual Member States to involve the Luxembourg Court is a highly significant advance in consideration of the fact that the matter involved belongs to the domain of State competence[36]. Considering, moreover, that the rulings of the Court, although directly addressing the referral courts and the States parties in

---

[34] Court of Justice, 7 December 2017, Case C-636/16 (*Wilber López Pasuzano c. Delegación del Gobierno de Navarra*).

[35] Cf. *supra* § 2 and regulation (EEC) n. 1612/68.

[36] Since the Treaty of Lisbon came into force, the traditional attitude of 'reluctance' or 'timidity' on the part of the national judges (as well as the European Commission) when it comes to involving the Court of Justice seems to be on the wane in a sector that is, moreover, politically sensitive and directly associated with state sovereignty.

the judgement, also have repercussions on the legislators and judges of the other EU countries, it would not be unrealistic to envisage sometime in the future the construction of a common standard for integration of TCNs that would ensure for them legal status – as the Tampere conclusions put it – 'comparable' with and 'as close as possible' to that of the EU citizens.

6.4 Conclusions and challenges for the near future: support for more operative integration of immigrants in order to achieve effective European integration

Leaving aside definitions inspired by wishful thinking, when we approach the subject of integration certain points must be kept very much in mind. To begin with, integration is not a requisite that is verifiable *uno actu* or measurable simply with a test, but is rather an interactive and dynamic process – a work in progress that should find implementation in everyday practice, above all within the host country, and receive concrete support from the reception facilities. Moreover, it needs to be recognised that, at least in the western world, a uniform and monolithic configuration of the society, to which immigrants should conform, is no longer corresponding to the realities. In fact, not only have the migratory flows from the third countries contributed little by little to demolishing this myth, but pluralism – linguistic, cultural, ethnic and religious – is a well-established value in the immigration territories themselves, to the extent that different solutions are imaginable for management of the integration of immigrants in States that recognise the presence of historical linguistic minorities, which could feel threatened by uncontrolled entry from third countries in the respective places of residence.

Faced with the economic and demographic challenges that Europe has to address in the present situation, the EU continues to point out that immigration is a resource for the individual Member States and for the Union as a whole. The integration of regular immigrants is seen

as «a driver for economic development and social cohesion»[37] and it is «in the common interest of all the Member States»[38] as being potentially able to promote sustainable and competitive economic growth. At the same time, integration constitutes a cultural wealth provided that the policies promoting it rest on respect of differences, protection of fundamental rights and guarantee of equal treatment. European secondary legislation has for some time been focusing on these principles, and more recently the expediency of simplifying migration procedures has emerged, ensuring, for example, that TCNs with settled residence in a Member State should enjoy a common range of rights equal to that of the national citizens[39]. These objectives should be achieved with the support of more effective integration measures through greater involvement of all the levels of governance – local, national and European. In particular, acquiring a command of the language is a fundamental prerequisite for effective inclusion in the host society, and stress is therefore placed on the need to organise courses in language and civic education as well as introductory programmes both in the host country and in the place of origin to provide migrants with adequate grounding before their departure and ensure that they are informed of their rights and obligations, including the duty to observe the rules and values of the society they wish to settle in. As regards coordination and dialogue between interested parties, the EU – which has no direct powers for intervention on matters of integration – has committed to offering its support to the development of a trilateral process between migrants, host society and country of origin.

Alongside these ambitious projects, which are awaiting full validation at the practical level, it is also to be borne in mind that respect

---

[37] See the *European agenda for integration of third country nationals,* Brussels, 20 July 2011, 2.

[38] See the *Action plan for the integration of third country nationals,* Strasbourg, 7 June 2016, 2.

[39] See directive 2011/98/EU which introduces a single application procedure for issue of qualification combining residence permit and work permit. The single permit should help simplify and harmonise the regulations at present in force in the various Member States, as well as facilitating verification of the regularity of residence and employment.

of human dignity, freedom, democracy, equality, the rule of law and human rights, including the rights of persons belonging to minorities, are «values […] common to the Member States in a society in which pluralism, non-discrimination, tolerance, justice, solidarity and equality between women and men prevail» (art. 2 TEU). It is on the basis of these fundamental values of the European Union that practices for integration of regular extra-EU immigrants should be modulated and revised by the individual Member States. It is, in fact, hard to deny the abyss that lies between the proclamation of the (liberal) values of democracy and pluralism allegedly inspiring the European systems and the (illiberal) suffocation of diversity[40] occurring whenever the relevant national policies disregard the basic rules of equality and respect of human rights, adopting disproportionate provisions that obstruct effective inclusion in the host society or discriminatory forms of treatment, as in the cases that have so far been brought to the attention of the Court of Justice.

Definition of the action to be taken by states on the entry, residence and naturalisation of TCNs now offers the Member States an extraordinary and possibly unique opportunity to reflect on themselves and on categories that had seemed to be eternal, beginning with those of sovereignty-people-nation. They could thus go on to determine what key changes they might embark upon and thereby determine how to cope with the inexorable reality of mass migration. Unless, of course – this, too, would be a choice – they mean to isolate themselves behind their respective borders, sheltering behind (visible or invisible, but certainly unrealistic and anachronistic) barriers. Migration, and the policy decisions they raise for the public powers and society in general, bring pressure to bear on identities – collective, national, constitutional and cultural – that had seemed to be thoroughly consolidated if not immutable. They also force governments to measure up to the changes that have taken place so far, and to take on the responsibilities for the generations to come; they are the mirror of our conscience, revealing the

---

[40] In this connection, Orgad (2015, spec. pp. 142 ff.), discusses 'Europe's Paradox of Liberalism'.

image of what we autochthonous Europeans really are[41].

However, the massive flows of migrants and the need to rewrite, and above all implement, the policy agenda on integration represent not only a testing ground for each of the individual Member States, but also for the States as participants in that vaster assembly which is the European Union. They hold a challenge for the supranational institutions and, ultimately, for the endurance of the Union. The Treaty of Lisbon provides the rules for a common immigration policy, but leaves the states' self-determination intact on matters of integration. The principle is reaffirmed in the directives we have examined, and yet action for promotion support and monitoring by the EU organisations is recognised as admissible and desirable. For its part, the Court of Justice has opportunely contributed to the identification of certain minimum standards by which it can at least be determined '*what integration is not*'.

Given this broad picture of the realities, it is primarily the task of the Member States to decide whether to invest in the itinerary embarked upon and provide contents for the formulas in the key documents which the EU continues to draw up, or leave them at the level of empty rhetoric. Proposals have even recently been advanced for greater integration amongst the Member States, or integration at diversified rates[42], or exit from the Union, to the extent of its conjectured and avertible disintegration: this is the formidable challenge the Member States are facing at this point in their history[43]. Coherent, realistic and

---

[41] As observed Orgad (2015, p. 234), «Immigration policy is a mirror of constitutional identity. Naturalization requirements – the criteria that "they" must fulfill in order to join "us" – define "our" way of living, form of thinking, and mode of behaviour. Much can be learned about collective identities by analyzing immigration and naturalization requirements. By investigating the legal ways to become a citizen, we can learn a great deal about who we are as a people».

[42] Cf. in this respect the Communication by the European Commission, *White Paper on the future of Europe. Reflections and scenarios for the EU 27 by 2025,* Brussels, 1 March 2017, illustrating five possible scenarios for the evolution of the European Union.

[43] Some years ago it was observed by Gross (2005, p. 161), that «European integration will not be complete as long as third-country nationals resident in the territory of the Union are not regarded as an integral part of the area of freedom, security and justice».

supportive response to the complex phenomenon of migration can contribute not only to the economic revival and cultural enrichment of the countries of the European Union, but also to truly promoting 'unity in diversity', reinforcing the bonds with the common European home, and reaffirming the sharing in those values discussed above, observance of which is also required of TCNs by the Member States.

References

Acosta Arcarazo, D. (2011). *The Long-Term Residence Status as a Subsidiary Form of EU Citizenship,* Leiden-Boston: Brill.

Acosta Arcarazo, D. (2017). *The Security of the Status of Long-Term Non-EU Residents in the EU: Some Thoughts on Case C-636/16 López Pastuzano,* available at http://eulawanalysis.blogspot.it/2017/12/the-security-of-status-of-long-term-non.html.

Boelaert-Suominen, S. (2005). Non-EU Nationals and Council Directive 2003/109/EC on the Status of Third-Country Nationals Who are Long-Term Residents: Five Paces Forward and Possible Three Paces Back, *Common Market L. Rev.*, *42*, 1023.

Block, L. and Bonjour, S. (2013). Fortress Europe or Europe of Rights? The Europeanisation of Family Migration Policies in France, Germany and the Netherlands*, European Journal of Migration and Law,* 15(2), 203-224.

Carrera, S. (2009). *In search of the perfect citizen?: the intersection between integration, immigration and nationality in the EU*. Leiden-Boston: Brill.

Carrera, S. (2014). Integration of Immigrants in EU Law and Policy: Challenges to Rule of Law, Exceptions to Inclusion. In Azoulai, L. and de Vries, F. (eds.), *EU Migration Law: Legal Complexities and Political Rationales*. Oxford: Oxford University Press.

Cerrina Feroni, G. and Federico, V. (eds.) (2017). *Società multiculturali e percorsi di integrazione: Francia, Germania, Regno Unito ed Italia a confronto*. Firenze: Firenze University Press.

Cerrina Feroni, G. and Federico, V. (eds.) (2018). Strumenti, percorsi e strategie dell'integrazione nelle società multiculturali. *Napoli, ESI*, 209-241.

de Vergottini, G. (1995). Verso una nuova definizione del concetto di minoranza. *Regione e Governo locale*, 23.

Groenendijk, K. (2006). Family reunification as a right under community law, *European Journal of Migration and Law*, 8(2), 215-230.

Groenendijk, K. (2014). Recent Developments in EU Law on Migration: The Legislative Patchwork and the Court's Approach, *European Journal of Migration and Law,* 16(3), 313-335.

Gross, T. (2005). Integration of Immigrants: The Perspective of European Community Law, *European Journal of Migration and Law,* 7(2), 145-162.

Hailbronner, K. and Klarmann, T. (2016). Family Reunification Directive 2003/86/EC. In Hailbronner, K. and Thym, D. (eds.), *EU Immigration and Asylum Law. A commentary*. München: Beck.

Jesse, M. (2016). Integration measures, integration exams, and immigration control: P and S and K and A, *Common Market Law Review,* 53, 1065.

Murphy, C. (2010). The Concept of Integration in the Jurisprudence of the European Court of Human Rights, *European Journal of Migration and Law*, 12(1), 23-43.

Orgad, L. (2015). *The Cultural Defense of Nations. A Liberal Theory of Majority Rights*. Oxford: Oxford University Press.

Piergigli, V. (2013). *Integrazione linguistica e immigrazione. Approcci e tendenze nel diritto comparato europeo*, *Federalismi.it*, 1-25.

Piergigli, V. (2018). L'integrazione degli immigrati da paesi terzi nel diritto sovranazionale: limiti e potenzialità dell'Unione europea, *Rivista AIC,* 3, 1-26.

Strazzari, D. (2015). Integrazione linguistica e immigrazione in ordinamenti composti plurilingue. Tendenze dal diritto comparato europeo, *Le Regioni*, 43(4), 857-902.

Strazzari, D. (2016). Linguistic Integration Test and Federalism: A Comparative Analysis, *European Journal of Migration and Law,* 18(4), 442-466.

Thym, D. (2016a). Long Term Residents Directive 2003/109/EC. In Hailbronner, K. and Thym, D. (eds.), *EU Immigration and Asylum Law. A commentary.* München: Beck.

Thym, D. (2016b). Towards a Contextual Conception of Social Integration in EU Immigration Law: Comments on P & S and K & A, *European Journal of Migration and Law,* 18(1), 89-111.

Wiesbrock, A. (2010). *Legal Migration to the European Union.* Leiden-Boston: Brill.

Woelk, J., Guella, F. and Pelacani, G. (2016). *Modelli di disciplina dell'accoglienza nell''emergenza immigrazione': la situazione dei richiedenti asilo dal diritto internazionale a quello regionale.* Università degli Studi di Trento.

Chapter 7

Financialisation of non-financial corporations, globalization and capital accumulation

Giovanni Scarano

7.1 Introduction

One of the major movers, if not the prime mover, of the modern process of economic growth is, as everybody knows, capital accumulation, which ultimately depends on the share of social surplus devoted to investment. In capitalist economic systems, moreover, profits are the main part of social surplus on the income distribution side, and thus they are also the most prominent source of potential investment in new capital goods.

Unfortunately, in the past fifty years the investment/profits ratio has shown a declining trend (Stockhammer 2006). Since the late 1980s, however, non-financial corporations, while reducing their accumulation of capital goods, have progressively increased their financial investment (Stockhammer 2004).

Thus, it becomes important to understand whether there is a connection between decreasing real investment and growing financialisation by non-financial corporations. Unfortunately, investigation into this potential connection is made very difficult by the present weakness of investment theory.

In fact, even though trends over time and volatility of aggregate investment in fixed capital should be central to understanding aggregate fluctuations in economic activities, investment theory has been traditionally weak on these matters. Moreover, in the early 1960s Jorgenson had already noted there was a great gap between economic

theory and econometric practice in the literature on business investment in fixed capital (Jorgenson 1963). Econometric models, indeed, are often a simply attempt to find empirical correlations between aggregate investment and other economic variables, sometimes without any rigor in the underlying theoretical foundations.

Obviously, as physics shows, mass phenomena can usually reveal rules of aggregate functioning that cannot be easily reduced to atomistic behaviours. However, physicists can discover these rules with empirical experiments in the lab. Unfortunately, things in economics are more complex, because economic phenomena cannot be reproduced in labs and empirical rules can be discovered only by measuring ex-post statistical correlations, which may often be interpreted in various ways and by means of different causal relationships. Moreover, some variables, like interest rates and investment decisions, can act on each other through different mechanisms, producing contradictory net effects.

For a long time economists have been trying to explain the aggregate investment demand empirically simply by using scale variables. The most famous attempt of this kind, as is well known, is the *accelerator model* by Clark (1917), built on the basis of the relation existing between the first differences of a simple fixed coefficient inverse production function. However, despite the empirical success of this model, since the late 1960s there have been many attempts to introduce the cost of capital as explanatory variable in econometric models specified starting from the optimisation problem of a perfectly competitive firm (Hall and Jorgenson 1967).

Meanwhile, other economists highlighted the predominance of liquidity variables over the interest rate for short-run investment, introducing elements of portfolio choices into the investment theory (Tinbergen 1939; Meyer and Kuh 1957). Subsequently, according to the Tobin approach, investment has been recognised as an increasing function of q, the ratio of the financial value of the firm to the market cost of its capital goods, which is, in fact, closely connected to the companies' rate of profit (Brainard and Tobin 1968; Tobin 1969; Hayashi 1982; Caballero 1999).

Yet very few models highlight the fact that investment in capital

goods depends mainly on corporate savings decisions, which are closely connected to the features of corporate governance and the forms of competition, and strategic competitive behaviours. In most advanced economies, in fact, retained earnings constitute the dominant source of finance (Mayer 1988; Tirole 2006).

Most models of investment decisions utilised in macroeconomic models, in fact, take free or perfect competition as explicit or implicit assumption. However, the oligopolistic structure of most real markets leads to corporate strategic behaviours that can produce very different results. Strategic decisions, connected with agency problems, can play a major role in producing financialisation and timing the rhythms of real investment.

In the following pages, after a brief survey of the role played by the interest rate, in all its different aspects, within investment theory, the paper deals with some contributions, both mainstream and heterodox, that analyse the effects of corporate governance and strategic behaviours on portfolio management and investment decisions in big corporations, seeking to determine how these effects might play a major role in producing growing liquidity holdings and financialisation, and how they can be influenced by the new opportunities created by free movements of capital and the economic dynamics of the emerging countries. The main objective is to understand whether these models can explain the tendency to place growing shares of social surplus in speculative financial channels, thereby contributing to long-term real stagnation of at least one part of the world.

7.2 Investment decisions theories and the interest rate

7.2.1 The foundations of Neoclassical investment theory

Neoclassical economists are accustomed to bringing together under the heading of real investment decision criteria all the rational choices that involve a trade-off between the present and the future (Hirshleifer 2008), with apparently no distinction between consumers and firms' behaviours. In this way, investment is eventually reduced to the problem of optimizing

consumption patterns over time. This obviously presupposes that all the economic agents can be reduced to consumers or their representatives, essentially because households are assumed to be the end owners of all the production factors and firms are viewed as mere agents of them.

Of course, in this schematised picture of the economy, the relation between investment, simply seen as the amount of reduction in current consumption, and the market interest rate has to be governed by the consumers' structure of intertemporal preferences. Thus, assuming a universal psychological preference for the present on the part of consumers (households), investment will be a decreasing function of the market interest rate, which is the average measure of this preference.

In the simplest models of investment decisions of this kind, it is usually assumed that, at every point in time, a firm has only two alternatives in utilising its net revenues. It can disburse them to its owners as income or invest them to produce a greater amount of net revenues at some future date. Thus, by reducing owners' current income at time t and investing its retained earnings, the firm can increase future income of the owners themselves. It is in this way that firms are no more than the simple agents for their owners' intertemporal consumption choices (Branson 1979).

7.2.2 Some alternative models

However, in the economic literature there are a number of models, both neoclassical and heterodox, in which the real investment decision-makers are firms that have no shareholders to satisfy (Romer 2012). In this case, the connection with intertemporal utility equilibrium is not so immediate. In fact, investment decisions follow from entrepreneurs' choices of techniques finalised to maximise their profit or, symmetrically, to minimise their production costs. The intertemporal utility equilibrium connected to them can still be demonstrated in the context of a general equilibrium scheme, thanks to the role played by market prices and interpretations of them in terms of utility, but the entrepreneurs' choices between present and future cannot be immediately described in terms of utility, because the latter is not the direct objective to be maximised by entrepreneurs.

According to Irving Fisher's neoclassical theory of capital, for example, firms should choose their production plans so as to maximize utility over time, and this, under certain conditions, would lead to maximization of the net worth of the enterprise as the objective for optimal capital accumulation (Jorgenson 1963). However, this idea of firms maximising utility is the trivial result of the idea that market prices, at their equilibrium level, are measuring social marginal utilities. Thus, maximising the difference between revenues and costs – i.e. profit – equals maximising social utility. This final result, however, is no intentional outcome pursued by entrepreneurs, but only an impersonal effect achieved by the market's invisible hand.

Nevertheless, in this theoretical perspective capital accumulation is the result of providing inputs – i.e. capital services – to the productive process, to maximise profit in accordance with given production functions. Thus, it is the outcome of changes in techniques made by firms in response to changes in external decision parameters.

In neoclassical models in which the investment decision is the direct consequence of the entrepreneurs' choice of techniques, therefore, the interest rate is simply the price of utilising capital. Thus, its changes determine changes in the optimal technique chosen. If the interest rate goes down, then the choice moves towards more capital-intensive techniques, increasing the stock of capital desired and creating investment demand.

On this front, the main difference between neoclassical and neo-Ricardian models lies in attaining or not attaining *well-behaved* production functions that can guarantee optimal general equilibria. The Cambridge capital controversy proved that there was no regular relation between the use of jelly capital – i.e. financial capital – and labour productivity.

In neo-Ricardian models, changes in the interest rate have effects on the choice of techniques by means of changes in income distribution that produce changes in the long-period prices. The changes of techniques that produce capital accumulation are still the consequence of minimising production costs, yet the connection between the interest rate and the cost of utilising capital is more indirect (Pasinetti 1966).

In both these kinds of models, however, real investment, as accumulation of capital, is viewed as a passive behaviour, because it is only the result of firms adapting techniques to exogenous changes in market or social parameters.

Thus, real investment as a result of changes in the choice of techniques shows capital accumulation behaviour as a mere passive effect, led both by changes in technology and consumption preferences, as well as the relative availability of production factors in neoclassical models, or by changes in distributive variables and exogenous aggregate demand in neo-Ricardian, Keynesian and post-Keynesian models (Crotty 1992; Bonifati 2016). However, this theoretical result is the consequence of the assumption of perfect or free competition, under which firms are aware they cannot modify their decision-making parameters.

7.2.3 Strategic behaviours and Schumpeter's hypotesises

Under oligopolistic competition hypothesises, however, things could be very different. In this case, firms could be following strategic behaviours aiming to gain key positions in the market for successive competition battles, such as a growing share of the total supply in specific production sectors. From this point of view, for example, firms could forgo present normal profit for greater expected future extra-profit, based on greater monopoly power. These kinds of strategic behaviour could, therefore, increase aggregate real investment as compared with equilibrium saving decisions, producing disequilibria in the financial markets and pressures on interest rates, as was probably the case in Europe during the 1960s (Lamfalussy 1968).

A close connection between real investment and strategic behaviours clearly emerges from Schumpeter's analysis. According to his approach, in fact, capital accumulation is the consequence of 'innovation' – i.e. the result of active strategies of reducing own production costs as compared with those of their competitors on the part of 'entrepreneurs', which are creative managers and capitalists or genial inventors well financed by forward-looking banks and other financial intermediaries. Thus, innovation appears as a series of creative actions by 'enterprises' – i.e. innovative firms – intrinsically embodied in new

investment goods, which push the economy out of its stationary equilibria, inducing adjustments in the choice of techniques, and then more investment, also in the non-innovative firms (Schumpeter 1939).

From this point of view, capital accumulation is still connected with cost-minimising behaviours, but in a context of active and strategic competition for the conquest of dominant or monopolistic positions in the market, in order to make the highest possible extra-profit. There is no role played by the interest rate, apart from the risk premium required by financial operators, which could in some case discourage the entrepreneurs from seeking higher extra-profits.

7.2.4 Interest rates and financial investment

In neoclassical economics, however, the interest rate, at its equilibrium level, is also a measure of the marginal productivity of capital – i.e. the marginal rate of profit. Thus, investment quantity depends on capital profitability too, obviously compared with the available quantity of savings supplied, which in turn depends on the intertemporal consumer preferences.

In recent years, after the financial crisis of 2007-2008, the idea is returning that the central banks have to influence market interest rates to drive them toward levels consistent with their best estimation of the equilibrium real interest rate – i.e. the Wicksellian interest rate (Bernanke 2015). Thus, according to mainstream approach, the objective conditions of the economic systems, and in particular return on capital invested, are again what ultimately determine the real rate of return for savers and financial investors.

This draws attention to the role of the interest rates as rate of return on financial assets. From this point of view, the term structure of interest rates is very important in determining portfolio choices, and then the alternative between financial and real investments.

In Keynes's theory, investment is the driving force of income levels and their fluctuations (Minsky 1975), and this, in turn, depends on the propensity to hoard, and therefore, on the monetary policy and expectations concerning yield of capital assets (Keynes 1937). Thus, Keynes's theory of investment connects the fluctuations of real

investment to variables which are determined in the financial markets and primarily to the rate of interest (Minsky 1975). The monetary theory of the interest rate proposed by Keynes, in fact, underlined its meaning as required rate of return on financial investment. From this point of view, if speculators are not viewed as intrinsically different from other types of investors, the interest rate may be in competition with the expected rate of return on real investment – i.e. the rate of profit – within portfolio choices. New savings can be invested in a financial or real way depending on comparison between the interest rate on financial capital and the rate of profit with the risk premium deducted.

Thus, Keynes argued that even though aggregate saving depends essentially upon aggregate income level, the rate of interest plays a secondary role by influencing how economic agents allocate their savings among different potential shares of their portfolios.

7.2.5 Marx's conception of investment

Marx's conception of investment, as capital accumulation, is, then, quite another issue, but very relevant to our task of analysis. According to Marx, in fact, capital accumulation is the mere result of the self-valorisation process of capital. If capital value is, by nature, in search of additional value, then the surplus value, which is of the same nature as the original capital value, will in turn be searching for its additional value. Thus, capital, by its very nature, organises a continuum process of self-valorisation that increases its value, until an overproduction crisis temporarily arrests it.

This conception, which might sound somewhat metaphysical to those who conceive of real investment only as a way to adapt production techniques to new market conditions, in fact fully complies with the 'pursuit of shareholder value' principle, put in place by the managers of big corporations, who certainly cannot be suspected of being Marxists.

In Marx's works, moreover, there is no reference to the fact that the rate of profit can be a subjective incentive for accumulation of capital or play a role in changing aggregate investment decisions. According to him, the average rate of profit plays a role only in capitalist competition, stimulating movements of *money capital* – i.e. financial capital – from

one productive branch to another and determining a redistribution of total surplus value proportionally to the *money capital* invested in capital goods and anticipated wages. Thus, different rates of profit in different productive branches can reduce real investment in those with the lowest rate, but increase it in those with the highest rate. Obviously, analogous movements of *money capital* can also occur between financial branches or between real branches and financial branches. And Marx also thought the rising mass of profits that exceeds the possibilities for new industrial investment in periods of prosperity first accumulates as a huge mass of credit and then ends up by taking speculative paths, because the production process usually appears only as «a necessary evil for the purpose of money-making», and at the first opportunity, 'fits of giddiness' are unleashed in which capitalists «try to accomplish the money-making without the mediation of the production process» (Marx, [1885]1992, p. 137).

7.2.6 Some intermediate conclusions

Thus, even though the interest rate, in itself, can play a minor role in determining real investment, if corporate managers act as financial investors, then the term structure of interest rates could play a major role in determining investment in capital goods as a share of companies' portfolios. At this point, the way managers make their portfolio choices becomes crucial. Yet above all, what happens to the relation between investment decisions and the term structure of interest rates if corporations are not a veil but decision-makers guided by objectives other than utility maximization?

This issue, of course, is closely connected to corporate governance problems.

In the first three post-war decades, the role of shareholders in corporations was severely limited by heavily restrictive financial regulation and capital flows control, which were the political reactions to the financial and real crisis of the 1930s.

In the 1950s and 1960s, according to Baran and Sweezy (1966), giant corporations usually aimed at financial independence through retained earnings. They were able to borrow from financial institutions

and markets, but were not normally forced to act so and could avoid subjection to control by financial corporations and outside shareholders. In this kind of corporations, managers were a self-perpetuating group that identified itself with the corporation and its fate. The board of directors and the chief executive officers were 'organization men' and the control rested securely in their hands. Their major objectives were the corporation market share and its strategic positions in the market.

However, this situation has been changing since the late 1970s, through the progressive erosion of financial regulation by means of the invention of new financial instruments, such as junk bonds and other high-risk and high-return securities. Moreover, up to 1982 the *Securities and Exchange Commission* (SEC) could counteract massive stock repurchases as illegal attempts to manipulate stock prices by the companies. Since the end of 1982, instead, during the deregulation onset of the neoliberal phase, the SEC has partially liberalised stock repurchases, provided that they be less than 25% of the average daily trading volume over the previous four weeks and the buybacks be carried out at neither the beginning nor the end of the trading day (Lazonick 2013).

By means of this financial deregulation, the financial markets have progressively exerted increasing pressure on non-financial corporations (NFCs), by means of hostile takeovers first, and then with the 'shareholder revolution', characterised by a growing presence of institutional investors within their shareholding (Lowenstein 2004; Orhangazi 2008). French regulationists have been emphasising corporate governance since the 1970s, because the pursuit of 'shareholder value' is closely associated with the short-termism of non-financial corporations (Boyer 2000; Grahl and Teague, 2000; Aglietta 2000; Aglietta and Breton 2001), and Lazonick and O'Sullivan (2000) have perceptively shown the connections between shareholder value and company downsizing throughout the neoliberal phase of capitalist development (Lapavitsas 2011).

According to Stockhammer, the 'shareholder revolution' is one of the main features of the present neoliberal era, which has produced radical changes in corporate behaviour in the name of creating 'shareholder value'. According to him, this revolution has been the

consequence of the financial liberalization and the emergence of very liquid share markets in the 1980s and 1990s, together with the successive rise in shareholders' capability to influence public company managers by means of the creation of 'a market for corporate control'. The managements of large corporations, in fact, would have committed themselves to increasingly producing shareholder value because of the expanded possibilities for financial investors to use the capital market to estimate and compare performance of their corporations and to discipline them with the threat of hostile takeovers. In this new context, the managers of large corporations could easily be replaced by shareholders if corporate performance proved inadequate in creating value for them (Stockhammer 2006).

Thus, the threat of growing control by large financial intermediaries in public companies could be an incentive for managers to change their investment behaviours, increasingly orienting them towards short-term profit investment and discouraging 'extravagant' or long-term strategic investments. This change is also supported by an incentives system for managers that closely connects their wages to the company stock prices, encouraging financial operations like share repurchasing by means of retained earnings, which would thus be subtracted from investment in new capital-goods and technologies.

However, this tendency to produce an increasing shareholder value could not only be the result of new forms of corporate governance and new financial intermediaries, but rather the traditional way to maximise the equity capital self-valorisation in a different competition environment and given new financial investment opportunities. This puts the emphasis on other transformations of the capitalist system in its neoliberal phase, which have been in part gathered under the label of financialisation.

7.3 Financialisation

Today the term financialization is usually used to refer to three different, even though interconnected, phenomena. The first is the reduction of

reliance on bank loans by large non-financial corporations and their growing autonomous ability to raise funds in financial markets. The second is the expansion of banks' mediating activities in financial markets and their tendency to lend mainly to households. The third is the increasing involvement of households in the financial markets, as both debtors and asset holders (Orhangazi 2008; Lapavitsas 2011).

Thus, at the macroeconomic level, financialisation in practice simply becomes synonymous with the expanding financial sector within the economic system.

However, the concept can have a deeper and more interesting meaning at the firm level, where it can highlight the changes in the behaviours of the managers of non-financial corporations and their new relations with the financial markets.

An important part of the literature on the relation between financialization and investment draws on Keynesian and Minskian approaches, which emphasize the importance of financial factors in corporate investment (Eichner and Kregel 1975; Minsky 1986; Skott 1989; Crotty 1990, 1992; Lavoie 2014; Davis 2017). In the last three decades, however, a new kind of phenomenon has powerfully been emerging. Mainly in the US, but also in continental Europe, non-financial corporations have been increasingly investing in financial assets and creating own financial subsidiaries, deriving increasing shares of their income from this kind of pure financial activities (Stockhammer 2004; Orhangazi 2008). In the same period, NFCs have increased transfers of earnings to the financial markets in the forms of interest payments, dividend payments and, mainly, stock buybacks. Thus, according to some analysts, these transformations, in close synergy with the previously examined evolution in corporate governance, have produced radical changes in the objectives of top managements, favouring an increasing propensity to substitute real investment with short-term financial investment in the process of corporate investment decision-making.

Moreover, according to Sawyer (2017), financialisation has changed the relations between the financial sector and the real sector because the passage of ownership of non-financial corporations into the hands of financial corporations has emphasised the 'pursuit of

shareholder value'. This could again connect the effects of financialisation to those previously examined with regard to corporate governance changes. In pursuit of higher corporate performance, defined as meeting financial markets' expectations for quarterly earnings per share, American companies have conducted great stock repurchases to increase their own corporations' stock prices (Lazonick 2013). In this way, trillions of dollars have been subtracted from innovation and job creation over more than three decades.

Thus, a financialized mode of corporate resource allocation could have been produced through the principle of maximizing shareholder value, and corporate executives have been incentivised in this direction by their stock-based compensation.

Financial control, however, has traditionally been viewed also as a particular organization controlling model by top managers (Fligstein 1990). Thus, the 'pursuit of shareholder value' could simply be the effect of traditional maximising self-valorisation of the capital of owners with a controlling shareholding, who utilise mass-shareholders and their financial intermediaries as less-secured claimholders, with a minor role played by the 'shareholder revolution'.

As pointed out above, since the late 1980s non-financial corporations, while reducing their accumulation of capital goods, have progressively increased their financial investment (Stockhammer 2004). The decreasing fixed capital formation rate in many countries, both developed and developing, has probably been the effect of growing uncertainty, risk and volatility on the real investment performance (Demir 2009). Thus, the relation between fixed investment, uncertainty, increasing integration of international capital markets, the widening gap between real and financial sector transactions and corporate portfolio choice seems to be a very important factor. Successively to financial liberalization, in fact, NFCs have been facing portfolio choice problems in their investment decisions between fixed and financial assets and increasing availability of alternative financial investments can channel NFCs' retained earnings to short-term financial investments instead of long-term fixed capital formation, and thus contribute to deindustrialisation.

If companies are viewed as common financial investors, Tobin's portfolio choice theory points out the substitutability of real and financial assets also in their portfolios, depending on the respective rates of returns. Increasing risk and uncertainty, combined with capital market imperfections, higher real interest rates and increasing rates of return in the financial markets, may encourage short-term financial investments over long-term real investments. In fact, according to Tornell (1990), an uncertain environment can usually encourage NFCs to invest in more liquid assets, which at the same time offer comparable or higher rates of return, at the expense of their real fixed assets.

In the first decade of the 2000s a growing part of the literature on financialisation focused on increasing rates of return on financial capital over those on fixed capital and at the same time increasing acquisition of short-term financial assets by NFCs in high-income OECD countries, providing empirical evidence of a structural change in their portfolio decisions (Stockhammer 2004; Crotty 2005; Dumenil and Levy 2005; Epstein and Jayadev 2005). Other studies suggest that also NFCs in developing countries take into account alternative financial investment opportunities when making their decisions on physical investment (Demir 2009). All these empirical works, moreover, have demonstrated that the increasing financial activity of NFCs has had a negative effect on their real investment (Stockhammer 2004; Orhangazi 2008; Demir 2009).

Thus, financialisation constitutes a radical change in corporate management behaviour that can produce major changes in investment demand for single firms. However, even though financial investment can be alternative to physical investment at the level of the single corporation, at the macroeconomic level the phenomenon shows a clear fallacy of composition. Indeed, financial investment only transfers liquidity from one agent to another one. Thus, it could transfer financial resources from firms with bad investment opportunities to others with better profitability prospects. From this point of view, it should have no macroeconomic effect, apart from increasing efficiency (Tobin 1997).

There are only two possible macroeconomic effects, global and local. The former is substantially reducible to hoarding or a speculative

demand for money. The latter, instead, is closely connected to capital transfers between different countries. In both cases, the macroeconomic problem does not emerge from finacialization in itself, but from its relations with other features of economic systems. From a Marxist point of view, it is real accumulation that determines the parameters on which finance runs, even though the latter can cause counter effects on the former (Itoh and Lapavitsas 1999).

7.4 Liquidity holding

Hoarding money and very liquid assets by NFCs, as in Marx's theory of crisis, could therefore be once again the first mover of decreasing aggregate investment that actually lies behind the multiform appearances of financialisation. And this phenomenon, of course, is closely connected to the role of money in business fluctuations.

  The cash balance policy of enterprises was already attracting attention in the late 1920s, immediately before the Wall Street crash of 1929. At that time corporations utilised part of their previous accumulated cash balances to buy securities in the stock exchange market in order to profit from the high call loan rate. So they were accused of contributing to the boom on the stock exchange, but also of feeding the financial markets at the expense of the other markets (Scarano 2016).

  In the 1930s, corporate cash holdings were again an object of heated discussion because, according to some economists, corporations held 'idle' cash (cash balances in excess of current operation needs), contributing to the stagnation of the economic system.

  According to a seminal study by Lutz (1945), in the years of the Great Depression the ratio of cash plus marketable securities to payments rose sharply. This ratio diminished from 1933 to 1937, and then rose again in 1938, because of a new business contraction. According to Lutz, the ratio of cash plus marketable securities to payments showed a rise in periods of business contraction and a fall in periods of expansion, and the movement of these free liquid funds showed an inverse correlation with the profit rate of the large corporations.

Thus, according to him, in the 1930s the large manufacturing corporations held 'free' liquid funds because they were hoarding money. The major reason for this was that the big corporations were largely independent of bank credit, so that their bank debts did not absorb the 'surplus' cash originated through liquidation of inventories and receivables during the depression. This behaviour suggested that in a system in which companies are not largely indebted to the banks, business fluctuations may be reflected in a change in the velocity of circulation of firms' deposits.

According to Lutz, during the period of depression in the thirties, medium-sized and small corporations had a small increase in their cash holdings, because great losses and the retirement of bank debt absorbed the liquid funds that would otherwise have been accumulated. Thus in the thirties, 'hoarding' was mainly a matter of large manufacturing corporations.

The corporations' cash balances were fed by partial liquidation of inventories and receivables and by the postponement of replacements for fixed equipment. Most of the additional cash was paid out mainly through distribution of dividends.

Coming back to the present, since 2002 the gross saving in the OECD corporations has been progressively exceeding their fixed investment (OECD 2007; André *et al.* 2007; Scarano 2015). Only one third of this increase in undistributed profits was generated by the non-financial sector, but this sector contributed to the increase in corporate net lending more than the financial sector.

The large-scale expansion of corporate net lending in the non-financial sector has been interpreted as partly due to the cyclical downturn since 2001 (OECD 2007). Some econometric works, in fact, suggest a significant influence of the business cycle on the corporate net lending between 2001 and 2005 (André *et al.* 2007).

However, the rise in net lending is the result of two different tendencies: falling corporate investment and increasing corporate saving share. The weakness in corporate investment, compared with GDP, can usually be largely the consequence of the business cycle, but other causes, more structural in nature and thus longer-lasting, have been detected in

the progressive decline of the relative price of capital goods, in lower depreciation rates, in lower growth trends and increasing net foreign investment abroad.

If growing corporate cash balances might seem relatively trivial after the financial crisis of 2007-2009, its beginning seven years before the crisis was a warning signal that, in 2012, led *The Economist* to write about 'hoard instinct' and 'dead money' (The Economist, 2012).

Corporations usually utilise their gross savings for depreciation, new investment, acquisitions, paying off debts and share repurchases. The change in the cash balance should normally be just the residual after spending. However, they can also hold liquid balances for precautionary, speculative and transactional reasons. The precautionary motive obviously prevails when they fear unforeseen fluctuations.

After the financial crisis, companies were certainly keen to accumulate more substantial cash balances to face up to the credit crunch. But most of companies were waiting to invest and make acquisitions because of uncertainty following the crisis. The beginning of the phenomenon, as from the early 2000s, suggests that uncertainty in the non-financial sectors really came to dominate the global economic scene as from the burst of the 90s bubble.

7.5 The effects of globalization

As everybody knows, there are a number of different ways to define globalization, each of which underlines different aspects of a progressive worldwide integration process between people, companies, and governments. However, here we prefer to confine our attention to its major economic features, which can be summarised as free trade improvement and a progressive worldwide liberalization of the movements of labour and capital.

Truly, so far, labour movements have only been liberalised on a very limited regional basis. Just as free trade has been only partially implemented under the umbrella of the WTO, with many surviving, even though minimised, tariff regimes, and countless non-tariff barriers.

In fact, the main successful economic field of globalization seems to be free movements of capital, yet today capital account liberalization is the theoretical field where economics largely fails in explaining actual events in the real world.

According to neoclassical theory, free capital flows should only be a form of intertemporal trade and then their functioning rules should be no different from those of free trade. Thus, free flows of external capital should contribute to smoothing consumption and production paths, improving social welfare. However, in the real world the result seems to be the exact opposite. Free movements of short-term capital, such as portfolio flows and short-term bank loans, have so far been related to a long series of serious economic and financial crises because of their volatility and exposure to surges in and sudden withdrawals from the financial markets.

Thus, successively to economic and financial crises in Asia, Latin America and Russia in the late 1990s, many economists underlined the possible dangerous effects of these kinds of capital movements for developing countries. Instead, long-term capital flows, such as FDI, were usually regarded as more positive for the long-term economic growth of developing countries, because they are generally more stable and can improve their production capacity and technology (Stiglitz 2000). Thus, the economic literature analysing the effects of liberalization of capital flows on the developing countries usually highlights the difference between short-term and long-term flows.

However, free movements of capital can produce significant effects on the developed economies, too. Much less analysis has been dedicated to these effects, but they can play a major role in producing the present tendency to stagnation in this kind of economies, and they are closely connected with another major phenomenon of our time: financialisation by non-financial corporations, which can greatly contribute to reducing their real investment in the developed countries, contributing to decreasing their growth rate and increasing their unemployment rate. In this context, however, the distinction between short-term and long-term capital flows can be less evident and significant.

Free movements of capital, moreover, can play a major role in financialisation of NFC from two different points of view.

If real investment depends on the term structure of interest rates over the full range of financial and real investment opportunities, then real investments in the developed countries also depend on the differential between their rates of returns and the rates of returns on real investments in the developing or emerging countries. However, this differential does not only act by means of FDI, but also by means of the possibility of financial investment in foreign securities, associated with real investment in foreign countries. Financial globalization, multiplying the potential range of financial instruments available to big corporations' portfolios and creating new ways to indirectly access the high profits produced in the emergent markets, can play a major role in changing the portfolio composition. Moreover, the managers of 'financialised non-financial corporations' can decide to substitute direct national real investments with financial investments in foreign corporations, thus also obtaining a greater liquidity for their portfolios.

However, this kind of investment does not necessarily have to go through the acquisition of equity, but can also be made by acquiring debt securities. Thus, decreasing capital controls can influence the very structure of countries' external liabilities, leading to substitution of FDI with equity inflows and external debt, which can, in turn, have a sizeable impact on the financial stability of debtor countries, significantly raising currency mismatch and making them more vulnerable to financial contagion (OECD 2012b).

Furthermore, financial investments by non-financial corporations are usually very different from the traditional forms of takeover and corporate holding because their profitability depends not only on the ratio between profits and invested capital, but also on the terms of capitalisation of the expected future profits realised through the financial markets. Thus, the growing liquidity of non-financial corporations' portfolios can contribute to heightening the usual volatility of the rates of return on financial assets as well as the vulnerability to contagion-induced financial shocks (OECD 2012a). Moreover, countries with a large financial sector have a riskier financial account structure, compared,

for instance, with commodity-exporting countries, which show a safer financial account structure.

All this obviously increases the overall uncertainty of financial investment profitability itself. And this growing uncertainty, in turn, leads to a greater tendency to money hoarding by non-financial corporations, which reduces the mass of real investment directly or indirectly financed by them.

7.6 Conclusions

Thus, ultimately, even though the interest rate can play a minor role in explaining real investment, in a world in which corporation managers behave as financial investors, the term structure of interest rates can play a major role at the firm level in determining investment in capital goods as a share of companies' portfolios. This structure will obviously depend on the available financial alternatives and the risk premiums that prevail in their markets.

However, the 'risk premiums' in imperfect markets are not necessarily the statistic measure of the riskiness of loans and contributions of capital, as the mainstream theory assumes, but can rather be the result of the power relations between financial capital and industrial capital, as classical economists and Marx himself thought. Thus, the average rate of return on financial assets can be the measure of the comparative profitability in utilising capital in a financial or productive way.

From this point of view, the major variable in explaining investment decisions is not the interest rate on money markets, strictly controlled by central banks, but rather the average 'risk premium' as a measure of the average profitability in financed real activities. And this measure is closely correlated with Tobin's q.

Of course, at the macroeconomic level, in closed economies, the average profitability can only have effects in capital movements from one sector to another, in direct or financial form, but cannot explain the absolute level of real investment. In open economies, however, the

differences in average profitability between different countries can matter, reducing capital sources for real investment in one country by means of capital transfers, in direct or financial forms, towards other countries.

Moreover, in portfolio choices by corporations, beside the rates of return, the liquidity degree of the assets can also be a very important determinant, in close connection with business fluctuations. Thus, portfolio choices by corporations also mainly depend on the uncertainty degree of their economic environment, which can induce hoarding phenomena that are, ultimately, the real prime mover of decreasing aggregate investment.

References

Aglietta, M. (2000). Shareholder value and corporate governance: some tricky questions, *Economy and Society*, 29(1), 146-59.

Aglietta, M. and Breton, R. (2001). Financial systems, corporate control and capital accumulation, *Economy and Society*, 30(4), 433-66.

André, C., Guichard, S., Kennedy, M. and Turner, D. (2007). Corporate Net Lending: A Review of Recent Trends, *OECD Economics Department Working Papers*, 583.

Baran, P. and Sweezy, P.M. (1966). *Monopoly Capital. An Essay on the American Economic and Social Order*, New York: Monthly Review Press.

Bernanke, B.S. (2015). Why are interest rates so low?, available at http://www.brookings.edu/blogs/benbernanke/posts/2015/03/30 -why-interest-rates-so-low.

Bonifati, G. (2016). Investimenti, consumi e occupazione. Capacità produttiva, domanda effettiva e distribuzione del reddito nel lungo periodo, *Economia & Lavoro*, 1, 77-107.

Boyer, R. (2000). Is a finance-led growth regime a viable alternative to Fordism? A preliminary analysis, *Economy and Society*, 29(1), 111-45.

Brainard, W.C. and Tobin, J. (1968). Pitfalls in financial model building, *American Economic Review*, 58(2), 99-122.

Branson, W.H. (1979). *Macroeconomic Theory and Policy*, New York: Harper and Row.

Caballero, R.J. (1999). Aggregate investment. In Taylor J.B. and Woodford, M. (eds.), *Handbook of Macroeconomics*, vol. 1B. Amsterdam: Elsevier, 813-862.

Clark, J.M. (1917). Business acceleration and the law of demand: a technical factor in economic cycles, *Journal of Political Economy*, 25(3), 217-235.

Crotty, J. (1990). Owner-manager conflict and financial theories of investment instability: a critical assessment of Keynes, Tobin and Minsky, *Journal of Post Keynesian Economics*, 12(4), 519-542.

Crotty, J. (1992). Neoclassical and Keynesian Approaches to the Theory of Investment, *Journal of Post Keynesian Economics*, 14(4), 483-497.

Crotty, J. (2005). The neoliberal paradox: the impact of destructive product market competition and 'modern' financial markets on nonfinancial corporation performance in the neoliberal era. In Epstein, G. (eds.), *Financialization and the World Economy*. Edward Elgar, 77-110.

Davis, L.E. (2017). Financialization and investment: a survey of the empirical literature, *Journal of Economic Surveys*, 31(5), 1-27.

Demir, F. (2009). Financial liberalization, private investment and portfolio choice: Financialization of real sectors in emerging markets, *Journal of Development Economics*, 88(2), 314-324.

Dumenil, G. and Levy, D. (2005). Costs and benefits of neoliberalism: a class analysis. In Epstein, G. (eds.), *Financialization and the World Economy*. Edward Elgar, 17-45.

Eichner, A. and Kregel, J. (1975). An essay on post-Keynesian theory: a new paradigm in economics, *Journal of Economic Literature*, 13(4), 1293-1314.

Epstein, G.A. and Jayadev, A. (2005). The rise of rentier incomes in OECD countries. In Epstein, G. (eds.), *Financialization and the World Economy*. Edward Elgar, 46-74.

Fligstein, N. (1990). *The Transformation of Corporate Control*. Cambridge (Mass.): Harvard University Press.

Grahl, J. and Teague, P. (2000). The Régulation School, the

employment relation and financialization, *Economy and Society*, 29(1), 160-178.

Hall, R.E. and Jorgenson D.W. (1967). Tax policy and investment behaviour, *American Economic Review*, 57(3), 391-414.

Hayashi, F. (1982). Tobin's Marginal q and Average q: A Neoclassical Interpretation, *Econometrica*, 50(1), 213-224.

Hirshleifer, J. (2008). Investment decision criteria, *The New Palgrave Dictionary of Economics*, vol. 4. New York: Palgrave Macmillan, 553-557.

Itoh, M. and Lapavitsas, C. (1999). *Political Economy of Money and Finance*, London: Palgrave MacMillan.

Jorgenson, D.W. (1963). Capital Theory and investment behaviour, *American Economic Review*, 53(2), 247-259.

Keynes, J.M. (1937). The General Theory of Employment, *Quarterly Journal of Economics*, 51(2), 209-23.

Lamfalussy, A. (1968). *Les marchés financiers en Europe*. Paris: Presses Universitaires de France.

Lapavitsas, C. (2011). Theorizing Financialization, *Work, employment and society*, 25(4), 611-626.

Lavoie, M. (2014). *Post-Keynesian Economics: New Foundations*. Northampton (MA): Edward Elgar.

Lazonick, W. and O'Sullivan, M. (2000). Maximizing shareholder value: a new ideology for corporate governance, *Economy and Society*, 29(1), 13-35.

Lazonick, W. (2013). The Financialization of the U.S. Corporation: What Has Been Lost, and How It Can Be Regained, *Seattle University Law Review*, 36(2), 857-909.

Levine, R. (2003). More on Finance and Growth: More Finance, More Growth?, *Federal Reserve Bank of St. Louis Review*, 85(4), 31-46.

Lowenstein, R. (2004). *Origins of the Crash: The Great Bubble and Its Undoing*. Penguin Books.

Lutz, F.A. (1945). *Corporate Cash Balances, 1914-43: Manufacturing and Trade*, NBER Books.

Marx, K. ([1885]1992). *Capital. A Critique of Political Economy. Volume Two*, London: Penguin Books.

Mayer, C. (1988). New issues in corporate finance, *European Economic Review*, 32(5), 1167-1189.

Meyer, J.R. and Kuh, E. (1957). *The investment Decision: An Empirical Study*. Cambridge (Mass.): Harvard University Press.

Minsky, H.P. (1975). *John Maynard Keynes*. New York: McGraw-Hill.

Minsky, H.P. (1986). *Stabilizing an Unstable Economy*. New Haven: Yale University Press.

OECD (2007). OECD *Economic Outlook*, 82.

OECD (2012a). International Capital Mobility: Which Structural Policies Reduce Financial Fragility?, *OECD Economic Policy Papers*, 2.

OECD (2012b). International Capital Mobility: Structural Policies to Reduce Financial Fragility?, *OECD Economics Department Policy Notes*, 13.

Orhangazi, Ö. (2008). Financialisation and Capital Accumulation in the Non-Financial Corporate Sector: A Theoretical and Empirical Investigation on the US Economy: 1973-2003, *Cambridge Journal of Economics*, 32(6), 863-886.

Pasinetti, L. (1966). Changes in the rate of Profit and Switches of Techniques, *Quarterly Journal of Economics*, 80(4), 503-517.

Romer, D. (2012). *Advanced Macroeconomics*. New York: McGraw-Hill-Irwin.

Sawyer, M. (2017). Financialization and Economic and Social Performance. In Gemzik-Salwach, A. and Opolski, K. (eds.), *Financialization and the Economy*, Taylor and Francis.

Scarano, G. (2015). Corporate Savings and the 2007-2009 Financial Crisis: A Warning for the European Banking Union. In Calciano, F., Fiordelisi, F. and Scarano, G. (eds.), *The Restructuring of Banks and Financial Systems in the Euro Area and the Financing of SMEs*, London and New York: Palgrave Macmillan, 175-187.

Scarano, G. (2016). Connections between corporate governance, corporate savings and business cycles in the economic literature around the Great Depression, *Dipartimento di Economia Università degli studi Roma Tre, Working papers*, 207.

Schumpeter, J.A. (1939). *Business Cycles. A Theoretical, Historical and Statistical Analysis of the Capitalist Process*. New York and London: McGraw-Hill Book Company.

Skott, P. (1989). *Conflict and Effective Demand in Economic Growth*, Cambridge: Cambridge University Press.

Stockhammer, E. (2004). Financialisation and the Slowdown of Accumulation, *Cambridge Journal of Economics*, 28(5), 719-741.

Stockhammer, E. (2006). Shareholder value orientation and the investment-profit puzzle, *Journal of Post Keynesian Economics*, 28(2), 193-215.

The Economist (2012). Dead money, *Finance and Economics section*, November 3rd, Weekly edition.

Tinbergen, J. (1939). *Statistical Testing of Business Cycle Theories: Part I: A method and its application to investment activity*. New York: Agaton Press.

Tirole, J. (2006). *The Theory of Corporate Finance*. Princeton and Oxford: Princeton University Press.

Tobin, J. (1969). A general equilibrium approach to monetary theory, *Journal of Money, Credit and Banking*, 1(1), 15-29.

Tobin, J. (1997). Comment. In Pollin, R. (eds.), *The Macroeconomics of Savings, Finance and Investment*. University of Michigan Press, Ann Arbor.

Tornell, A. (1990). Real vs. financial investment: can Tobin taxes eliminate the irreversibility distortion?, *Journal of Development Economics*, 32(2), 419-444.

Chapter 8

A multidimensional performance indicator to evaluate the effectiveness of Italian university education

Silvia Terzi and Francesca Petrarca

8.1 Introduction

In the present paper we suggest the use of Alkire-Foster (2009, 2011) dual cut-off method as a measure of multidimensional achievements and thus of performance.

The Alkire Foster (AF) methodology has been introduced as a multivariate deprivation measure, and it has led to the currently used Multidimensional Poverty Index (MPI). It consists in aggregating (possibly weighted) information on deprivations across all dimensions for each person, identifying multidimensionally poor persons on that basis, and subsequently aggregating across poor people to obtain a poverty measure. Thus, it reflects the joint distribution of deprivations.

The relevance of understanding interconnections among multiple deprivations was highlighted in the 2009 Report of the Commission on the Measurement of Economic Performance and Social Progress, which stresses the particular relevance of joint distribution when studying disadvantage: «For example, the loss of quality of life due to being both poor and sick far exceeds the sum of the two separate effects, implying that governments may need to target their interventions more specifically at those who cumulate these disadvantages» (p. 55).

In contrast «A marginal method reflects population deprivations within dimensions but does not look across dimensions for the same person, and cannot reflect the extent of associations among deprivations» (Foster 2011, p. 303).

In this paper we illustrate some interesting features of this methodology and suggest its use to measure performance and/or quality or customer satisfaction both in business environment and in public organizations. In particular we resort to AF dual cut-off method to evaluate the effectiveness of the Italian university education.

In fact, theoretical concepts like satisfaction or performance are generally recognized as multidimensional latent constructs exactly as non-income poverty; to define a synthetic indicator of performance (or satisfaction) we need to aggregate the sub indicators related to the different dimensions the multivariate indicator is based on. Aggregating different dimensions, it would be extremely useful to embed in the performance indicator some information concerning the association between the single dimensions. This is the main reason for which we suggest AF methodology. The second reason is due to the fact that sub indicators concerning performance are often measured on an ordinal scale, when not binary (success/failure).

In our case study we have a set of several indicators measuring different types of good performance for undergraduate and graduate courses of Roma Tre University. Let us assume that we can set a reference value for each indicator, indicating a good standard. Of course, we could count in how many indicators each under-graduate or graduate degree programme reaches the reference value and then define as effective a degree programme that reaches reference values in at least k% of the indicators. This is the basic idea. The overall performance indicator is obtained counting how many degree programmes are well performing (for each Department or each University).

8.2 The methodology

Assume we have different areas of performance (or poverty as in the original context) and no natural definition of an aggregate variable. The different areas or *dimensions* could be – and indeed will often be – measured on an ordinal scale, such as: customer's satisfaction/ appreciation in the evaluation of the quality of services. For each dimension, one could define a specific threshold/reference value or *cut-off* (as Alkire and Foster call it) and identify who is above and who is below each of these one-dimensional thresholds.

The second step consists in establishing a second reference value (or second level cut-off) usually indicated with *k*, to define as multidimensionally effective (poor in the original context) the unit that exceeds the first threshold in at least *k* dimensions or key-indicators. In other words, the second cut-off value defines how many successes a unit must record in order to be defined as effective *tout court*. If we set *k=d* this would lead us to the intersection-based approach, i.e. to consider multidimensionally effective the units that reach or exceed the reference thresholds in all key indicators.

Vice versa, if we set the second cut-off value equal to *1* (i.e. *k=1*) this would lead us to the union based setting: an effective unit is successful in any of the key indicators. For *1<k<d* we have intermediate solutions; and this is one of the advantages of the method.

Let $w_j (j = 1, ..., d)$ be the weight applied to the j-th dimension, and let $\sum w_j = d$, so that the weights w_j of the different dimensions add to the total number of areas d. Let $c_i (i = 1, ..., n)$ be the weighted number of achievements reached by the i-th unit; choose a performance cut-off k such that $0 < k \leq d$, and define multidimensionally effective the unit whose achievement count c_i is greater than k. Let q be the number of effective units, and let $c_i(k)$ be the count of the (weighted) achievements only for the effective units. P_0, the performance indicator can be defined as: $P_0 = \sum c_i (k)/nd$, i.e. the weighted average of the number of achievements in the population.

P_0 can also be expressed as a product between two measures: the

incidence of effective units (H) and the intensity of achievements (*A*); more precisely: $P_0 = H \times A$, where $H = q/n$ and $A = \sum c_i(k)/dq$.

It is logical to expect that, as the cut off *k* varies, both the degree of incidence and the intensity will change. More specifically, as the second cut-off *k* increases, H is reduced because fewer and fewer units will be able to obtain a sufficient number of achievements; but at the same time, the positive variation of *k* increases *A*, producing an opposite effect on the final indicator P_0.

On the contrary, by choosing a lower value for *k*, the increase in incidence contrasts with the reduction in intensity, with an uncertain effect on P_0, an effect that depends on the individual univariate distributions.

It is important to underline two other important properties of the methodology proposed by Alkire and Foster: multidimensional monotonicity and decomposability by subgroups.

The first property implies that if an additional achievement is recorded for a statistical unit, the overall index increases.

The decomposability, on the other hand, is based on the fact that if there were two distinct populations *x* and *y* (for example two different Universities, or two different departments or degree programmes of the same University), of n_x and n_y units, the index P_0 referred to the union of the two populations will be the average of $P_0(x)$ and $P_0(y)$ weighted with their respective numerosity.

8.3 An example application

Assume that we want to evaluate student's appreciation of the study programmes held by the Economics Department, based on the results of the questionnaires on students' opinions (OPIS). The Economics Department offers one undergraduate study course and three master courses. The students that attend lectures are asked 13 questions, concerning three distinct macro-environments.
For the first macro environment the questions are:

-        Was the knowledge you already had sufficient to understand the topics of the examination programme?
-        Is the study load of the course proportionate to the credits awarded?
-        Is the (indicated and available) didactic material adequate for the study of the subject?
-        Have the examination procedures been clearly defined?

For the second macro-environment:
-        Are the timetables for lessons, exercises and other educational activities respected?
-        Does the lecturer stimulate/motivate interest in the discipline?
-        Does the lecturer expose the topics clearly?
-        Are supplementary teaching activities (exercises, tutorials, workshops, etc. ...), where existing, useful for the learning of the subject?
-        Was the teaching conducted consistently with what was stated on the website of the degree course?
-        Has the principal lecturer held regular lessons?
-        Is the lecturer available for clarifications and explanations?

For the third macro-environment:
-        Regardless of how the teaching was carried out, are you interested in the topics covered in the teaching?
-        Are you on the whole satisfied with this course?

        For each question the student is asked to indicate a degree of satisfaction (totally unsatisfied, rather unsatisfied, rather satisfied, totally satisfied).

        We chose to give the same weight to each question; and to set as first level threshold a percentage of 'totally satisfied' at least equal to 50%.
        As far as the second level threshold k is concerned, we analyzed two different settings: *k=10* and *k=8*.
        Recalling that $P_0 = \sum c_i(k)/nd$, is the weighted average of the

number of achievements in the population; $H = q/n$ is the incidence of effective units; $A = \sum c_i (k)/dq$ is the intensity of achievements; and furthermore $P_0 = H \times A$, we set $k=8$ and computed P_0, H and A for all study courses but also for undergraduate course and for master courses. Then we computed the performance indicator separately for the four different disciplinary areas (that we have called *W, X, Y, Z*). The results are collected in table 8.1.

Table 8.1 – *Performance indicators for k=8, for different course levels (Overall, Undergraduate and Master), for four different disciplinary areas (W, X, Y, Z).*

k=8				
Area	*Study programme*	**P$_0$**	**H**	**A**
All disciplinary areas	*Overall*	*0,21*	*0,27*	*0,79*
	Undergraduate	*0,22*	*0,29*	*0,75*
	Master	*0,21*	*0,25*	*0,84*
W	*Overall*	*0,29*	*0,38*	*0,77*
	Undergraduate	*0,35*	*0,48*	*0,72*
	Master	*0,24*	*0,29*	*0,83*
X	*Overall*	*0,16*	*0,19*	*0,87*
	Undergraduate	*0,22*	*0,27*	*0,82*
	Master	*0,1*	*0,1*	*1*
Y	*Overall*	*0,19*	*0,22*	*0,85*
	Undergraduate	*0,06*	*0,08*	*0,77*
	Master	*0,52*	*0,6*	*0,87*
Z	*Overall*	*0,04*	*0,06*	*0,62*
	Undergraduate	*0*	*0*	*not a number*
	Master	*0,06*	*0,09*	*0,62*

In Table 8.2 we report the results for k=10.

Table 8.2 – *Performance indicators for k=10, for different course levels*
(Overall, Undergraduate and Master), for four different disciplinary areas (W, X, Y, Z).

k=10				
Area	**Study Programme**	P_0	**H**	**A**
All disciplinary areas	*Overall*	*0,14*	*0,16*	*0,88*
	Undergraduate	*0,14*	*0,16*	*0,85*
	Master	*0,15*	*0,16*	*0,92*
W	*Overall*	*0,16*	*0,18*	*0,91*
	Undergraduate	*0,18*	*0,20*	*0,88*
	Master	*0,15*	*0,16*	*0,94*
X	*Overall*	*0,16*	*0,19*	*0,87*
	Undergraduate	*0,22*	*0,27*	*0,82*
	Master	*0,10*	*0,10*	*1,00*
Y	*Overall*	*0,19*	*0,22*	*0,85*
	Undergraduate	*0,06*	*0,08*	*0,77*
	Master	*0,52*	*0,60*	*0,87*
Z	*Overall*	*0*	*0*	*not a number*
	Undergraduate	*0*	*0*	*not a number*
	Master	*0*	*0*	*not a number*

We next tried to see whether the poor performance of disciplinary area Z is a critical performance or not. So, we defined a *new* threshold (first order cut-off) corresponding to a % of satisfied (rather satisfied and totally satisfied) of at least 50% and computed the frequency distribution of the number of thresholds achieved:

Table 8.3 – *Frequency distribution of the number of thresholds achieved for area Z*

Thresholds achieved	frequency
8	1
9	2
10	1
11	5
12	12
13	91
Total	**112**

Most lecturers (81%) reach and exceed *all* first order cut-off values, all of them achieve at least 8 thresholds. This means that for k=8 and for this lower first order cut-off value, all courses in disciplinary area Z are above the standard of adequacy we have set.

References

Alkire, S. and Foster, J. (2009). Counting and Multidimensional Poverty Measurement, *Journal of Public Economics*, 95(7-8), 476-487.

Alkire, S. and Foster, J. (2011). Understandings and Misunderstandings of Multidimensional Poverty Measurement, *Journal of Economic Inequality,* 9(2), 289-314.

Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures, *Econometrica, 52*(3), 761-766.

Menghini, M. and Terzi, S. (2016). L'indice di performance ICM per l'analisi della competitività delle imprese: una proposta e un'indagine empirica per l'Italia, *Rivista economica del Mezzogiorno*, 30(1), 71-96, doi: 10.1432/83372, ISBN: 978-88-15-26318-6.

Chapter 9

The debate on capacity adjustment in the Classical-Keynesian approach: a brief historical account

Attilio Trezzini and Antonella Palumbo

9.1 Introduction

This work explores an analytical issue which has attracted much attention within the theories that explain economic growth as a demand-led phenomenon, namely the issue of the adjustment of productive capacity to the level and growth of aggregate demand. We will provide a historical account of the debate on the matter, particularly focusing on the analytical developments internal to the Classical-Keynesian approach. The latter, which is part of the category of theories that aim to extend the Keynesian principle of the crucial role of aggregate demand in the determination of output to the analysis of the long-run tendencies of the economy, is characterized by the fact that it combines an analysis of growth along 'Keynesian' lines with the modern revival of Classical Political Economy – following and developing the proposal of Garegnani (1978-9).

The capacity adjustment issue has proven particularly relevant in the Classical-Keynesian approach, given the theoretical and methodological characteristics of the latter, and has given rise to much literature and a lively discussion. In addressing what originally appeared merely as a logical problem and in trying to correctly define the specificity of the Classical-Keynesian approach in comparison with other demand-led growth approaches, the participants in the debate have acquired in our opinion progressive awareness of the complexity of the object of analysis, which has induced them to adjust both method and theory to such complexity.

If the debate is still open, since neither theoretical elaboration nor empirical verification have given conclusive answers, yet in our view it has produced many insights on the actual dynamics of the process of growth and opened some potentially fruitful theoretical routes.

The paper is structured as follows. Section 9.2 is devoted to a brief reconstruction of the approach. Section 9.3 illustrates the capacity adjustment question and the way it has been addressed by the two main strands of thought that can be identified in the approach. Section 9.4 explores the methodological questions that arise in the analysis of growth. Section 9.5 is devoted to the question of the so-called 'Harrodian instability' and, connectedly, to the analysis of investment, and highlights the open issues within the approach. Section 9.6 concludes.

9.2 Keynes plus Sraffa: the principle of effective demand and the theory of distribution

Keynes's and Sraffa's theoretical elaborations proceeded quite independently of each other, and have given rise to different traditions of thought with apparently very little in common. While many Keynesian authors focus on monetary factors, the short period, the role of uncertainty and expectations, instead the approach that stems from Sraffa's work is frequently regarded as dealing exclusively with normal long-period positions, relative prices, real magnitudes. However, this characterization is at best partial, and the potentiality for a fruitful cross-fertilization between Keynes's and Sraffa's thought is high.

The group of authors that in Cambridge were in contact with either or both, most notably Joan Robinson and Nicholas Kaldor in the first generation of scholars, and Luigi Pasinetti and Pierangelo Garegnani later, while becoming progressively aware of the insuperable theoretical limits of the neoclassical theory, knew that complete abandonment of the latter was necessary for the full development of the most innovative part of Keynesian theory. The principle of effective demand, stating the plurality of possible equilibria and the non-existence of any automatic

tendency towards full employment, is in fact deeply incompatible with neoclassical theory (Garegnani 1978-9).

Soon after the publication of the *General Theory*, instead, the successful attempt by Hicks (1937) and others to reinterpret Keynes's theory along more traditional lines, meant that in few years a new consensus, the 'neoclassical synthesis', established in macroeconomic theory, in which Keynes's most innovative conclusions were either denied or so 'watered down' (Pasinetti 1974) as to be unrecognizable. Effective demand had become relevant only to understand fluctuations, and active demand management only as a policy tool aimed at stabilization. Outside the short period, the conclusions of neoclassical theory – namely the tendency to full employment and the exclusive preoccupation with supply-side forces as determinants of growth – were restated. This was possible, according to Garegnani (1978-9), exactly due to the presence in Keynes's analysis of those principles – marginal products and the substitution among factors on which they are based – which allowed neoclassical theory to regard full employment as the normal result of the action of market forces.

The attempt to fully recover Keynes's message and make it into a theory of growth – based on the idea that aggregate demand determines economic growth and no in-built tendency exists in the system to grow along a full-employment path – thus implies both the need to get rid of all neoclassical theoretical legacy, and the need for an alternative theory of value and distribution. This theoretical attitude is common to all the different strands of thought that form the composite field of demand-led growth theories, although each of them either adopts different theories of distribution or solves differently the analytical questions involved. Some strands of thought within the approach derive instead directly from Kalecki's own independent formulation of the principle of effective demand[1], giving rise to the so-called 'Kaleckian' or 'neo-Kaleckian' growth models – also sharing the demand-led growth perspective and the refusal of the neoclassical theory of distribution.

A relevant role in the development of the demand-led growth

---

[1] See the papers of the 1930s collected in Kalecki (1971).

approach has been played by Harrod's (1939) well-known analysis of the 'inherent instability' of a dynamic system – caused by the possible inconsistency between the 'warranted' rate of growth (defined by the ratio between the propensity to save and the capital/output ratio and implying normal utilization of capacity), the natural rate of growth (defined by the exogenous growth in population and productivity) and the actual one.

The so-called 'Cambridge' theory of distribution (Kaldor 1955-6, 1957; Robinson 1962; Pasinetti 1962), which is one of the early developments of the demand-led growth approach, was born out of the need to address such instability question. Kaldor (1955-6), for example, states that the variability of the collective propensity to save due to the changes in distributive shares implies stability of the full-employment growth path. The latter might be reached either as a result of a strong private incentive to accumulation or as an effect of deliberate public policies[2]. So the system is not doomed to perpetually and violently oscillate between cumulative contraction and inflationary expansion.

However, the strict link these authors postulate between accumulation and distribution is ultimately based on the assumption that the ratio between output and capacity is fixed, i.e. that normal utilization of capacity prevails continually, which implies inelasticity of output to independent changes in aggregate demand – and the need for normal distribution to react to demand changes in order for savings to adjust to investment. While the 'Keynesian Hypothesis' (the independence of investment from saving; Kaldor 1955-6) is respected, the Cambridge theory of distribution thus implies continuous normal utilization and inelasticity of output to demand – two characteristics that the subsequent literature has shown to be particularly at odds with the principle that demand is independent of capacity and drives growth (Garegnani 1992).

Starting from such critical analysis of the Cambridge theory of distribution, the different strands of thought that compose the demand-

---

[2] Robinson (1962) explores different possible paths of growth of the system, not necessarily characterized by full employment of labour.

led growth approach have addressed these questions – how to express analytically the principle of demand-led growth, how normal and actual utilization of capacity interact, how to treat the determinants of investment and deal with Harrodian instability – though providing different answers. We will not review here this rich literature, made of such different approaches as those represented by the Harrodian (Skott 2019; Fazzari *et al.* 2013), Kaldorian (Thirlwall 2011) and neo-Kaleckian models of growth (Rowthorn 1981; Dutt 1990; Amadeo 1986; Badhuri-Marglin 1990; Lavoie 2014, 2016). All these models, including other less clearly classifiable ones that however share the same demand-led perspective, are usually also defined as 'Post-Keynesians' (see Setterfield 2019 for a synthesis of different approaches, and Park 2012 and Commendatore *et al.* 2003 for surveys).

Rather, we will explore the above-mentioned analytical questions through the lens of one particular approach, the Classical-Keynesian one.

9.2.1 The Sraffa-Keynes synthesis and the Classical-Keynesian approach

The classical theory of value and distribution, as rediscovered and reformulated by Sraffa (1960), constitutes a possible theoretical basis, alternative to the neoclassical one and not prone to the same theoretical inconsistencies, on which to found the long-period theory of effective demand. Although Sraffa's and Keynes's worlds are apparently alien to each other, yet there are many reasons to try this analytical route. Pasinetti (1974) pointed out the deep analogy between Keynes's method of analysis and that proper to classical economists, especially as regards the characteristic separation of theory into different stages. On a similar vein, Vianello (2007) points out, as common elements, the fundamental role of conventional and institutional factors and the recognition of a 'hierarchy' among economic relations, with some having less generality and a more limited applicability than others. In a seminal contribution, Garegnani (1978-9) argued that, while neoclassical theory postulates the spontaneous tendency to full employment as a necessary implication of the working of the system of prices, the classical theory of value and distribution is instead 'open' with respect to the theory of output. It regards in fact output levels as givens in determining relative prices and

the rate of profit. This implies that the theory refrains from establishing any sort of *necessary* connection between output and prices, or between accumulation and distribution. This 'degree of freedom' implies that classical theory is compatible with a theory of output in which underutilization of resources is a normal result of market forces. Analogously, Keynes's principle of effective demand only entails that investment – determined independently in a different part of the analysis – is independent of current output, but does not necessarily presupposes a specific theory of investment. Thus the principle of effective demand may be combined with a theory of investment different from that adopted by Keynes – and a different set of relations regarding the determination of prices and distribution.

The Classical-Keynesian approach stems from these insights, and particularly builds on Garegnani's contributions to this field of analysis (see especially 1992, and 1962 for some early insights). In the recent decades, much literature has flourished starting from these premises (see below for references).

As its basic proposition, the Classical-Keynesian approach states that output is elastic to changes in demand. Such elasticity, in the long period, rests both on the possibility to vary the utilization of installed capacity, and on the creation of new capacity in response to high demand, or the destruction of existing capacity (which is obtained when gross investment is short of reintegration) following long phases of underutilization.

Much attention has been devoted to the analysis of those components of demand which are liable to driving the growth process. Some authors particularly underscore the role of exports, or government expenditure for public consumption, or credit-financed private consumption; others stress also the possible role of technological investment as an independent influence on accumulation, while investment plays the double role of source of demand and vehicle for the creation of capacity.

It is worth describing briefly the characteristics of the Classical-Keynesian approach that directly derive from the Classical theory of value and distribution and that distinguish this approach from the other

strands of analysis in the Post-Keynesian camp. From Classical political economy, the Classical-Keynesian approach inherits in the first place a conflict view of distribution, and in the second place the idea that accumulation and normal distribution are sufficiently independent of each other so as to be studied separately. As already mentioned above, this is an essential characteristic of the approach directly connected to the structure of classical theory, in which output and the real wage are given when determining relative prices and the rate of profit.

To say that accumulation and distribution have to be studied separately does not amount, however, to excluding that many different mutual interrelations may exist between them. For example, a phase of rapid accumulation may increase the bargaining power of workers thus producing higher wages – also positively affecting aggregate demand. However, since such influences are not necessary but contingent, in different historical contexts wages may stagnate in phases of intense accumulation. No general and univocal relations may be established.

Connectedly, the Classical-Keynesian approach inherits from classical theory, in the third place, its characteristic method of analysis. Garegnani (1984) has defined as the 'core' of classical theory that limited analytical field in which relative prices and the rate of profit are determined, by means of a deductive method of analysis. Outside the core, however, in different parts of the analysis, the 'data' of the core (the real wage, the social product, the state of technical knowledge) are analysed and their determinants sought for. Here the method is a mixture of deduction, observation and generalization from experience; the relationships have no universal applicability but are extremely context-sensitive; the interrelations and directions of influence are plural; the historical, political and social forces play an essential role in shaping economic outcomes. This 'historical' method of analysis (Smith 2012) is the one appropriate to study the complexities of the growth processes (see below, section 9.4).

9.3 The debate on capacity adjustment

The basic propositions of the Classical-Keynesian approach have been stated in some seminal contributions (Vianello 1985; Ciccone 1986; Garegnani 1992).

The founding principle is the recognition of the system's capability to create or destroy resources in response to changes in demand. This recognition, according to these authors, immediately poses a methodological question, namely the incompatibility between a theory in which growth is genuinely demand-led and the representation of long-run tendencies by means of *steady-state models*. The latter were already frequently used within other demand-led growth approaches, partly as an effect of the way Harrod (1939) had posed the instability question. It seemed that steady growth – a situation in which all relevant variables grow at a constant and uniform rate, so that demand and capacity grow *pari passu* with constant normal utilization – could represent the only situation in which entrepreneurs, being 'content with what they are doing', would have no incentive to change the rate of accumulation.

However, steady-state models necessarily exclude by construction that deviations between demand and capacity are possible and output is elastic to changes in demand – thus being inconsistent with the very basic proposition of the approach, the primacy of demand – and induce to establish relations between variables that are actually valid only in steady growth and lose their significance otherwise. It was the steady-state hypothesis that, according to Vianello (1985), deceived the 'Cambridge' authors into assuming the long-run inelasticity of output and postulating the necessity of changes in the distributive shares to accommodate changes in the incentive to invest. Outside the steady state, as shown by Garegnani (1992), changes in the rate of accumulation may well affect the actual capital/output ratio and the actual rate of profit, without necessarily affecting normal distribution. Moreover, the steady growth paths cannot be regarded as centres of gravitation for the actual growth paths (Ciccone 1986), since no mechanism can be envisaged bringing back the system to steady growth after an accidental deviation.

While severing the strict link between accumulation and distribution envisaged by the Cambridge theory of distribution, and recovering the functional independence between the two in accord with the Classical approach, these authors also addressed the issue of capacity adjustment.

It should be noted in fact that the earliest elaborations in the Classical-Keynesian approach (Garegnani 1962; Eatwell 1979), had assumed, without further elaboration, the idea that full adjustment between capacity and demand (output) was a property of the long-period positions. Eatwell (1979), for example, defined 'normal' output as entailing full adjustment, and this seems to derive quite directly from the definition of normal prices. The latter are determined by referring to the 'normal' conditions of production, which include the dominant techniques (i.e., the best and most efficient techniques among those widely used) and the normal utilization of capacity. It is also worth noting that the early contributions, namely Garegnani (1962) and Eatwell (1979), in line with most of the Keynesian literature, assumed that investment expenditures could be classified in two different components, namely an autonomous one related to technical innovation, and an induced one representing the mechanism of adjustment of capacity to demand. The potential contradiction between the two assumptions, which was later pointed out by Serrano (1995), apparently had not yet been detected (see section 9.5 below).

The issue of adjustment was raised in the later literature mentioned above. Vianello (1985) noted that any autonomous change in demand would produce a long phase of either over- or under-utilization of capacity, so that, *even assuming* that the system would sooner or later converge to a new 'fully adjusted position', average utilization would necessarily be different from normal over a long period. Garegnani (1992) noted that the full adjustment hypothesis (either continuously or on average) necessarily entails, instead of the autonomy of investment envisaged by the Keynesian Hypothesis, the full dependence of investment on capacity savings – thus, if the latter are entirely exogenous and determined by supply forces, no autonomous role for demand in the growth process. Ciccone (1986), by focusing especially on the

fluctuations that normally characterize the actual working of any market economy, showed the possibility of a systematic discrepancy between normal capacity utilization and the average ex-post long-run utilization of capacity, also highlighting various reasons, especially related to the durability of fixed capital, to maintain that the tendency to full adjustment would never imply the actual realization of such adjustment[3].

This does not amount to denying the validity and usefulness of normal prices and normal distribution, even if such magnitudes are determined by referring to normal utilization. The normal rate of profit is the one that investors expect to realize on newly installed capacity, since they will have sized the flow of investments to fit their expectations about future demand. Gravitation of market prices towards long-period prices thus happens on the gross investment flows, without requiring as a necessary condition that utilization of the whole stock of capacity gravitates contemporarily towards its normal level.

Starting from these contributions, a lively debate has originated within the approach, dwelling especially on the proper representation of the growth process (i.e., discussing the analytical usefulness of the 'fully adjusted positions' in the analysis of growth), and on the proper specification of the investment function[4]. Although with some

---

[3] The Kaleckian models (Rowthorn 1981; Amadeo 1986; Dutt 1990) have also grasped the inconsistency, proper to the 'Cambridge' theories, between the autonomous role of demand and the assumption of a continuously normal capacity utilization. In such models the autonomy of demand in the growth process shows in the endogenous determination of capacity utilization. As is known, however, these models retain the steady-state assumption thus assuming that the system may *persistently* realize a (constant) different-from-normal utilization. This implies the absence of *any* correction of non-normal utilization on the part of entrepreneurs – an assumption quite at variance with the equilibrium nature of steady-growth paths (Committeri 1986; Park 1997). The second generation of Kaleckian models, starting from Badhuri-Marglin (1990), propose different ways to overcome the rigidities of the first-generation models as regards the relations between distribution and growth, but most contributions do not address the utilization paradox. For discussion of possible adjustment mechanisms – based on adjustment of normal to actual utilization – see Lavoie (2014). For a general critical discussion of Kaleckian models of growth, see Cesaratto (2015).

[4] For a reconstruction of the debate see also Aspromourgos (2013).

simplification, and possibly not doing justice to all the authors involved, we classify the contributions to this debate in two main strands of analysis, one assuming that growth should be studied by making use of 'fully adjusted positions', and the other denying any relevant analytical role for them.

9.3.1 The full adjustment hypothesis and the Sraffian supermultiplier

As representative of the first strand of analysis, we may quote Eatwell (2012) and the many contributions that have proposed and developed the so-called 'Sraffian supermultiplier' models (Serrano 1995; Bortis 1997; Dejuàn 2005, 2013; Cesaratto, Serrano and Stirati 2003; Cesaratto 2015; Freitas and Dweck 2013; Freitas and Serrano 2015; Girardi and Pariboni 2015, 2019; Serrano *et al.* 2019). The idea that actual utilization necessarily converges towards normal is also to be found in White (2006), although he regards the adjustment as problematic and difficult to realize, and in Park (1997, 2012), who however supposes that normal utilization may change according to the average utilization actually realized in the past.

For the relevant role they have assumed in the literature, we will focus on the Sraffian supermultiplier models.

The Sraffian supermultiplier (which elaborates, in a different analytical context, on the original supermultiplier put forward by Hicks 1950), is a formula connecting autonomous demand and output. The former is identified, in these models, with the 'unproductive' components of aggregate demand that supposedly do not depend on output: autonomous consumption, both public and private, and exports. Capacity-creating investment is instead regarded as fully induced by demand, since it represents the means by which the stock of capital changes and capacity adjusts to the requirements of production. Along with the propensity to consume, another propensity appears in the supermultiplier, i.e. the 'propensity to invest' out of income. The formula thus allows to define long-run output as a multiple of autonomous demand. When the propensity to invest is defined so as to ensure capital stock adjustment, by construction such an output level

entails normal utilization of capacity[5].

Dynamically, the rate of output growth is determined entirely by the rate of growth of autonomous demand. If the full adjustment hypothesis entails, as Garegnani had noted, that investment does in fact adjust to capacity savings, yet in this model capacity savings depend, in their turn, on demand, being affected by the share of autonomous demand in output (Serrano 1995).

The Sraffian supermultiplier model shows logically that, *even assuming full adjustment*, there would be no reason to conclude that normal distribution has to react to changes in the rate of accumulation. The system may accommodate a different rate of growth, determined by the exogenous growth of autonomous demand, thanks to induced changes in the autonomous demand-output ratio.

As regards the issues raised in the above-mentioned contributions by Ciccone (1986), Vianello (1985) and Garegnani (1992), i.e. the critique to steady-state models and the possibility of defining long-period positions regardless of the full adjustment hypothesis, these arguments are accepted by the supermultiplier authors. Freitas and Serrano (2015) clearly state that the supermultiplier formula should not be read as a steady-state model[6]. Rather, it expresses the relationships that have to hold in a fully adjusted situation, although the system normally operates out of such positions. Actual utilization – they maintain – is normally determined endogenously; in a very long period, however, the system would converge towards a fully adjusted position in which all the relations envisaged by the supermultiplier model hold. The process of adjustment of capacity to demand, they admit, is slower and more tortuous than the process of gravitation of market prices towards their normal values; however, the tendency to adjustment is strong enough to materialize, at a certain point. Fully-adjusted positions, which cannot

---

[5] See below for a possible different interpretation of the supermultiplier formula, not entailing full adjustment when the propensity to invest is exogenous.

[6] In an early critical assessment, Trezzini (1995, 1998) had pointed out that, if the Sraffian supermultiplier were interpreted as a steady-state model, demand would actually not be independent of capacity since autonomous demand would be forced to grow at a rate entirely determined by supply conditions.

represent, even on average, the actual state of the economy, are however bound to be realized sooner or later, which implies that they retain a relevant analytical role.

This account highlights the crucial role of the mechanism through which capacity adjusts to demand. Following Freitas and Serrano (2015) and Girardi and Pariboni (2015), we may sketch it as follows. Capacity-creating investment, as seen above, is entirely induced and represents the adjustment mechanism. However, the reaction of investment to misalignments between capacity and demand cannot be too strong, since otherwise the excessive investment flows would produce divergence instead of convergence. Investment cannot thus react too strongly nor too soon to demand changes. Although the adjusting flows might produce false signals to investors, it is assumed that the latter ultimately form their expectations by observing the rate of growth of aggregate autonomous demand and slowly adapting to it. Convergence, it follows, requires that a certain rate of growth of autonomous demand persists long enough to let the investors' expectations adjust[7].

It can be maintained that these hypotheses – the constancy of the rate of growth of autonomous demand for long periods and the low sensitivity of investment to under- or over-utilization – are unduly restrictive. Especially, they seem to derive from the formal properties required for convergence rather than from a direct analysis of the way investment decisions are actually taken in a market economy[8].

However, no doubt the supermultiplier models highlight a fundamental question, which has to be addressed not only in the Classical-Keynesian approach but also, more in general, in demand-led growth theories, i.e. the convergence question. We take up the issue again in section 9.5 below.

9.3.2 The flexible utilization approach
The second line of analysis along which the issue of adjustment of

---

[7] For a recent reassessment of the convergence question, see Serrano *et al.* (2019) discussed in section 9.5 below.

[8] See below, section 9.5, for a fuller discussion about the determinants of investment.

capacity to demand has been addressed is the 'flexible utilization approach'[9]. Either reflecting on the above-mentioned questions raised by Garegnani (1992), Ciccone (1986) and Vianello (1985), or through independent routes, some authors (Trezzini 1995, 2013, 2017; Ciccone 1987, 2011; Vianello 1989; Ciampalini and Vianello 2000; Smith 2012, 2018; Parrinello 2014; Palumbo and Trezzini 2003; Trezzini and Palumbo 2016; Palumbo 2015) have developed an approach based on the total rejection of the use of fully adjusted positions as theoretical tools for the analysis of growth.

According to this approach, the autonomy of demand implies that, although capacity is installed on the basis of *expected* demand, once capacity is installed *actual* demand (its fluctuations and its trend) is independent of the existing capacity.

This position was initially argued as a theoretical reasoning only. Both the Keynesian and the Sraffian critiques imply the absence of any mechanism ensuring that the level of aggregate demand is such as to determine normal utilization of existing capacity. As seen above (section 9.3), even assuming that the system starts from a fully adjusted position and reverts to another one after the whole effects of a change in the pace of accumulation have taken place, utilization would be different from normal along the whole adjustment path, and thus on average, due to the effect of the adjusting investment flows (Vianello 1985). Such variability in utilization is the only possible way in which demand may deviate from capacity and adjustment of capacity to demand may occur (Garegnani 1992; Trezzini 1995; Park 1997). Also the supermultiplier authors, as mentioned above, have recognized that phases characterized by non-normal utilization are essential to hypothesize any tendency of capacity to adjust to demand. However, the two different approaches draw different conclusions from this same consideration.

According to the authors of the flexible utilization approach, fully adjusted positions, not being representative of the average trend described by the development of real economies over time, have no role

---

[9] So labeled in Trezzini and Palumbo (2016). Cesaratto (2015) defines the same approach as the 'First Sraffian Position'.

to play in the analysis. This position, it is worth noting, does not imply denying the existence of a tendency of capacity to adjust to demand. Rather, this tendency is considered so relevant as to be the very means through which demand affects supply factors. However, although constantly at work, such tendency is not effective enough to overcome entirely the other influences on investment flows and actually produce full adjustment.

The flexible utilization literature identifies various reasons for such a conclusion. First, the process of adjustment of capacity to demand necessarily implies investments that create not only capacity but also additional demand, which alters the very trend of demand to which capacity tends to adjust (Trezzini 1995, 2013; White 2006). Second, capacity does not react instantaneously to changes in demand and in some circumstances may even not react at all to persistent deviations of actual utilization from normal. This is not only due to wrong expectations, but rather, especially, to the durability of fixed capital and the high elasticity with which it can be used (Ciccone 1986, 2011). Firms may sustain a different from normal utilization on their installed capacity for long and will correct capacity only after observing systematic and lasting over-or underutilization. Which implies that only *average long-period* different-from-normal utilization would affect the decision to invest, since temporary deviations are regarded as absolutely normal (Palumbo and Trezzini 2003; Trezzini and Palumbo 2016; Hein, Lavoie and van Treeck 2011).

Also due to such elasticity, firms may regard not a single degree of utilization but a whole range of them as normal or at least 'acceptable' (Dutt 1990; Hein, Lavoie and van Treeck 2011; Parrinello 2014). Changes in (average) actual utilization within this range could thus imply no adjustments.

Moreover, since normal (desired) utilization can be thought of as the average of the utilization degrees that firms expect to achieve over the entire lifetime of the new plants they are installing, when they compare this desired average degree with the actual average utilization that is realized over fluctuations, they might not be induced to revise investment decisions if existing capacity proves adequate to satisfy the

expected peaks of production (Trezzini 2017).

Being durable, plants and equipment can continue to be used even when the presence of new (superior) production techniques makes them obsolete but not to the point of causing their abandonment. These plants, which can be defined as 'fossils', do not generate normal profits and will not be reproduced; their under- or over-utilization will not be corrected until they are discarded altogether and completely replaced by new capital.

9.4 On method:
how to study growth without normal utilization

As seen above (section 9.3.1) a complete adjustment of capacity to demand would require that autonomous expenditure grows for a *practically unlimited period* at a constant rate and that entrepreneurs form their expectations and investment decisions not on the actual levels of their own demand, but on the basis of the regular evolution of the autonomous components of aggregate demand (Trezzini 1998; White 2006). Technical progress inducing changes in the ratio between output and fixed plants would disturb the adjustment and is normally assumed away.

Constancy of some conditions is generally assumed for the determinants of the theoretical positions towards which, according to theory, the market tends to converge or gravitate. The assumption of such constancy, however, cannot be arbitrary but has to reflect an actual 'relative persistence' of the determinants of the theoretical position if compared with other, more transitory, factors (Trezzini 2013)[10]. Not only is this property absent in the determinants of fully adjusted positions, since no realistic reasons justify the hypotheses of unlimited constancy in the growth of autonomous demand and in technical

---

[10] Such relative persistence, according to classical theory, is proper to the technical and social conditions of production that determine normal prices, which justifies the attention the theory devotes to them.

conditions – if not the need to impose a particular behaviour to variables in order to ensure adjustment. Moreover, the persistence required by the full adjustment hypothesis, due to the length of time it entails, cannot be properly defined as relative persistence but rather as absolute constancy – all the more unwarranted (Trezzini 1998).

A more fruitful approach to the study of growth should be based, we think, on the extraction of laws and regularities from the observation of reality, which would allow the definition of the properties of theoretical positions.

In section 9.3 above we have recalled Ciccone's (1986) analysis according to which full adjustment is not a necessary requirement for the long-term positions used in the study of prices and distribution. Normal prices prevail in the system by virtue of competition and are determined by techniques, defined in terms of normal utilization, that are relevant only for flows of newly installed capacity. Due to the durability of fixed capital, the adjustment of the entire capital stock never actually takes place, but this does not prevent actual prices to gravitate towards their normal levels (Ciampalini and Vianello 2000). When using comparison between normal positions, the unnecessary additional hypothesis of full adjustment of the whole stock of capacity to demand should thus be avoided.

A fundamental implication of the flexible utilization approach is that growth must be conceived as a 'path dependent' process. The autonomy of demand from existing capacity implies not only the variability of utilization over time, but also the impossibility to assume that the system follows predetermined trajectories. The absence of any mechanism ensuring full or normal utilization of resources means that it is impossible to know in advance the (theoretical) level and composition of the output of the next period. Each actual level of demand and output redefines the way in which resources can grow from that moment on. In each period the spectrum of possibilities is redefined. In each single period, the effective level of production activity redraws the set of possibilities for the future.

9.4.1 Historical analysis and the role of models

How should, in this picture, growth paths be studied? The nature of the phenomenon guides the choice of method. As noted above (section 9.2.1), determinants of output and accumulation are analysed outside the core of classical theory, i.e. in that vast field of analysis in which no general universal quantitative relationships may be established between variables, but complex systems of influences that involve historical, political and social forces as essential determinants. A 'historical' method of analysis is thus required (Mongiovi 2011; Smith 2012), which allows to take into account the possibility that the relations between variables change in intensity and even in direction, according to the specific historical circumstances. Not only the use of steady-state models but also, more generally, the use of formal models as such should be avoided as far as the general analysis of growth is concerned, since they are not flexible enough to represent the plurality of contradictory influences, nor their sensitivity to context. Formal models may rather serve to exemplify particular phenomena studied in isolation, provided the complexity of the growth process is always taken into account and not assumed away.

9.5 Harrodian instability and the determinants of investment

As seen above (section 9.2), demand-led growth theories have to address the question of Harrodian instability, whereby the reaction of investment to divergences between actual and normal utilization may produce explosive behaviour of the system.

Within the Classical-Keynesian approach, the supermultiplier models (see section 9.3.1 above) address Harrodian instability through the introduction of autonomous unproductive expenditure. Serrano *et al.* (2019) show that the presence of autonomous demand is sufficient to ensure 'static' stability (i.e. to guarantee that the direction of investment flows is consistent with adjustment between capacity and demand) provided that capacity-creating investment is entirely induced and a rate of growth of autonomous demand persists for long enough for

investment flows to produce, albeit only in the very long run, adjustment of the whole capacity. For the latter to be actually realized ('dynamic' stability), the further hypothesis of low sensitivity of investment to demand is needed. Thus, the full adjustment hypothesis requires a specific theory of investment, which sees it as fully induced and not affected by any other influence that could jeopardise the possibility of adjustment.

The most important implication of this approach concerns the analysis of technical progress. Any hypothetical flow of investment determined solely by the introduction of technical innovations would, in the generality of the cases, have capacity-creating effects. As such, it would however render superfluous the quantity of induced investment that would have created the corresponding amount of productive capacity. Technological investments thus *crowd out* a corresponding amount of induced investment (Cesaratto *et al.* 2003), which implies that the volume of investment is determined exclusively by the capacity adjustment mechanism, the availability of new products and techniques only determining the physical forms that such investment takes.

The flexible utilization approach entails an entirely different analysis of the determinants of investment. The capacity adjustment process is regarded as slow and contrasted by divergent forces, such as irregularity in the pattern of final demand, wrong expectations and the demand effects of adjusting flows (see above, section 9.3.2), which implies that adjustment is normally incomplete. This view suggests that product and process innovations can exert an independent influence on investment, and that technological investment does not automatically crowd out induced investment. Technological investment, then, both is an additional source of demand that drives the growth process and is able to create capacity that only partially replaces existing capacity (or capacity that would have been created anyway).

In accordance with many historical analyses, technological innovations are therefore considered an additional determining force of accumulation. The flexible utilization approach conceives investment as determined by a plurality of factors. The expected evolution of demand, under-utilization or over-utilization of existing capacity, and the

independent influence exerted by technological progress are the main determining causes, but institutional and political factors such as the functioning of financial markets, forms of state direct intervention or regulation, labour market institutions, public spending and infrastructure policies exert such a relevant influence on these factors that they can be considered as independent determinants of investment (Bonifati 2011; Mongiovi 2011).

Such an analysis of investment, and the conclusion, proper to the flexible utilization approach, that output and capacity may grow for considerable periods at different average rates, do however entail the need to address, also within this approach, the issue of Harrodian instability. It is in fact necessary to identify mechanisms that explain why the changes in actual utilization do not generate divergent, potentially explosive paths of growth of demand and capacity. In other words, a decreasing (increasing) utilization of capacity, even if possible for prolonged periods, must at a certain point be corrected and revert into a process of opposite sign.

In fact, this issue is still open within the flexible utilization approach. Attempts to solve this problem are developing in several directions, all based on the attempt to address the study of accumulation starting from observation of the way in which firms actually take investment decisions. One relevant fact of reality is the fluctuating character of the growth process; this, together with a closer critical analysis of the factors producing Harrodian instability, shows the possibility of conceiving non-explosive theoretical positions, even characterized by different-from-normal capacity utilization (Trezzini 2017)[11]. A second empirical regularity has to do with the tendency of firms to produce investment in excess of that justified by demand, both in phases of prolonged expansion and in stagnation, due to the independent stimulus to introduce innovations. In view of the fact that cumulative overexpansion of capacity is never observable, it is possible to assume that the system tends spontaneously to alter (or even produce

---

[11] Outside the Classical-Keynesian approach, see also the penetrating remarks in Setterfield (2019).

abrupt variations in) the rate of technical obsolescence (see the sugge-
stions in Garegnani 1992, p. 53). Observation of the reality of historical
processes suggests that the economy systematically gets rid of whatever
excess of capacity it may have generated (through destruction of existing
capital), avoiding thus any type of Harrodian instability. These insights,
however, have yet to be properly developed form a theoretical point of
view (see Palumbo and Trezzini forthcoming) and empirically tested.

9.6 Concluding remarks

It is clear, from the previous account, that the two main strands of
thought that we have identified within the Classical-Keynesian approach
offer on the whole some useful insights for the analysis of growth and
many interesting questions, although some of the crucial and most
challenging analytical issues are still unsettled.

      The potentiality for development of the approach is relevant,
given the current state of growth theory. The devastating effects of the
Great Recession on output and capacity are undeniable, and have
vindicated the demand-led growth perspective since even in the
mainstream camp the effects of aggregate demand on resource creation
or destruction are increasingly considered. The authors sharing the
demand-led growth perspective have offered in this respect theoretical
and empirical analyses dealing with a series of relevant issues, such as
modelling the demand effects on labour force and productivity (Fazzari
et al. 2020), testing the permanent supply-side effects not only of deep
recessions but also of phases of high growth of demand (Girardi *et al.*
2017), offering alternative methods for estimation of capacity and
potential output (Fontanari *et al.* 2020), analysing the interrelations
between the dramatic fall in wage shares and the stagnating dynamics of
demand (Storm 2017), and so on. Although the various models and
analyses may be open to discussion or criticism in various respects, on
the whole they try to address a set of questions that should be crucial for
anyone interested in understanding the true dynamics of growth.

      This implies that to address in a satisfying way the analytical

questions that have to be solved in the demand-led growth approach is essential to increase the possibility that the approach is more widely recognized as theoretically well-founded and empirically relevant.

Coming back to the analytical questions raised within the Classical-Keynesian approach, one fruitful aspect of the debate is the fact that the authors taking part in it seem to have progressively acquired a deeper knowledge of the complex questions at stake and that some cross-fertilization has taken place between the two different strands of analysis.

As a result of the debate, the supermultiplier authors have recognized the validity of many arguments that originally they did not consider as very relevant, given that their attention was focused on the possibility that the existence of autonomous demand solved the logical puzzle of the full adjustment hypothesis being in apparent contrast with the assumed independence of demand. They have recognized that non-normal utilization is a pervasive characteristic of the growth process, although they assume that fully adjusted situations have a role to play in the analysis, and have developed a non-normal-utilization model. The flexible-utilization authors, on the other hand, have recognized the need of addressing directly the question of convergence of the adjustment mechanism, even if they do not express it in the form of a mathematical function.

A growing number of contributions aim to test empirically the main conclusion of the approach. Particularly, the supermultiplier propositions are tested in a series of contributions, among which Freitas and Dweck (2013) and Girardi and Pariboni (2015). Evidence seems to prove quite convincingly that higher output growth is generally associated with faster growth in 'autonomous' demand, although, as for example Girardi and Pariboni (2015) admit, there is some econometric uncertainty about causation. It has to be noted that empirical analysis necessarily requires a number of assumptions and some empirical interpretation of theoretical concepts that sometimes may be unsatisfying. In particular, 'autonomous' demand has to be defined on the basis of available data rather than on purely theoretical terms, so that the whole of government consumption, exports and residential investment (and sometimes also credit-financed consumption) are defined as autonomous, although it is entirely reasonable to suppose that some parts of

them (or some of them in some periods) are not autonomous but determined by the trend of the economy (see Skott 2019, for a thorough critical analysis of the notion of autonomous demand). On the other hand, empirical evidence clearly shows frequent changes in the estimated supermultiplier, which the authors in the approach regard as compatible with the notion of the 'flexible' supermultiplier (Freitas and Dweck 2013; Freitas and Serrano 2015). All in all, while evidence seems to support the general idea of the Classical-Keynesian approach (and of the demand-led growth perspective) of a fundamental role of demand in economic growth, it offers no conclusive answers to the superior validity of any of the two main strands of analysis within the approach.

In conclusion, what seems to be particularly fruitful is the theoretical flexibility of the Classical-Keynesian approach, which allows to introduce in the analysis of growth some aspects of the reality of accumulation processes that tend to be neglected in the demand-led growth literature. The process of growth in its historical realizations seems to be characterized, much more than it is assumed in growth models, by structural, institutional or technological changes. Moreover, it is intrinsically irregular, especially as regards the irregular fluctuations observable in demand and output components. Fixed capital is in reality an aggregate of heterogeneous capital goods with their own irreversible physical peculiarity, which implies that capacity adjustment cannot take the simple form envisaged in models. Not only may normal utilization correspond to a whole range of values instead than a single one, but the reaction of firms to the divergence between actual and normal utilization may be much more complex and less automatic than generally assumed. The capital-output ratio will not be constant; the durability and elasticity of use of fixed capital also imply that obsolescence is not independent of the growth process, but, on the contrary, strongly influenced by the economic phase.

The explicit consideration of these and possibly other aspects of real growth processes within theoretical analysis is, in our opinion, the route along which the Classical-Keynesian approach should develop, so as to further increase its ability to address the actual complexities of the growth phenomenon.

References

Amadeo, E.J. (1986). The Role of Capacity Utilization in Long-period Analysis, *Political Economy,* 2, 147-60.

Aspromourgos, T. (2013). Sraffa's System in Relation to Some Main Currents in Unorthodox Economics. In Levrero, E.S., Palumbo, A. and Stirati, A. (eds.), *Sraffa and the Reconstruction of Economic Theory,* vol. III. Palgrave Macmillan.

Bhaduri, A. and Marglin, S. (1990). Unemployment and the real wage: the economic basis for contesting political ideologies. *Cambridge Journal of Economics*, 14, 375-393.

Bonifati, G. (2011). The capacity to generate investment: an analysis of the long-term determinants of investment. In Ciccone, R., Gehrke, C. and Mongiovi, G. (eds.), *Sraffa and Modern Economics,* vol. II. London: Routledge.

Bortis, H. (1997). *Institutions, Behaviour and Economic Theory: A Contribution to Classical-Keynesian Political Economy*, Cambridge: Cambridge University Press.

Cesaratto, S. (2015). Neo-Kaleckian and Sraffian Controversies on the Theory of Accumulation, *Review of Political Economy*, 27, 154-182.

Cesaratto, S., Serrano, F. and Stirati, A. (2003). Technical Change, Effective Demand and Employment, *Review of Political Economy,* 15, 33-52.

Ciampalini, A. and Vianello, F. (2000). Concorrenza, accumulazione del capitale e saggio del profitto. Critica al moderno sottoconsumismo. In Pivetti, M. (eds.), *Piero Sraffa. Contributi per una biografia intellettuale*, Roma: Carocci.

Ciccone, R. (1986). Accumulation and capacity utilization: some critical considerations on Joan Robinson's theory of distribution, *Political Economy: Studies in the Surplus Approach*, 2, 17-36.

Ciccone, R. (1987). Accumulation, capacity utilization and distribution: a reply, *Political Economy: Studies in the Surplus Approach*, 3, 97-111.

Ciccone, R. (2011). Capacity utilization, mobility of capital and the classical process of gravitation. In Ciccone, R., Gehrke, C. and

Mongiovi, G. (eds.), *Sraffa and Modern Economics,* vol. II. London: Routledge.

Commendatore, P., D'Acunto, S., Panico, C. and Pinto, A. (2003). Keynesian Theories of Growth. In Salvadori, N. (eds.), *The Theory of Economic Growth: a Classical Perspective*. Aldershot, Elgar.

Committeri, M. (1986). Some comments on recent contributions on capital accumulation, income distribution and capacity utilization, *Political Economy*, 2, 161-186.

Dejuàn, Ó. (2005). Paths of Accumulation and Growth: towards a Keynesian Long-Period Theory of Output, *Review of Political Economy*, 17, 231-252.

Dejuàn, Ó. (2013). Normal Paths of Growth Shaped by the Supermultiplier. In Levrero, E.S., Palumbo, A. and Stirati, A. (eds.), *Sraffa and the Reconstruction of Economic Theory,* vol. II, Palgrave Macmillan.

Dutt, A.K. (1990). *Growth, Distribution and Uneven Development*, Cambridge: Cambridge University Press.

Eatwell, J. (1979). Theories of Value, Output and Employment, *Thames papers in political economy*, London.

Eatwell, J. (2012). The Theory of Value and the Foundations of Economic Policy: In Memoriam Pierangelo Garegnani, *Contributions to Political Economy,* 31, 1-18.

Fazzari, S.M., Ferri, P., Greenberg, E.G. and Variato, A.M. (2013). Aggregate demand, instability, and growth, *Review of Keynesian Economics*, 1, 1-21

Fazzari, S.M., Ferri, P. and Variato, A.M. (2020). Demand-led growth and accommodating supply. *Cambridge Journal of Economics*, 44(3), 583-605.

Fontanari, C., Palumbo, A. and Salvatori, C. (2020). Potential Output in Theory and Practice: A Revision and Update of Okun's Original Method*, Structural Change and Economic Dynamics,* 54, 247-266.

Freitas, F. and Dweck, E. (2013). The Pattern of Economic Growth of the Brazilian Economy 1970-2005: A Demand-Led Perspective.

In Levrero, E.S., Palumbo, A. and Stirati, A. (eds.), *Sraffa and the Reconstruction of Economic Theory,* vol. II. Palgrave Macmillan.

Freitas, F. and Serrano, F. (2015). Growth Rate and Level Effects, the Stability of the Adjustment of Capacity to Demand and the Sraffian Supermultiplier, *Review of Political Economy*, 27, 258-281.

Garegnani, P. (1962). *Il problema della domanda effettiva nello sviluppo economico italiano*, Svimez, Roma.

Garegnani, P. (1978-9). Notes on Consumption, Investment and Effective Demand, Part I and II, in *Cambridge Journal of Economics* 2, 335-354; 3, 63-81.

Garegnani, P. (1984). Value and distribution in the classical economists and Marx, *Oxford Economic Papers,* 36, 291-325.

Garegnani, P. (1992). Some Notes for an Analysis of Accumulation. In Halevi, J., Laibman, D. and Nell, E. (eds.), *Beyond the steady-state*. Basingstoke & London: Macmillan.

Girardi, D. and Pariboni, R. (2015). Autonomous demand and economic growth: some empirical evidence, *Centro Sraffa Working Papers*, n. 13.

Girardi, D. and Pariboni, R. (2019). Normal utilization as the adjusting variable in Neo-Kaleckian growth models: A critique, *Metroeconomica*,70, 341-358.

Girardi, D., Paternesi Meloni, W. and Stirati, A. (2017). Persistent Effects of Autonomous Demand Expansions. *Institute for New Economic Thinking*, Working Paper, 70.

Harrod, R.F. (1939). An Essay in Dynamic Theory, *The Economic Journal*, 49, 14-33.

Hein, E., Lavoie, M. and van Treeck, T. (2011). Some instability puzzles in Kaleckian models of growth and distribution: a critical survey, *Cambridge Journal of Economics*, 35, 587-612.

Hicks, J.R. (1937). Mr. Keynes and the classics, *Econometrica*, 5, 147-159.

Hicks, J.R. (1950). *A Contribution to the Theory of the Trade Cycle*. Oxford, Clarendon.

Kaldor, N. (1955-6). Alternative Theories of Distribution, *Review of Economic Studies* 23, 83-100.

Kaldor, N. (1957). A model of Economic Growth, *Economic Journal*, 67, 591-624.

Kalecki, M. (1971). *Selected Essays on the Dynamics of the Capitalist Economy 1933-1970*, Cambridge: Cambridge University Press.

Lavoie, M. (2014). *Post-Keynesian Economics. New Foundations*. Cheltenham: Elgar.

Lavoie, M. (2016). Convergence Towards the Normal Rate of Capacity Utilization in Neo-Kaleckian Models: The Role of Non-Capacity Creating Autonomous Expenditures, *Metroeconomica*, 67, 172-201.

Mongiovi, G. (2011). Demand, Structural Interdependence, and Economic Provisioning, *American Journal of Economics and Sociology*, 70, 1147-1174.

Palumbo A. (2015), Studying Growth in the Modern Classical Approach: Theoretical and Empirical Implications for the Analysis of Potential Output, *Review of Political Economy* 27, 1-26.

Palumbo, A. and Trezzini, A. (2003). Growth without normal capacity utilisation, *European Journal of the History of Economic Thought,* 10, 109-35.

Palumbo, A. and Trezzini, A. (forthcoming). Overinvestment and depreciation: a reflection on investment decisions in demand-led growth models.

Park, M.-S. (1997). Accumulation, Capacity Utilisation and Distribution, *Contributions to Political Economy,* 16, 87-101.

Park, M.-S. (2012). Alternative routes of adjustment of saving to investment in the long period. In Salvadori, N. and Gehrke, C. (eds.), *Keynes, Sraffa and the Criticism of Neoclassical Theory: Essays in Honour of Heinz Kurz.* Taylor and Francis.

Parrinello, S. (2014). A Search for Distinctive Features of Demand-led Growth Models, *Centro Sraffa Working Papers,* 2.

Pasinetti, L.L. (1962). Rate of Profit and Income Distribution in Relation to the Rate of Economic Growth, *The Review of Economic Studies*, 29, 267-279.

Pasinetti, L.L. (1974). The economics of effective demand. In Pasinetti, L.L. (eds.), *Growth and income distribution*. Cambridge: Cambridge University Press.

Robinson, J. (1962). *Essays in the Theory of Economic Growth*. New York: St Martin's Press.

Rowthorn, R. (1981). Demand, Real Wages and Economic Growth, *Thames papers in political economy*, London.

Serrano, F. (1995). Long-period effective demand and the Sraffian supermultiplier, *Contributions to Political Economy,* 14, 67-90.

Serrano, F., Freitas, F. and Bhering, G. (2019). The Trouble with Harrod: The fundamental instability of the warranted rate in the light of the Sraffian Supermultiplier, *Metroeconomica*, 70, 263-287.

Setterfield, M. (2019). Long-run variation in capacity utilization in the presence of a fixed normal rate, *Cambridge Journal of Economics,* 43, 443-463.

Skott, P. (2019). Autonomous demand, Harrodian instability and the supply side, *Metroeconomica,* 70, 233-246.

Smith, M. (2012). Demand-led Growth Theory: A Historical Approach, *Review of Political Economy*, 24, 543-573.

Smith, M. (2018). Demand-Led Growth Theory in a Classical Framework: Its Superiority, Its Limitations, and Its Explanatory Power, *Centro Sraffa Working Papers*, 29.

Sraffa, P. (1960). *Production of commodities by means of commodities*. Cambridge: Cambridge University Press.

Storm, S. (2017). The New Normal: Demand, Secular Stagnation and the Vanishing Middle-Class, *Institute for New Economic Thinking*, Working Paper, 55.

Thirlwall, A.P. (2011). Balance of payments constrained growth models: history and overview, *PSL Quarterly Review*, 64, 307-351.

Trezzini, A. (1995). Capacity utilisation in the long-run and the autonomous components of aggregate demand, *Contributions to Political Economy*, 14, 33-66.

Trezzini, A. (1998). Capacity utilization in the long run: some further considerations, *Contributions to Political Economy*, 17, 53-67.

Trezzini, A. (2013). The Meaning of Output Trends in the Analysis of Growth. In Levrero, E.S., Palumbo, A. and Stirati, A. (eds.), *Sraffa and the Reconstruction of Economic Theory,* vol. II. Palgrave Macmillan.

Trezzini, A. (2017). Harrodian Instability: A Misleading Concept, *Centro Sraffa Working Paper,* 24.

Trezzini, A. and Palumbo, A. (2016). The theory of output in the Modern Classical Approach: main principles and controversial issues, *Review of Keynesian Economics*, 4, 503-522.

Vianello, F. (1985). The pace of accumulation, *Political Economy*, 1, 69-87.

Vianello, F. (1989). Effective Demand and the Rate of Profits: Some Thoughts on Marx, Kalecki and Sraffa. In Sebastiani, M. (eds.), *Kalecki's Relevance Today.* London: Macmillan.

Vianello, F. (2007). Reviewing a Review, *Review of Political Economy*, 19, 249-261.

White, G. (2006). Demand-led growth and the classical approach to value and distribution: Are they compatible? In Salvadori, N. (eds.), *Economic Growth and Distribution: On the Nature and Cause of the Wealth of Nations*. Cheltenham: Elgar.

Chapter 10

What are the risk drivers in the banking system? The fundamental betas of European banks

Daniela Venanzi

10.1 Introduction

In-depth understanding of bank risk is important for a range of financial market participants. It is of interest of regulation and supervisory authorities, who are responsible for maintaining the financial system stability: understanding which are the determinants of bank risk is crucial in order to better address country economic and financial policies as well as decision making of industry operators. Furthermore, it is of interest of the financial market operators (banks, investors, etc.), because the most of their decisions are influenced by the determinants of bank risk. It is enough to think of the relevance of estimating the bank cost of capital, which depends on systematic risk, in order to assess, for example, if bank profitability is adequate (compared to the return requested by the risk borne), or for estimating the cost of capital in M&As or asset management operations[1].

The international literature on bank risk is very wide-ranging. The various strands connect the total risk of bank (mainly, credit risk and bankruptcy risk) to different categories of determinants, among which bank characteristics, regulation policies, industry competition, deposit guarantee framework, etc.

This study, however, assumes a different perspective, partial and

---

[1] Recent studies show how are widespread (in Italy) worst practices in estimating beta for M&As among banks (Venanzi 2016).

focused on a particular measure of a bank's systematic risk. In fact, the object is to empirically estimate a model that explains the beta of the European commercial bank stocks by means of a set of economic and financial fundamentals, provided by their financial statements. Therefore, we want to explain the systematic risk (neither specific risk nor total one) of banks and specifically, a measure of this risk, that is beta. By utilising a sample of more than 100 European banks, whose main activity is the traditional financial intermediation business, in the 2006-2015 decade, we use the theoretical model of the fundamental beta, previously formulated for non-financial firms by the pioneering study of Beaver *et al.* (1970), which seems to be very meaningful for financial companies for the following reasons:

•   Capital Asset Pricing Model (hereinafter CAPM) is a pricing model that well functions in explaining the returns of bank stocks (Damodaran 2009);

•   the bank stock market (bank stocks are often leading in stock exchanges) is highly influenced by rumours and speculative behaviour which might distort the observed market prices, and therefore estimated historical betas; in addition, expectations about regulation and supervisory policies are further factors of return volatility;

•   therefore, given above, if beta is a good measure of a bank's systematic risk, it doesn't seem to be correctly estimated by regressing past market returns (in excess of the risk-free rate) in function of the corresponding excess returns of the market portfolio proxy, since stock market prices could be biased.

The fundamental beta is, therefore, an alternative measure, which estimates the *true beta* as a function of bank economic and financial fundamentals and cleans up the historical beta from errors.

This study contributes to extend the empirical findings on European banks (Baele *et al.* 2007; Haq and Heaney 2012), that appear more limited than those regarding US banks (Leung *et al.* 2015; Stiroh 2004, 2006; Stiroh and Rumble 2006; Stever 2007).

Moreover, the aim is also to verify the impact of size and business diversification on systematic risk: if, as we expect, both increase the

bank's risk, the empirical findings should lead operators and regulators to change their currently dominant attitude in favour of both M&As among banks and a shift of the bank's business towards investment activities, that are riskier, to the disadvantage of traditional lending activity to the real economy, with obvious implications on moral hazard behaviour from bankers for the well-known 'too big to fail'[2].

This study also wants to verify (see below) if the Basel coefficients stated for risk weighting the bank assets are actually able to measure risk correctly.

10.2 The theoretical and empirical framework

10.2.1 Measuring the systematic risk: the beta

Given the portfolio diversification theory, the relevant risk (that is the risk that needs to be remunerated by market) is the systematic risk, measurable in the modern finance theory from beta, which measures how the excess returns of a single stock or portfolio is sensitive to the variance of the excess returns of a well-diversified portfolio, that is an appropriate proxy of the market risk. Beta (Sharpe 1964; Lintner 1965), therefore, is the coefficient of the time-series linear regression among the past stock excess returns and the corresponding excess returns of the market portfolio proxy, as follows:

$$\left(R_i - R_f\right) = \alpha_i + \beta_i * \left(R_{mkt} - R_f\right) + \varepsilon_i \qquad (10.1)$$

$$\beta_i^{mkt} = \frac{cov_{i,mkt}}{\sigma_{mkt}^2} = \frac{\sigma_i * \sigma_{mkt} * corr_{i,mkt}}{\sigma_{mkt}^2} = \frac{\sigma_i * corr_{i,mkt}}{\sigma_{mkt}} \quad (10.2)$$

---

[2] If a large bank, with a complex structure, is experiencing severe distress conditions, its consolidation process increases the probability that liquidation/restructuring results to be more difficult or implemented in disorder. Since this kind of financial intermediaries implies that their problems could generate large and widespread risks, the concentration process can therefore involve an increase of probability that the distress conditions produce negative implications for the overall system.

Market portfolio beta is obviously equal to 1 and therefore, if a stock has beta larger than 1, it means that its returns vary more proportionally than market portfolio returns (*aggressive stock*), and, if not, the stock is less sensitive to systematic risk (*defensive stock*).

There are many criticisms about beta:

• *ex post* using of past returns implies (rather questionably) that the realised returns were consistent with the future expectations that investors had before the returns took place;

• return frequency adopted in the regression can distort beta estimation. In practice, monthly returns are used, in a 5-year period (60 monthly returns in total). Daily returns would increase the observation number and therefore shorten the time span of the estimation (limiting the bias of not respecting the coeteris paribus assumption), but they could under-evaluate beta of less liquid stocks as well as over-evaluate the more liquid ones, that are often target of speculative behaviours;

• the choice of an appropriate proxy of market portfolio is a critical aspect, since, according with Roll (1977), the tests of the CAPM (Black *et al.* 1972; Fama and MacBeth 1973; Fama and French 1992) must be interpreted with great caution. In fact, they merely imply that the market index that was selected was ex post efficient, but they do not prove that the true market portfolio is ex post efficient, but unfortunately, because the market portfolio contains all assets, marketable and nonmarketable, it is impossible to observe;

• betas change over time as far as firms change (Damodaran 1999): *i)* firms divest current businesses and invest in new businesses or acquire new firms. This process changes their business mix and therefore their beta; *ii)* they change their financial leverage, increasing or decreasing debt. In addition, decisions like paying dividends or buying-back shares change the financial leverage; analogously, variations in market value of equity or debt can cause relevant changes in financial leverage, also in short time; *iii)* more generally, firms tend to grow over time and in the same time their operating cost structures change, causing the change of their betas;

• multi-factor models (Fama and French 1992, 1993) undermined the CAPM/beta validity, showing that there are systematic risk

components other than beta, in particular the *size-effect* (small firms have higher betas but also, coeteris paribus, stock returns higher than returns requested by CAPM) and the *book-to-market-value* premium (high book-to-market-value firms are considered riskier because with low growth perspectives, less profitable and low dividend paying): therefore, two more factors of systematic risk.

10.2.2 The fundamental beta

Many studies (Beaver *et al.*, 1970; Bildersee 1975; Eskew 1979; Jarvela *et al.* 2009) affirm that corporate financial statements contain data and information that can be used for measuring risk. The question is: are risk measures based on accounting data related to risk measures based on market data, in particular to beta? If stock market prices reflect the firm fundamentals, then these fundamentals could be used to explain different betas among stocks. Therefore, it is very relevant to know which fundamentals affect beta more significantly, both to orient decision making in terms of risk implications, and to utilise them for estimating stock/portfolio beta, given the many criticisms that undermine beta estimation based on historical market data.

The fundamental beta approach estimates beta based on identifying the main drivers of systematic risk. This approach was firstly introduced by Beaver, Kettler and Scholes (BKS) study in 1970, which oriented many subsequent studies.

The study is based upon an analysis of 307 firms listed at the New York Stock Exchange (NYSE) in the period 1947-1965. The nineteen-year period was further divided into two subperiods of ten years (1947-1965) and nine years (1957-1965), respectively. The partitioning of the total time period will permit an analysis of the stationarity of the relationships over time and an examination of the ability of accounting data to forecast into a future period.

Firstly, the authors used time series regressions for ex post empirical estimate of systematic risk (stock excess returns versus NYSE index excess returns); a separate regression was computed for each security and for each subperiod: therefore, 307 regressions were

computed for each of the two subperiods, resulting in a total of 614. Secondly, BKS identified the relevant accruals able to explain systematic risk by analysing correlations among betas and accounting fundamentals (correlations were also conducted at the portfolio level, since the portfolio, rather than the individual security, is the relevant decision-prediction entity for investors).

Finally, BKS utilised the accounting data as instrumental variables in forming estimates of beta in period one that will reduce or eliminate the errors in the observed historical beta. This was directed to compare the ability of accounting risk measures in period one (the fundamental beta) to forecast the market-determined risk measure (historical beta) in period two.

Historical beta (β_H) is an estimate subjected to the error (ω) of *true beta* (β_T), which we cannot observe directly:

$$\beta_H = \beta_T + \omega \qquad (10.3)$$

The instrumental variables approach states that, although the true beta may be directly unobservable, it is linearly related to n observable variables, z_1 through z_n (called instrumental variables):

$$\beta_T = \varphi_0 + \varphi_1 * z_1 + \ldots + \varphi_n * z_n \qquad (10.4)$$

where z_i are the accounting fundamentals and φ_i are the sensitivities of the true beta to these variables. Analogously, we can estimate from the following cross-section linear regression equation the sensitivities of historical beta (which is observable) to the instrumental variables:

$$\beta_H = c_0 + c_1 * z_1 + \ldots + c_n * z_n + \omega \qquad (10.5)$$

The error term ω reflects error in β_H.

Therefore, removing the error (ω) from β_H we obtain the estimate of true beta, that is the fundamental beta (β_F):

$$\beta_F = \beta_H - \omega \qquad (10.6)$$

The multiple correlation coefficient in BKS implied an R^2 (a measure of the explaining power of the model) of about 45%. An extremely low R^2 would probably indicate that the wrong instruments were chosen. On the other hand, extremely high correlation would result in a fundamental beta essentially equal to the historical beta, which would defeat the purpose of attempting to remove measurement errors in the last.

Finally, the ability of both β_H and β_F in period one to forecast the market-determined risk measure in period two was analysed. The following relationship between the *true betas* of the two sub-periods is assumed:

$$\beta_2 = \delta_{10} + \delta_{11} * \beta_1 \ where \ \delta_{10} = 0 \ and \ \delta_{11} = 1 \qquad (10.7)$$

Beta is assumed to be stable along the period: this hypothesis is confirmed by analysing the correlation between historical betas in the two sub-periods considered, at portfolio level.

BKS empirical findings reveal the consistently superior performance of the instrumental variables approach in forecasting risk measure than the historical beta: the mean of the squared errors as well as the mean of the absolute value of the errors are consistently larger for the naive model (historical beta). The margin of superiority increases at the portfolio level (61 portfolios of 5 securities each): in order to form the portfolios, the securities were ranked according both to the magnitude of historical beta and instrumental beta. In all cases, the fundamental beta has a better forecasting ability than historical beta (the mean absolute error is about half). Moreover, the instrumental model had a lower error than the naïve model in the tail areas (operationally defined to be the upper and lower deciles at the individual level and the upper and lower quartiles at the portfolio level) and this is very useful, since they are probably the areas where accurate forecasts are most needed.

However, the approach presents some limits. Firstly, in the cross-

section regression, it should be better to use homogeneous firms, for example belonging to the same industry, because it seems reasonable that risk drivers vary across sectors: therefore, the instrumental variables set might be industry-specific.

A second criticism concerns the flaws of accounting data: Shan *et al.* (2013) show that the accrual variability might depend not only on risk of innate accounting variables (i.e. fundamentals influencing business risk), but be subject to managerial discretions either to signal private (predictive) information or to manipulate earnings opportunistically. The authors further decompose accrual variability into fundamental and discretionary components and examine whether these two components have distinct effects on stock return volatility. The final result is that the effect of the discretionary component on future stock return volatility is substantially lower and economically insignificant; therefore, fluctuations of stock returns mainly reflect a firm's fundamental uncertainty rather than managerial manipulation.

10.2.3 The determinants of beta in non-financial firms

The main drivers of systematic risk emerging from BKS' s study and the related strand are as follows:

•   **dividend payout** (BKS 1970; Eskew 1979; Jarvela *et al.*, 2009), measured as the sum of cash dividends paid out divided by the earnings available for common stockholders. The emerging link is negative: firms with low payout ratios are riskier. The belief can be rationalized by the signalling theory, according to which managers have better information about firm than outside investors and therefore can provide information about firm conditions to the market through the dividend policy. As well known in the international literature (Lintner 1956; Fama and Babiak 1968; Bharati *et al.* 1998), firms follow a policy of dividend stabilization (i.e., firms are reluctant to cut back, once a dividend level has been established), and the payout ratio can be viewed as a surrogate for management's perception of the uncertainty associated with the firm's earnings;

•   **growth** (BKS 1970; Bildersee 1975; Eskew 1979), measured as natural logarithm of the ratio of the terminal asset size divided by the

initial asset size. The expected relationship is positive: in a competitive economy the excessive earnings opportunities of any firm will erode as other firms enter, then it can be argued that these excessive earnings streams are more uncertain (i.e. volatile) than the *normal* earnings stream of the firm. In addition, growth is negatively associated with payout: firms with lower payout ratios, ceteris paribus, will have higher growth rates. Yet it was argued above that low payout implies greater riskiness. If so, then growth rate would be positively associated with risk;

• **leverage**: as debt is introduced, the earnings stream of the common stockholders becomes more volatile (Modigliani and Miller 1958). According to the second proposition of Modigliani and Miller theory, the levered cost of capital of a firm increases as far as the market value of debt divided market equity increases

$$r_e = r_u + \frac{D}{E} * (r_u - r_d) \tag{10.8}$$

where:
r_e = *levered cost of capital*; r_u = *unlevered cost of capital*; r_d = *cost of debt*.

According with CAPM, the relationship between returns lead to the corresponding relation between betas:

$$\beta_e = \beta_u + \frac{D}{E} * (\beta_u - \beta_d) \tag{10.9}$$

• **liquidity** measured by current ratio (current assets divided current liabilities). We expect a negative link with beta: liquid assets or current assets have a less volatile return than noncurrent assets. Larger the liquidity, less probable the bankruptcy. However, liquidity in excess is disadvantageous as far as taxes are concerned and could generate agency costs: in fact, entrenched managers can use large free cash flows inefficiently;

• **size** measured by the natural logarithm of total assets (the log

transformation was used because its distribution more nearly conforms to the properties of symmetry and normality): it is widely believed that larger firms are less risky than smaller firms. In terms of default risk, the evidence indicates that the frequency of failure is lower for the large size classes. Moreover, larger firms are more diversified and if individual asset returns are less than perfectly correlated, larger firms will have lower variance of rate of return than smaller firms. In terms of portfolio theory, however, as long as the investor can diversify out of the individualistic risk, he is indifferent to whether an individual firm is an efficient portfolio in and of itself. Many studies (Gu and Kim 1998; Titman and Wessels 1988) show that the systematic risk of larger firms is less than the smaller since they are able to better face the adverse economic changes and better diversification opportunities. In addition, larger firms can realize scale economies and therefore reduce the incidence of direct bankruptcy costs on company value (Ang *et al.* 1982; Warner 1977). Finally, Fama and French (1992, 1993), found that market returns remunerate a '*small minus big*' premium for systematic risk. We expect a negative impact of size on beta;

• **variability in earnings**, measured (BKS 1970; Bildersee 1975; Eskew 1979; Jarvela *et al.* 2009) by the standard deviation of an earnings-price ratio (i.e., income available for common stockholders to market value of common stock outstanding):

$\sigma_{E/P} = \sqrt{\frac{1}{T}\left[\sum_{t=1}^{T}\frac{E_t}{P_{t-1}} - \left(\frac{\bar{E}}{\bar{P}}\right)^2\right]}$. This variable affects negatively beta;

• **accounting beta**: it can be derived in a similar manner to the market beta, that is from a time series regression with the firm's earnings-price ratio as the dependent variable and some economy-wide average of earnings-price as the independent variable:

$$\beta_{accounting} = \frac{cov\left(\frac{E_t}{P_{t-1}}, M_t\right)}{var(M_t)} \qquad (10.10)$$

where $M_t = (\sum_{i=1}^{N} E_{it}/P_{it-1})/N$. It is positive the expected link, but it for each security will be estimated on a small number of observations,

which implies that estimates will be subject to a large amount of sampling error (earnings are available only yearly).

The cited empirical studies obtain statistically significant results, consistent with the expected signs.

Relevant is the Jarvela *et al.* (2009) study, which aims at verifying if BKS approach was still valid in recent years: they obtain consistent results, except for some variables. In particular, the *dividend payout* does not explain beta for larger companies. This evidence can be explained by the fact that the larger companies, although financially strong and able to pay dividends, can adopt discretionary dividend policies and investors did not evaluate a lower payout negatively (i.e. as insufficient cash flows generated), but as a strategy based on other reasons; on the contrary, for small firms, a low payout is interpreted as a symptom of larger risk, because they are considered more financially vulnerable. In addition, earnings volatility has a very weak impact on beta, although statically significant. The authors explain this evidence through the speculative bubbles that distort E/P ratios in the last years.

10.2.4 The drivers of bank beta

Many recent studies focus on identifying drivers of bank systematic risk. We will analyse the main fundamentals emerging from these studies.

Firstly, **diversification**. Banks are allowed to diversify functionally. From a regulatory perspective, they can combine commercial banking, securities, insurance and other financial activities in a conglomerate organizational form. The European regulatory framework allows a more diversification degree than US banks, longer regulatory constrained. Baele *et al.* (2007) discuss costs and benefits of diversification in term of profitability and risk.

First, the formation of financial conglomerates would be beneficial if there are positive cost and/or revenue effects from combining various financial services activities. Consolidated revenues would be improved if the income-generating capacity of the combined institutions is enhanced. Similarly, the operating costs of financial conglomerates would be lower relative to specialized banks if integration

leads to operational synergies, e.g. through economies of scope. The sharing of inputs such as labour, technology and information across multiple outputs constitutes the major source of such potential cost savings. Second, banks possess information from their lending relations that may facilitate the efficient provision of other financial services, including securities underwriting or insurance. Similarly, information acquired through securities or insurance underwriting can improve loan origination and credit risk management. Thus, financial conglomerates could enjoy economies of information that boost performance and market valuations. Third, the potential for functional diversification may improve corporate governance through the working of the takeover market. When cross-activity mergers are allowed, managers of financial firms incur stronger monitoring by the takeover market.

From the risk dimension, standard portfolio theory predicts that the combined cash flows from non-correlated revenue sources should be more stable than the constituent parts. Securities and insurance activities have the potential to decrease conglomerate risk, but the effect largely depends on the type of diversifying activities that bank holding companies undertake. However, we know that diversification can simply pursue by investors at individual portfolio level.

On the cost side, agency costs may arise due to the complexity of the conglomerate organization. Diversification of activities in a conglomerate structure could intensify agency problems, between insiders and outsiders, but also between the divisions of the conglomerate and between the conglomerate firm and its customers in the form of conflicts of interest. Managers may pursue diversification to enhance their ability to extract private benefits, even when diversification would lower the market value. The question is whether or not internal mechanisms can be designed to align interests or whether external discipline can alleviate some of the agency problems. In addition, on the costs side, regulatory costs associated with multiple supervision can be invoked.

If theoretically it is unclear whether or not the potential benefits of functional diversification are larger than the costs, empirically many studies (Stiroh and Rumble 2006; Stiroh 2004, 2006; DeYoung and Roland 2001; Baele *et al.* 2007; Demirguc-Kunt *et al.* 2010) show a

significant positive link between non-interest income (non-interest income captures all income streams that functionally diversified banks generate by providing a broad array of financial services) and the volatility of market returns or accounting earnings. Diversification generates an increased exposure to non-interest activities, which are much more volatile than interest-generating activities. The above empirical findings show that more diversified banks have a higher exposure to changes in market sentiment (e.g. because of their reliance on investment banking) or economy-wide shocks. As far as the idiosyncratic risk is concerned, evidence from European banks reveals that an increasing reliance on non-interest income decreases a bank's idiosyncratic volatility; however, this relationship is nonlinear. Once a bank becomes too exposed to non-traditional banking activities, its bank-specific risk increases. The impact on bank total risk of diversification would result positive. US bank studies (Stiroh 2006) reveal a linear and positive linkage between diversification and risk, both systematic and total.

Size is another determinant of risk. Differently from non-financial firms, banks' equity betas are positively related to size. Small banks appear to make safer loans than large banks. As a result, individual loans at small banks exhibit less sensitivity to market movements (and to other risk factors) than large bank loans. However, due to small banks' inability to diversify, the total equity volatility of large and small banks is the same (given the high regulatory degree in this industry). This evidence depends partially on the effect of diversification: banks grow through diversifying their activities. Stever (2007) shows, in addition, that small banks may lend to similar sectors and asset types as large banks, but they make loans with lower credit risk. They may require more collateral per loan or have superior information on borrower risk (since small banks have both a smaller number of loans and less groups of firms to which they can lend, they can pursue a better monitoring of their borrowers)[3].

---

[3] ECB official data (ECB 2018), on the contrary, show that in the supervised 110 banks, the average NPL ratio decreases with increasing size (in the first quarter 2018,

Haq and Heaney (2012) observe that the regulatory protection of larger banks could result in large banks becoming 'too big to fail' and this could increase the incentive for large banks to undertake riskier activities (i.e. a moral hazard behaviour), particularly the riskier non-interest generating activities. Large banks could also be more sensitive to general market movements than small banks leading to a positive relation between bank systematic risk and size. Their study, conducted on a sample of 117 European commercial banks (from 15 European countries) in the 1996-2010 period, highlights that size significantly increases systematic and total risk, while decreases the idiosyncratic one. These results are substantially confirmed in Baele *et al.* (2007) study, regarding European banks too.

As far as studies on US banks are concerned (Stiroh 2006 and Leung *et al.* 2015), the size impact is positive on beta, but negative on total risk: it means that, differently from European banks, in US banks the negative impact of size on idiosyncratic risk overcomes the positive one on beta.

Other fundamentals of beta are the following:

1.   **capital adequacy** CET1, measured as Tier 1 (i.e. 'core capital' which consists primarily of common stock, reserves and retained earnings), divided by risk-weighted assets (RWA), based on Basel Accord standard weights, or internal ratings (IRB=*Internal Rating Based*), when banks are authorized to adopt them. It is the basis par-excellence of the micro-prudential supervisory framework of Basel Committee. Haq and Heaney (2012) assume a negative relationship between CET1 and risk (systematic, idiosyncratic and total). Banks generally maintain a capital buffer to absorb losses that arise from their loan portfolio, adjusting the buffer as the risk of

---

from 12.45% in banks with assets less than 30 billion euros to 4.08% in banks with assets larger than 330 billion and to 3.35% in global systemically important banks). These data are, however, biased by country effect, as ECB recognizes itself (since the country mix in each size class differs). In addition, it is an average, weighted to the size (i.e. it is not necessarily representative of the banks in each size class, if banks are very different in size). In my study (Venanzi 2017) on a sample of about 450 Italian commercial banks (using single balance sheets), a statistically significant relationship between size and NPL ratio does not emerge, but smaller banks are more frequent in more virtuous clusters for credit quality.

banks to hold capital to protect them against the cost of financial distress, agency problems and to curtail the risk shifting benefit arising from deposit insurance. However, the impact of capital regulation on bank risk is ambiguous. For example, in an agency problem framework, higher capital standards help to reduce the risk of the bank's assets, however, with the bank issuance of equity to meet the new standards, bank insider effort reduces as their equity stake decreases. Moreover, bank capital regulation suggests that higher capital levels may induce banks to increase asset portfolio risk and the probability of default. They propose a 'U-shaped' relation between bank capital and bank risk, thus reconciling the two opposing views on the effect of bank capital on bank risk. That is, for low levels of capital, as a bank's capital increases, it takes on less risk, reflecting the disciplinary effect of bank capital, but as capital continues to rise, banks eventually reach a point where further increases in bank capital result in increasing risk. The authors argue that this turning point occurs when banks start to take on more profitable, albeit potentially riskier, investments, either because the probability of bank default is very remote or because, in the event of bankruptcy, the bank can shift the cost of default onto the state insurance on deposits (moral hazard problem). Their empirical results support the nonlinear linkage of CET1 with beta. In Leung *et al.* (2015) study on US banks, CET1 decreases the total and idiosyncratic risk, but not the systematic one (the nonlinear relationship is not tested, however);

2.   **off-balance sheet items** (bank guarantees attached to commercial letters of credit, loan commitments and stand-by letters of credit, derivative obligations, etc.). Greater levels of regulation and increased competition have resulted in banks developing non-traditional activities which, while not appearing on the balance sheet, do create contingent assets and liabilities, which are difficult for investors and regulators to be assessed in terms of risk implications. It has proven difficult for investors and regulators to identify the actual level of risk. The off-balance sheet activities of most concern here are the contingent liabilities of the banks where the bank must honour guarantees when required. The theoretical literature assumes that they increase bank risk and empirical findings prove this impact;

3.   **dividend payout**, whose impact on beta is assumed negative, as in non-financial firms (see 10.2.3);

4.   **incidence of non-performing loans**: Leung *et al.* (2015) affirm that banks with stronger risk control had lower non-performing loans. Therefore, the determinant serves as a proxy of bank efficiency in risk monitoring. Empirical evidence supports the expected positive linkage only with idiosyncratic and total risk, but not with systematic one;

5.   **operational inefficiency**, in terms of cost-to-income ratio (i.e. the ratio of all operating expenses as a fraction of the sum of net interest and non-interest revenues) (Baele *et al.* 2007). Better performing banks in terms of superior technology and more skilled management (Baselga-Pascual *et al.* 2015) are perceived less risky by market; in addition, operational efficiency should protect banks from unexpected volatility of profits. No particular effect, however, is expected on systematic risk and empirical evidence supported this assumption.

10.3 The empirical test on a sample of European commercial banks

10.3.1 Tested hypotheses

We empirically tested the following hypotheses on a sample of European commercial banks.

H1. Size increases beta.

Large banks lend more aggressively and extend more credit than small banks, and therefore, as a consequence, on average their loans have a lower success rate.

In addition, banks grow through diversifying their activities and undertaking riskier activities. This behaviour favours moral hazard attitude: large banks become 'too big to fail' and this could increase the incentive for large banks to undertake riskier non-interest generating activities. Moreover banks, becoming larger through diversification, are more interconnected to the whole financial system.

H2. Diversification increases beta

Diversification generates an increased exposure to non-interest activities (financial investments, trading, insurance, etc.), which are much more volatile than interest-generating (more traditional) activities.

H3. Systematic risk decreases with increasing dividend payout

According with the signalling theory (Lintner 1956; Fama and Babiak 1968; Bharati *et al.* 1998), if managers have better information about firm than outside investors, they can provide information about firm conditions to the market through the dividend policy. Therefore, payout ratio can be viewed as a surrogate for management's perception of the uncertainty associated with the firm's earnings: higher payout indicates higher expected earnings and less bankruptcy risk.

H4. RWA (divided by total assets) increases beta

RWA (*risk-weighted assets on total assets*) measure assets weighted by the Basel II-III coefficients: it represents a very important indicator of bank capital adequacy (since supervisory authorities base the regulatory capital requirements by using this figure). Therefore, this ratio should sum up 'in a nutshell' the most important risk drivers for a certain financial activity. Risk-weighting coefficients (fixed by the Basel Committee rules), if appropriately measured, should increase with increasing risk of assets. However, the expected relationship could be of the opposite sign if the risk-weighting framework was biased (as emerging from recent studies in Europe).

In fact, Basel rules could fail in measuring risk and erroneously direct the bankers' decisions. Some studies show that the risk-weighting coefficients penalise exposures to corporates (non-financial firms) in comparison with exposures to governments, banks and central bank institutions, stating for the former risk-weighting coefficients higher than the latter, with the same creditworthiness. It seems an inexplicable choice, since it underestimates the enhancing systemic effect that bank or government defaults/bankruptcies generate, if compared to similar events concerning non-financial firms. This distortion, highlighted by many studies (Angelini 2016), has been confirmed by the following findings

216

from Mediobanca-Ricerche & Studi annual report on international banks (Mediobanca-R&S 2014; Barbaresco 2015): counter-intuitive positive linkage of RWA/total assets ratio with loans on asset incidence, on the one hand, and negative with derivatives incidence on tangible net worth, on the other. These findings show that the Basel risk-weighting coefficients penalise customer loans in comparison with other assets (including derivatives). Furthermore, the Basel II-III framework boosts internal rating (i.e. measured by banks) in comparison to the standard rating system, assigned by rating agencies. Regulatory capital requirements could be less tightening in the former. It is important to consider that a self-regulation mechanism (by means of internal ratings) or a power delegation to rating agencies (by means of standard ratings) are introduced in this way. They represent factors of further risk, which depends on suitability of the utilised models by delegated parties and related fiduciary relationships in a context already weakly based on reliance; moreover, they are very complex statistical models, for which effective monitoring/validation by regulators is likely to be very difficult. Internal ratings would undervalue bank risk, as recent studies by supervisory authorities show (Behn *et al.* 2016; Cannata *et al.* 2012).

H5. Capital adequacy lessens systematic risk

According with Haq and Heaney (2012), a higher Tier1 (the main component of equity capital) functions as a capital buffer for absorbing potential future losses and reducing distress costs (through less debt). In addition, higher net worth means less agency costs.

Differently from previous studies, here we don't measure capital adequacy using CET1. Firstly, since CET1 is highly correlated with RWA (Pearson coefficient is -38%); secondly, because the assumed distortion in calculating RWA (see above) could also bias CET1 (so explaining why this indicator was not able to distinguish between virtuous and bankrupt banks, in recent crisis).

Here we will use TEXAS ratio (net NPL on tangible equity). Since the sample in this study includes only commercial banks (see infra 10.3.2), whose dominant business is credit intermediation, loans are the most part of assets, credit risk is the main component of total risk and

therefore capital adequacy might be better measured by how equity capital faces up to *shortfalls* deriving by NPL write-down. In addition, among the determinants of beta (see hypothesis 9 below) leverage is included (since it is calculated using Tier1, it can make up for the lack of CET1).

Higher TEXAS ratio increases a bank's systemic risk. Tangible equity capital excludes intangibles, which are assets of uncertain valuation and differ among banks, if they grow internally rather than through mergers (goodwill).

H6. Operational inefficiency does not affect systematic risk

According with Baele *et al.* (2007), better performing banks in terms of superior technology and more skilled management (Baselga-Pascual *et al.* 2015) will be perceived less risky by market, but no impact is expected on systematic risk (otherwise, inefficiency increases a bank's idiosyncratic risk).

H7. Opaque assets (i.e. assets of subjective and doubtful value) increase beta

Bank investment in opaque assets has a stronger impact on bank risk than transparent assets. Asset opacity is not appreciated by market. We assume that the intangibles are opaque because their value written in balance sheet is discretionary. In fact, their value is often calculated by means of models based on subjective estimations, not directly verifiable.

H8. Derivatives increase beta

Derivate assets have high risk (amplified in comparison to underlying assets), and they are opaque in balance sheet value. Derivatives are available for a few banks only (this explains a smaller sample than original). Obviously, derivatives are physiologically used for hedging bank portfolio, but it is impossible to distinguish among hedging and speculative uses.

H9. Leverage amplifies systematic risk

Leverage (total assets on Tier1), not considered by the Basel II framework since considered a rough indicator, showed to be more

effective than more common indicators (for example in comparison with CET1) in forecasting bank distress in recent crisis (BCBS 2014). In fact, Basel III re-introduces it among indicators to be controlled.

However, bank leverage has a different meaning than in non-financial firms. Many studies on fundamental beta of banks do not include leverage among regressors, or, when they do, statistically non-significant coefficients are obtained: see, for example, Haq and Heaney (2012) and Leung *et al.* (2015).

Moreover, leverage might affect idiosyncratic risk more than systematic one: if a bank monitors total risk, the leverage impact on systematic risk could be negative, since higher systematic risk can induce strategies that don't enhance leverage. In addition, if leverage affects returns, its impact on beta might be distorted by beta-return relationship; or, if banks become larger through increasing leverage, leverage effect on beta can be absorbed by size impact (Bhagat *et al.* 2015)[4].

10.3.2 The sample

The dataset consists of 149 listed banks from 17 European countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, and the United Kingdom) in 2006-2015 period (it is a balanced sample: 10 year observations for each bank): a sample of 1490 bank-year observations, in total.

From the universe of commercial listed banks (data provider is Thompson Reuter Eikon - Datastream Equities and Worldscope Fundamentals) we selected only banks with the first SICcode (i.e. the dominant business by revenues) equal to 6029 (*Commercial Banks, NEC*) or 6022 (*State Commercial Banks*) or 6035 (*Savings Institutions, Federally Chartered*), institutions that offer similar commercial banking services (in detail, 1400 bank-year observations with the first code, 20 for the second e 70 with the third).

However, the sample homogeneity in terms of dominant

---

[4] ECB studies/statistics (ECB 2019) highlight a positive linkage between leverage and size.

business does not exclude other business lines in bank activity[5].

Further detail on the sample is provided in Tables 10.1, 10.2 and 10.3.

The sample is very variegated in terms of size, as dispersion coefficients show (see Table 10.2). On average, assets amount to 263 billion euros, and the half of sample has assets lower than 23.5 billion. Even though sample includes only listed banks, there is not lack of smaller ones: the 5% of distribution has assets less than 2 billion and employees less than 62.

Table 10.1 – *Sample by country*

Country	Sample bank number	Datastream bank number
Austria	6	7
Belgium	2	8
Denmark	21	22
Finland	2	5
France	18	18
Germany	9	15
Greece	5	6
Ireland	2	3
Italy	16	16
Netherlands	2	3
Norway	20	23
Poland	10	15
Portugal	2	2
Spain	6	8
Sweden	4	4
Switzerland	17	27
United Kingdom	7	9
Total	**149**	**191**

Table 10.3 shows the weight on assets of credit intermediation activity: the sample is very homogeneous in terms of dominant business,

---

[5] For example, the following SIC codes: 6211 (*Security Brokers, Dealers, and Flotation Companies*) for 33 banks, 6282 (*Investment Advice*) for 30 banks, 6311 (*Life Insurance*) for 20 banks, etc.

as two dispersion coefficients show. The traditional business of collection of savings and lending is dominant, consistently with their profile of commercial banks: if we exclude the distribution tails, loans to customers are at least 40% of assets and deposits about 25%; in at least a half of banks, 75% and over 50%, respectively.

Table 10.2 – *Sample by size*

	Total assets (billion euros)	*Number of employees*
mean	263.5	17,999
median	23.5	1,838
5° percentile	2.0	62
95° percentile	1,717	119,530
coefficient of variation	2.62	2.35
interquartile range/median	4.27	6.52

Table 10.3 – *Sample by incidence of credit intermediation activity*

	Loans on total assets	*Deposits on total liabilities*
mean (%)	72.2	52.0
median (%)	75.6	52.7
5° percentile (%)	37.9	23.2
95° percentile (%)	90.8	80.6
coefficient of variation	0.23	0.34
interquartile range/median	0.28	0.52

10.3.3 Variables, tested model and statistical methodology
Datastream beta is yearly calculated, by using time series regressions of 60 monthly logarithmic returns (in a 5-year period) of each bank with respect to local index returns (e.g. of the corresponding listing country stock market) (if multi-listed stocks, the first listing market is considered).

Table 10.4 shows the determinants of beta, highlighting the related methods of calculation.

In addition to variables discussed in section 10.3.1, some interactions among variables are also considered:
•   between RWA intensity and *dummy rating*, respectively

dummy_IRB (equal to 1 when bank asstes are higher than 50 billion euros) and *dummy_STD* (equal to 1 when assets do not overcome 50 billion). This dummy should distinguish banks that utilise internal ratings from banks that adopt standard ratings. This is a proxy, that is based on size for distinguishing rating system for calculating RWA, because effective data about adoption are not available (the size threshold is derived by a limited sample of banks for which the adopted rating system was known). We want to verify if the counter-intuitive relationship between beta and RWA might depend on adoption of one or the other risk-weighting system. Behn *et al.* (2016) show that the *probability of default* (PD) of banks utilising IRB system are lower than those utilising standard system. Recently, also the Basel Committee highlighted the pitfalls of IRB system, introducing appropriate correction mechanisms in the regulatory framework;

• between PAYOUT and TEXAS ratio, on the one hand, and *dummy_CRISIS*, on the other, to verify if the impacts have differed during the years of the recent financial crisis (2008-2010). For example, we can assume that, according with signalling theory, dividend distribution can affect beta more strongly, since it is a more credible signal during crisis, or alternatively, the impact of NPLs on risk could be enhanced.

We include country dummies (Italy as reference basis), as fixed effects, and the variable *GDP_index* (index number of GDP, per country-year) which serves as proxy of both time effect (probably better than fixed effect) and country economic scenario (in terms of cycle and inflation).

The tested model is the following:

$$\beta_{i,j,t} = \alpha + \varphi * [DETERMINANTS_{i,j,t}]$$
$$+ \varphi_{RWA} * dummy\_rating_{i,j,t}$$
$$* RWA_{i,j,t} + \varphi_{TEXAS} * dummy\_crisis_{i,t} * TEXAS_{i,j,t} + \varphi_{PAYOUT}$$
$$* dummy\_crisis_{i,t} * PAYOUT_{i,j,t} + \alpha_P * COUNTRY_j$$
$$+ \alpha_{P,T} * GDP\_index_{j,t} + u_{i,j,t} \tag{10.11}$$

α and φ parameters are respectively the intercept and the coefficient vector of k determinants of beta, $[DETERMINANTS]$ is the matrix of the assumed determinants of bank systematic risk. $u_{i,j,t}$ is the term of error.

The model was tested using *pooled* OLS[6] (from GRETL package): error estimation (heteroskedastic and auto-correlated in series) uses HAC methodology (*Heteroskedasticity and Autocorrelation Consistent*) and therefore can be considered robust (Arellano 2003).

Table 10.4 – *Determinants of beta*

determinants	symbol	calculation
size	SIZE	ln (total assets)
diversification	DIV	noninterest revenues/total revenues
dividend payout	PAYOUT	paid dividends/net income
operational inefficiency	INEFFICIENCY	(operating costs – provisions for credit losses)/total revenues
opacity of assets	OPACITY	intangibles/total assets
derivatives on total assets	DERIVATIVES	derivatives/total assets
risk-weighted assets intensity	RWA	risk-weighted assets/total assets
leverage	LEVERAGE	total assets/TIER 1
texas ratio	TEXAS	non-performing loans (net of related provisions)/tangible equity capital
proxy internal ratings	dummy IRB	=1 if total assets > 50 md
proxy standard ratings	dummy_STD	=1 if total assets ≤ 50 md
dummy crisis years	dummy_CRISIS	=1 for years 2008, 2009 and 2010
GDP index number	GDP_index	basis 2005, nominal values
dummies country	DCOUNTRY_Austria…..DCOUNTRY_UK	=1 if belonging to country

10.4 Results

10.4.1 Descriptive statistics
Table 10.5 summarizes the descriptive statistics of beta, Table 10.6 the correlation matrix and Table 10.7 the descriptive statistics of beta determinants. All are referred to the whole sample. Beta variance seems

---

[6] We want to estimate a model which can explain both longitudinal and cross-sectional variability of beta.

to be appropriate (as shown by two dispersion indicators) for estimating the effects of determinants.

Table 10.5 – *Descriptive statistics of beta*

mean	0.77
median	0.68
5° percentile	0.04
95° percentile	1.89
coefficient of variation	0.80
interquartile range/median	1.38

Table 10.6 – *Matrix of correlations*

	BETA	SIZE	DIV	PAYOUT	INEFFICIENCY	OPACITY	RWA	DERIVATIVES	TEXAS	LEVERAGE
BETA	1	.618**	.173**	-.199**	.171**	.339**	-.372**	.300**	.217**	-0.043
SIZE		1	.068**	.080**	.056*	.249**	-.430**	.337**	0.015	.156**
DIV			1	.060*	0.046	.337**	-.214**	.081**	0.034	-.266**
PAYOUT				1	-.201**	0.033	-.139**	-.096**	-.230**	.177**
INEFFICIENCY					1	.083**	-.233**	.075*	.159**	.203**
OPACITY						1	-.123**	.083**	.101**	-.126**
RWA							1	-.242**	-0.024	-.152**
DERIVATIVES								1	-0.05	.060*
TEXAS									1	.242**
LEVERAGE										1

*** sign = 0.01 (two tails) * sign = 0.05 (two tails)*

Table 10.7 – *Descriptive statistics of beta determinants*

	DIV	PAY OUT	INEFFI CIENCY	OPACITY	DERIVA TIVES (%)	RWA	LEVE RAGE	TEXAS	dummy IRB
mean	0.28	0.30	0.78	0.005	1.90	0.57	9.92	0.62	0.36
median	0.26	0.29	0.77	0.002	0.51	0.57	7.99	0.30	0.00
5th percentile	0.08	0.00	0.61	0.000	0.00	0.22	2.02	0.03	0.00
95th percentile	0.57	0.81	0.95	0.025	7.19	0.88	23.5	2.05	1.00
coefficient of variation	0.53	0.84	0.15	1.69	2.57	0.41	1.04	1.82	1.33
interquartile range/median	0.57	1.64	0.16	3.12	2.74	0.44	0.70	1.74	

Preliminary findings emerge from the correlation matrix, that confirm some formulated hypotheses: positive impact on beta of size, diversification, derivatives, asset opacity and NPL incidence. Some evidence confirms also the suspected distortions in risk-weighting of

assets, based on Basel framework. In fact, we can observe that RWA is negatively correlated to beta as well as derivatives incidences: correlation signs that appear counter-intuitive. We can see, furthermore, that large banks have more derivatives and diversification causes more opacity of assets (diversification is likely to induce acquisitions with goodwill). In addition, dividend payout is constrained by incidence of both NPLs (negative correlation of PAYOUT with TEXAS) and operational costs (negative correlation with INEFFICIENCY).

We cannot make a trend analysis on descriptive statistics of the determinants of beta, since data of some years are not available and, therefore, sample mix is not homogeneous over the years.

From Table 10.7 we can see that, on average, noninterest revenues are about a quarter of total; omitting distribution tails, we have a range of variation from 8% to 57%, that confirms the differences among the banks in the sample, in terms of relevance of investment assets, besides the dominant traditional business of credit intermediation.

PAYOUT is on average equal to 30%, the incidence on income of operational costs is 78%, derivatives are only 2% of assets, but in UK and German banks they reach average values four times greater and in South European countries (i.e. Italy, Greece, Portugal and Spain) assume lower values.

As far as RWA intensity is concerned, sample average is equal to 57%, with higher values in South European countries. Median value of leverage is about 8, which means a Tier1 equal to 12.5% of total assets.

Finally, TEXAS ratio has an average value of 60% (but the median value is half of mean); the indicator is higher than 100% (that means a shortfall of equity capital in case of write-off of the NPLs) in Greece, Ireland and Italy (for brevity, descriptive statistics per country are omitted).

10.4.2 Regression results

Table 10.8 sums up the regression results.

We tested different models, including various groups of determinants. The final sample (due to data availability) includes 112 banks (the 5 models in the table are comparable, because they use the same observations).

Results reveal that size, diversification, derivatives and NPL incidence increase a bank's systematic risk, confirming hypotheses 1, 2, 5 and 8.

Dividend payout, on the contrary, decreases beta (according to hypothesis 3), signalling to the market about expected earnings; the signal seems to be stronger in crisis years: in fact, in model 2, the coefficient of interaction variable dummy_CRISIS * PAYOUT is statistically significant.

Consistently with our tested hypotheses and international empirical evidence, beta is not affected by operational inefficiency.

Asset opacity never shows a statistically significant impact on beta, although the sign of relationship is as expected. Therefore, hypothesis 7 is not verified. However, the reason could be twofold: on the one hand, correlation matrix (Table 10.6) shows a strong positive relationship of opacity with size and diversification, proving that these two determinants absorb the opacity effect on beta; on the other hand, the proxy used is weak, since intangibles are too generical and non-analytical category, since they can include many different components (data on detail are not available).

Table 10.8 – *Regression results (pooled OLS)*

	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5
	coefficient	coefficient	coefficient	coefficient	coefficient
constant	−0.674704	0.0236626	−1.62594 ***	−0.357558	−1.30935 **
SIZE	0.153253 ***	0.1104 ***	0.136333 ***	0.167542 ***	0.125416 ***
DIV	0.457187	0.551611 *	0.727826 ***	0.30486	0.741111 ***
PAYOUT	−0.529938 ***	−0.423695 ***	−0.251153 ***	−0.510581 ***	−0.237620 ***
INEFFICIENCY	0.032376	0.179364	0.0236334	0.261173	0.049607
OPACITY	4.34145	309.878	5.18250	1.5754	4.64789
RWA	−0.329487 **		−0.344415 **	−0.476150 ***	
TEXAS	0.130673 ***	0.117336 ***	0.09144 ***	0.133557 ***	0.091237 ***
DERIVATIVES	0.012236 *	0.0137032 **	0.01495 **	0.0121203 **	0.014859 **
LEVERAGE	−0.00501 **	−0.005500 **	−0.00243 *		
GDP_index	−0.009067 ***	−0.009689 ***	0.0005646	−0.008995 ***	0.000218
dummy_IRB*RWA		−0.132633			−0.320366 *
dummy_STD*RWA		−0.515560 **			−0.407789 **
dummy_CRISIS*PAYOUT		−0.188305 **			−0.058716
dummy_CRISIS*TEXAS		0.0369362			0.012436
DCOUNTRY_Austria			0.050076		0.0624145
DCOUNTRY_Belgium			0.78158 ***		0.800342 ***
DCOUNTRY_Denmark			−0.154755		−0.140595
DCOUNTRY_Finland			−0.495281 ***		−0.491875 ***
DCOUNTRY_France			−0.133565		−0.105568
DCOUNTRY_Germany			−0.151470		−0.122210
DCOUNTRY_Greece			0.361257 ***		0.357091 ***
DCOUNTRY_Ireland			1.48462 ***		1.47753 ***
DCOUNTRY_Netherlands			−0.111722		−0.088375
DCOUNTRY_Norway			−0.294759 ***		−0.269803 **
DCOUNTRY_Poland			0.17217 *		0.172561 *
DCOUNTRY_Portugal			0.323187 ***		0.346641 ***
DCOUNTRY_Spain			−0.014855		−0.01036
DCOUNTRY_Sweden			−0.192124		−0.160551
DCOUNTRY_Switzerland			−0.305591 **		−0.266471 *
DCOUNTRY_UK			0.120591		0.140141
ln(LEVERAGE)				−0.23572 ***	−0.06078
adjusted R-squared	0.4994	0.5149	0.6833	0.5181	0.6831

*** sign=0.01 ** sign=0.05 * sign=0.10   DPAESE_Italy basis omitted

In contrast to hypothesis 9, leverage does not influence beta (see model 5), even when the logarithmic transformation is used – ln(LEVERAGE) – to linearise the relationship; however, the linkage is significant in models from 1 to 4, but the sign is opposite to what is expected: we can explain this evidence highlighting that leverage affects bank idiosyncratic risk and, therefore, if a bank monitors its total risk, when beta increases, the bank also reduces its total risk by means of leverage; the linkage, therefore, could be mediated by a third omitted variable and then of opposite sign (and reverse causal link). In addition, as correlation matrix shows (Table 10.6), leverage is positively correlated to size and therefore its impact on beta could be absorbed by the latter.

In addition, we can see (by comparing model 3 to 1 as well as model 5 to 4) how the impact of leverage on systematic risk is likely to be absorbed by the country effect: in fact, when country dummies are introduced, leverage coefficient becomes less significant; statistical data from ECB (ECB 2019) confirm a country characterization of leverage, and this fixed effect (i.e. structural effect), *time-invariant*, would be stronger in the model in comparison with time-varying values of leverage.

Finally, as discussed above, bank leverage has not the same meaning than in non-financial firms. Many studies on fundamental beta of banks do not include leverage among regressors, or, when they do, statistically not-significant coefficients are obtained: see, for example, Haq and Heaney (2012) and Leung *et al.* (2015).

The impact of RWA intensity on beta is counter-intuitively negative, which means that banks, that are perceived by the market as systematically riskier, present an RWA/total assets ratio lower and, conversely, banks with higher RWA intensity are perceived as less risky. This evidence confirms distortions of Basel risk-weighting framework, already discussed. When we distinguish by kind of model adopted (models 2 and 5), bias seems larger for banks adopting a standard system: in fact, the negative coefficient of the interaction variable dummy_STD*RWA is larger and more statistically significant.

The variable GDP_index, in regressions where it is statistically significant (where country fixed effects are omitted) negatively affects bank beta: in growing economies, market risk is lower (as previous

empirical findings confirm). However, its impact is absorbed by country fixed effects; the latter show (in comparison to Italy, used as benchmark) a lower beta, on average, in North Europe countries (Finland and Norway) and in Switzerland, and a higher beta in Belgium, Greece, Ireland, Poland and Portugal (i.e. more volatile countries): significant coefficients of country dummies must be interpreted as corrective of model intercept, which holds for Italian banks (omitted dummy) and for countries with insignificant coefficients of respective country dummy.

The apparently weak relevance of country's dummies (8 among 16) might seem a failure of the explanatory model, meaning that economic and political features of country are only weakly relevant in explaining beta. However, we have to consider the following issues: *a)* the sample countries are all members of the EU (all 28 member states), with the exception of Switzerland. Therefore, if we consider the global nature of financial systems, these countries are relatively homogeneous from the perspective of beta; *b)* betas in Datastream are calculated through the well-known time-series regressions between returns of bank stocks and returns of corresponding local market indices; in detail, beta is not an absolute measure of systematic risk, but rather a relative one, that is the stock return sensibility to market index of country whom banks belongs to: therefore, it could be theoretically neutral with respect to geographical differences among sample observations. We mean that the country economic and financial characteristics do affect returns and volatility of bank stocks, as well as other shares included in the market index, but not necessarily (or in a limited manner, anyway) the structural relationship between true beta and bank fundamentals; *c)* if we analyse the residual errors of model 5 (the most complete one), we do not observe higher errors in some countries in comparison to others, in particular not for Switzerland or the United Kingdom, which are countries that are potentially less homogeneous with respect to other sample countries.

The most complete model (model 5) shows a very good explanatory power: it explains more than two thirds of beta variance (both longitudinal and cross-sectional).

This study presents some limits, discussed as follows.

Firstly, the presence of missing values: from an original sample of 149 banks, we arrived at 112 banks, since some determinants (in particular incidence of derivatives and TEXAS ratio) were not available for some banks (however, we always had a minimum of 5 year observations per bank).

Secondly, proxies of some determinants could be measured more accurately, for example asset opacity: off-balance sheet items (according to some empirical studies) might improve the measurement of impact of this determinant on beta.

Thirdly, other explanatory variables of beta could be included among regressors. However, this inclusion could be problematic, since there is collinearity among economic and financial fundamentals of banks. A factor analysis can resolve this problem, by expressing determinants as latent factors, that are linear combinations of observable elementary variables (in this way it is possible to divide the multiple impact of some proxies among different determinants). However, this step can complicate the practical uses of fundamental beta.

10.4.3 Preliminary conclusions

From the empirical test on a sample of more than 100 European commercial banks in 2006-2015 decade, size and diversification of assets (which increases with increasing size) result to increase bank systematic risk. This empirical evidence should suggest that regulators (both European and national) correct their current orientation in favour of mergers and acquisitions among banks as a panacea for all the evils of the banking system, affirming that concentration increases system stability (Venanzi 2018).

As shown by many studies, increasing size incentivizes moral hazard behaviour of bank managers, related to the 'too big to fail' effect and creates a lot of interdependence among larger and complex financial institutions. Managers are encouraged to undertake riskier activities (that could generate much profits), relying on government protection.

Larger size, in addition, generates the following consequences: *i)* makes bank activities more complex and therefore more difficult to assess and monitor risk exposition, from both managers/internal controllers

and supervisory authorities; *ii)* improves interest conflicts of the banking system, since the most of them depend on the presence within the same institution of many and various activities, from commercial (deposit collection and customer lending), under government safeguard, to riskier ones, like asset management and proprietary trading (Walter 2004).

Larger size generally causes a business mix more oriented to activities that are different from traditional credit intermediation (trading, for example), increasing riskier non-interest revenues; the results obtained from this study indicate that diversification positively affects (in a statistically significant manner) a bank's systematic risk.

ECB 2018 Annual Report (ECB 2018) shows that the group of less risky banks (based on SREP classification) among the 119 global systemically important banks (supervised directly) have a more incidence of customer loans on total assets (64% versus 58%) and a lower of investments (14.5% vs 18%) and derivatives (6.7% vs 8.9%) in comparison with the group of banks with medium or high risk.

Moreover, this orientation to the consolidation among banks happens in the current context, characterized by a high concentration degree of European and global banking system, paradoxically improved by the recent crisis and consequent public bailouts (6 mega-mergers after 2007 in USA and 4 in Europe). The current landscape of bank institutions appears very scary for gigantism (Mediobanca-R&S 2014). The main 33 European banks had net assets (excluding derivatives) double than European GDP, on average, in 2004-2013 decade (assets of the main 13 USA banks were equal to 60% of GDP). In Switzerland and the Netherlands, the first banks in 2013 has assets about three times GDP, in France, Spain and United Kingdom from two to one and half, in Italy equal to GDP, in Germany 80%[7]. If we compare the average assets of European banks to those of European non-financial multinationals, the ratio is of 11.4 to 1 (Venanzi 2018).

Finally, from this study there appears evident the failure of Basel

---

[7] Comparing bank net assets to GDP is not correct, since they are not homogeneous figures (fund the former, flow the latter); however, it can be a sign (although rough) of the big size reached by financial intermediaries.

coefficients of asset risk-weighting to correctly measure bank risk: the emerging negative relationship of RWA with beta confirms the bias previously revealed by other European studies.

References

Ang, J.S., Chua, J.H. and McConnell, J.J. (1982). The Administrative Costs of Corporate Bankruptcy: A Note, *The Journal of Finance*, 37, 219-226.

Angelini, P. (2016). Le modifiche del quadro regolamentare e le sfide per le banche italiane, *ABI (Italian Banks Association) Conference 'Unione Bancaria e Basilea 3 – Risk and Supervision'*, Rome 21 June.

Arellano, M. (2003). *Panel Data Econometrics*. Oxford University Press.

Baele, L., De Jonghe, O. and Vennet, R.V. (2007). Does the stock market value bank diversification?, *Journal of Banking & Finance*, 31(7), 1999-2023.

Bhagat, S., Bolton, B. and Lu, J. (2015). Size, leverage, and risk-taking of financial institutions, *Journal of Banking and Finance*, 59, 520-537.

Barbaresco, G. (2015). Zibaldone su banche, Stato e regolamentazione, *ABI (Italian Banks Association) Conference 'La ristrutturazione delle banche e dei sistemi finanziari nell'area Euro'*, Rome 28 January.

Baselga-Pascual, L., Trujillo-Ponce, A. and Cardone-Riportella, C. (2015). Factors influencing bank risk in Europe: Evidence from the financial crisis, *North American Journal of Economics and Finance*, 34, 138-166.

BCBS (2014), *Basel III leverage ratio framework and disclosure requirements.*

Beaver, W.H., Kettler, P. and Scholes, M. (1970). The Association Between Market Determined and Accounting Determined Risk Measures, *The Accounting Review*, 45(4), 654-682.

Behn, M., Haselmann, R. and Vig, V. (2016). The limits of model-based regulation, *European Central Bank Working Paper Series*, 1928.

Bharati, R., Gupta, M. and Nanisetty, P. (1998). Are Dividends Smoothed Signals of Earnings Asymmetry? An Empirical Investigation, *International Journal of Business*, 3(2), 1-18.

Bildersee, J.S. (1975). The Association between a Market-Determined Measure of Risk and Alternative Measures of Risk, *The Accounting Review*, 50(1), 81-98.

Black, F., Jensen, M.C. and Scholes, M. (1972). The Capital Asset Pricing Model: Some Empirical Tests. In Jensen, M.C. (eds.), *Studies in the Theory of Capital Markets*. New York: Praeger.

Cannata, F., Casellina, S. and Guidi, G. (2012). Inside the labyrinth of Basel risk-weighted assets: how not to get lost, *Questioni di Economia e Finanza, Bank of Italy Occasional Papers*, 132.

Damodaran, A. (1999). Estimating Risk Parameters, *NYU Stern School of Business, Digital Archive*, available at https://archive.nyu.edu/bitstream/2451/26789/2/S-CDM-99-02.pdf.

Damodaran, A. (2009). Valuing Financial Service Firms, NYU Stern School of Business, *Digital Archive*, available at http://people.stern.nyu.edu/adamodar/pdfiles/papers/finfirm09.pdf.

Demirguc-Kunt, A., Detragiache, E. and Merrouche, O. (2013). Bank Capital: Lessons from the Financial Crisis, *Journal of Money, Credit and Banking*, 45(6), 1147-1164.

Demirguc-Kunt, A. and Huizinga, H. (2010). Bank activity and funding strategies: The impact on risk and returns, *Journal of Financial Economics*, 98, 626-650.

DeYoung, R. and Roland, K.P. (2001). Product mix and earnings volatility at commercial banks: evidence from a degree of total leverage model, *Journal of Financial Intermediation* 10(1), 54-84.

ECB (2018). *Supervisory Banking Statistics*, third quarter, available at https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisorybankingstatistics_third_quarter_2017_201801.en.pdf.

ECB (2019). *Supervisory Banking Statistics*, fourth quarter, available at https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisorybankingstatistics_fourth_quarter_2018_201904~4b71b5db1c.en.pdf.

Eskew, R.K. (1979). The Forecasting Ability of Accounting Risk

Measures: Some Additional Evidence, *The Accounting Review*, 54(1), 107-118.

Fama, E.F. and Babiak, H. (1968). Dividend policy: an empirical analysis, *Journal of the American Statistical Association*, 63(324), 1132-1161.

Fama, E.F. and French, K.R. (1992). The Cross-Section of Expected Stock Returns, *The Journal of Finance*, 47(3), 427-465.

Fama, E.F. and French, K.R. (1993). Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, 33(1), 3-56.

Fama, E.F. and MacBeth, J.D. (1973). Risk, Return and Equilibrium: Empirical Tests, *Journal of Political Economy*, 81(3), 607-636.

Gu, Z. and Kim, H. (1998). Casino firms' risk features and their beta determinants, *Progress in Tourism and Hospitality Research*, 4(4), 357-365.

Haq, M. and Heaney, R. (2012). Factors determining European bank risk, *Journal of International Financial Markets, Institutions and Money*, 22(4), 696-718.

Jarvela, M., Kozyra, J. and Potter, C. (2009). The Relationship Between Market And Accounting Determined Risk Measures: Reviewing And Updating The Beaver, Kettler, Scholes (1970) Study, *College Teaching Methods & Styles Journal*, 5(1), 1-10.

Leung, W.S., Taylor, N. and Evans, K.P. (2015). The determinants of bank risks: Evidence from the recent financial crisis, *Journal of International Financial Markets, Institutions and Money*, 34(1), 277-293.

Lintner, J. (1956). Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes, *The American Economic Review*, 46(2), 97-113.

Lintner, J. (1965). Security prices, risk, and the maximal risk from diversification, *The Journal of Finance*, 20(4), 587-615.

Modigliani, F. and Miller, M.H. (1958). The Cost of Capital, Corporation Finance, and the Theory of Investment, *The American Economic Review*, 48(3), 261-297.

Roll, R. (1977). A Critique of the Asset's Pricing Theory's Tests: Part I, *Journal of Financial Economics*, 4(2), 129-176.

Shan, Y., Taylor, S.L. and Walter, T.S. (2013). Fundamentals or Managerial Discretion? The Relationship between Accrual Variability and Future Stock Return Volatility, *Abacus*, 49(4), 441-475.

Sharpe, W.F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk, *The Journal of Finance*, 19(3), 425-442.

Stever, R. (2007). Bank size, credit and the sources of bank market risk, *Bank for International Settlements (BIS) Working Papers*, 238.

Stiroh, K.J. (2004). Diversification in banking: is noninterest income the answer?, *Journal of Money, Credit and Banking* 36(5), 853-882.

Stiroh, K.J. (2006). A portfolio view of banking with interest and noninterest activities, *Journal of Money Credit and Banking*, 38(5), 1351-1361.

Stiroh, K.J. and Rumble A. (2006). The dark side of diversification: the case of U.S. financial holding companies, *Journal of Banking & Finance*, 30(8), 2131-2161.

Titman, S. and Wessels, R. (1988). The determinants of capital structure choice, *The Journal of Finance*, 43(1), 1-19.

Mediobanca-R&S (2014). Dati cumulativi delle principali banche internazionali, available at www.mbres.it.

Walter, I. (2004). Conflicts of interest and market discipline among financial service firms, *European Management Journal*, 22(4), 361-376.

Warner, J.B. (1977). Bankruptcy Costs: Some Evidence, *The Journal of Finance*, 32(2), 337-47.

Venanzi, D. (2016). Banche: fusioni e worst practices nella valutazione, *Economia e politica–Rivista on line*, 8(12), available at https://www.economiaepolitica.it/banche-e-finanza/moneta-banca-finanza/banche-fusioni-e-worst-practices-nella-valutazione/.

Venanzi, D. (2017). Banche italiane: che sofferenze!, *Economia e politica-Rivista on line*, 9(14), available at https://www.economiaepolitica.it/banche-e-finanza/moneta-banca-finanza/banche-italiane-che-sofferenze/.

Venanzi, D. (2018). Banche e falsi miti, *Critica Marxista*, 2, 17-28.

This book collects some of the papers presented during the Research Days held at the Department of Economics of Roma Tre University, on May 16 and 17, 2019. Besides two plenary sessions — opening and closing — eight parallel sessions took place, during which as many as 30 contributions were presented. The topics mirrored the main research interests of the Department: micro and macro economic theory; fiscal, environmental and integration policies; personal and functional income distribution; international trade; finance and banking; quantitative methods.

Pasquale De Muro is associate professor of Economic Policy, Roma Tre University.

Saverio M. Fratini is full professor of Economics, Roma Tre University.

Alessia Naccarato is associate professor of Statistics, Roma Tre University.

ROMA
TRE
UNIVERSITÀ DEGLI STUDI