

4

FINANCIAL AND ECONOMIC REVIEW

December 2017
Vol. 16. Issue 4.

The Life Cycle Model of the Fertility Choice
in Hungary

Petra Németh

Possible Paths for GDP Per Capita –
Simulation with a Demographic Growth Model

Éva Berde – Izabella Kuncz

Fiscal Policy and the Business Cycle

György Molnár – Gábor Dániel Soós – Balázs Világi

Trade Credit or Bank Credit? – Lessons Learned
from Hungarian Firms between 2010 and 2015

Dániel Havran – Péter Kerényi – Attila Víg

Explanations of Asset Price Bubbles

Gergely Lakos – Tibor Szendrei

Financial and Economic Review

Scientific journal of the Magyar Nemzeti Bank

Chairman of the Editorial Board:

DÁNIEL PALOTAI

Editorial Board:

TAMÁS BÁNFI, PÉTER ÁKOS BOD, PÉTER CSILLIK, ÉVA HEGEDŰS, DAVID R. HENDERSON,
GYÖRGY KOCZISZKY, LEVENTE KOVÁCS, CSABA LENTNER, DIETMAR MEYER,
KOPPÁNY NAGY, GÁBOR P. KISS, ANETT PANDURICS, MIHÁLY PATAI, RÓBERT SZEGEDI,
RICHÁRD VÉGH, EYAL WINTER

Editor-in-Chief: BARNABÁS VIRÁG

Editor-in-Charge: ENDRE MORVAY

Proofreader: KENDALL LOGAN

Assistant: BERTA DRAPCSIK, ADRIENN TÁLAS, GÁBOR TÓTH

Publisher: Magyar Nemzeti Bank

Publisher in Charge: ESZTER HERGÁR

H-1054 Budapest, Szabadság tér 9.

www.hitelintezetiszemle.hu

ISSN 2415–9271 (Print)

ISSN 2415–928X (Online)

Cover design: MARIANNA IZSÓNÉ BIGAI

© Copyright: Magyar Nemzeti Bank

The papers in this issue contain the opinion of the authors which is not necessarily in line with that of the Magyar Nemzeti Bank.



4



FINANCIAL AND
ECONOMIC REVIEW



December 2017
Vol. 16. Issue 4.

FINANCIAL AND ECONOMIC REVIEW

The address of the Editorial Office: H-1054 Budapest, Szabadság tér 9.

Phone: +36-1-428-2600

Fax: +36-1-429-8000

Homepage: www.hitelintezetiszemle.hu

E-mail: szemle@hitelintezetiszemle.hu

Editorial Staff:

Barnabás Virág Editor-in-Chief, E-mail: viragb@mnb.hu

Endre Morvay Editor-in-Charge, E-mail: morvaye@mnb.hu

Published regularly in every three months.

HU ISSN 2415–9271 (Print)

HU ISSN 2415–928X (Online)

Page setting and printing:

Pauker – Prospektus – SPL consortium

H-8200 Veszprém, Tartu u. 6.

Contents

Vol. 16. Issue 4., December 2017

STUDIES

Petra Németh:

The Life Cycle Model of the Fertility Choice in Hungary 5

Éva Berde – Izabella Kuncz:

Possible Paths for GDP Per Capita – Simulation with a Demographic Growth Model. 36

György Molnár – Gábor Dániel Soós – Balázs Világi:

Fiscal Policy and the Business Cycle 58

Dániel Havran – Péter Kerényi – Attila Víg:

Trade Credit or Bank Credit? – Lessons Learned from Hungarian Firms between 2010 and 2015 86

Gergely Lakos – Tibor Szendrei:

Explanations of Asset Price Bubbles 122

ESSAY

Gusztáv Báger:

Operation of the International Monetary and Financial System – Structural Tensions of a “Non-system” 151

Mónika Kuti – Zsolt Bedő – Dorottya Geiszt:

Equity-based Crowdfunding. 187

Ákos Vajas:

Presentation of the Hungarian Publications of the Magyar Nemzeti Bank Published between 1950 and 1990 from a Visual Communication Perspective 201

BOOK REVIEWS

Gábor Kürthy: The Revolution of Money

(Tamás Bánfi: A pénz forradalma. A pénzteremtés elmélete és gyakorlata) 215

Péter Tóth: Has the US Economy Lost Its Long-term Growth Potential?
(Robert J. Gordon: Rise and Fall of the American Growth) 219

Dániel Papp: Impact of the Financial Crisis on the European
Constitutional System
(Kaarlo Tuori – Klaus Tuori: The Eurozone Crisis: A Constitutional Analysis) 225

CONFERENCE REPORTS

Péter Domokos – Róbert Hausmann – Gábor Szarka:
Report on the Competitiveness and Responsible Corporate Governance
Section Workshops of the Hungarian Economic Association’s 2017
Itinerant Conference. 229

Gábor Horváth: Report on the Digital Currencies International
Conference 239

Erzsébet Czinger: The Future of Old-Age Provision –
Demographic Change, Financial Services and Residential Property.
International Expert Meeting Summary 245

The Life Cycle Model of the Fertility Choice in Hungary

Petra Németh

In this paper, we model the Hungarian fertility choice at the micro level. We outline a life cycle model in which a representative family makes decisions on its number of children, the timing of having children and the mother's return to work after having children over the course of its life cycle. In constructing the model, we incorporate the factors influencing the decision to have children (with a particular emphasis on the family benefits scheme between 2006–2014) and how the behaviours of various household types (in terms of qualifications) differ from each other. According to the simulation results, among the family support tools presented, the family tax allowance scheme introduced in 2011 has a significant impact on the final number of children, maternal age and the timing of childbearing among low and medium education families. The measures introduced in 2014 provide incentive to have three children in all educational groups and foster the early labour market return of mothers. To our knowledge, no similar modelling attempts have been made so far in the Hungarian literature.

Journal of Economic Literature (JEL) codes: C61, D91, H31, J13, J18

Keywords: childbearing, family support, life cycle model

1. Introduction

The objective of this paper is to model the decision to have children in Hungary *at the micro level*, in order to expand the current body of material on this topic. In constructing the model, we incorporated the domestic factors influencing the decision to have children (with a particular emphasis on the family benefits scheme) and how the behaviours of various household types (in terms of qualification) differ from each other. To our knowledge, there have been no similar attempts to date in the Hungarian literature to create a model for the optimal timing of having children.

A substantial body of literature exists on the modelling of women's fertility and labour decisions, but the modelling technique vary widely. The foremost

Petra Németh is assistant professor at the Department of Macroeconomics of Corvinus University of Budapest. E-mail: petra.nemeth@uni-corvinus.hu

The Hungarian manuscript was received on 30 June 2017.

DOI: <http://doi.org/10.25201/FER.16.4.535>

representatives of one of the main directions describe women's fertility and sequential labour supply decisions in a *dynamic model*, also known as the *life cycle model* under specific circumstances and conditions. The common feature of these models is that they assume optimising economic agents and factor in all direct and indirect costs and benefits associated with having children. Solving or structurally estimating these types of dynamic models gives us answers to several key questions such as family policy allowances and the role of childcare institutions in life cycle decisions, as well as explanations for decreasing fertility, changes in female labour supply and the differences in the fertility and labour supply decisions between countries (Arroyo – Zhang 1997, Hotz et al. 1997, Francesconi 2002, Del Boca – Sauer 2009, Bick 2010, Keane – Wolpin 2010). In this model framework, we can also investigate not only the optimal number of children, but also the optimal time interval between births.

Hotz, Klerman and Willis (1997) and Arroyo and Zhang (1997) provide a broad overview of the structure, characteristics and solvability of the dynamic or life cycle model investigating the fertility decision, while the work of Gábor (2005) provides a summary in Hungarian. Francesconi (2002) estimated a dynamic structural model in which married women make labour supply and fertility decisions. Women differ according to their full-time or part-time work status. The model yielded an important result in several regards, of which the most relevant one for our topic is: if the mother works full time and has a child, she will achieve a significantly lower lifetime utility if she stays off the labour market for an extended period after childbirth compared to staying off the labour market for a short period. However, this difference is negligible among women with part-time work. Keane and Wolpin (2010) also used a structurally estimated life cycle model to quantify the extent to which the difference in preference, the available welfare services and differences in labour market opportunities determined the life cycle decisions of Spanish women with various characteristics. Del Boca and Sauer (2009) estimated a decision rule derived from the life cycle model using data from Italy, France and Spain and drew conclusions from this on the link between the institutional environment, labour market flexibility, childcare institutions and activity and fertility decisions. Bick (2010) investigated two reforms recently introduced in Germany using a life cycle model calibrated to German data. His main line of investigation is the role of subsidised early childhood childcare in the life cycle labour supply of married women and their willingness to have children. Based on the results of the model, Bick (2010) concludes that the labour market activity and fertility of women with children under three years of age is positively influenced by the availability of subsidised crèches.

Mainly tied to Bick's (2010) paper, we build a microeconomics-based dynamic or *life cycle model* to determine the optimal childbearing strategy for Hungary that

factors in the fertility choice, resuming work after having children and the temporal links between the factors shaping the decision to have children and the mechanism of action between them. Accordingly, we would like to lay the foundations of a life cycle model that takes into account and factors in to such extent as is practicable the *domestic economic and institutional factors* and *environment* that influence/determine a family's fertility decisions and the subsequent return to work. These factors are: daytime childcare options for young children; the mother and father's education, labour market status and labour income; family allowances and tax rules. Among these, we focus particularly on investigating the role of family policy tools on fertility decisions. We would like to especially emphasise that – for modelling purposes – we ignored every other factor that we knew has or may have a significant impact on fertility choice, including changes in values, the spread of new types of relationships, cultural and biological factors or the family's housing conditions. In summary, we outline the life cycle decisions of women having completed their studies, but still of childbearing potential given the specific support and paid leave system, while the model's resolution is based on comparing the direct and indirect costs and benefits of having children.

The model ultimately shows that given a specific support environment, *when to have children and how many children are optimal with how much labour market absence* for families is characterised by different parameters (different education). In other words, the model gives us an answer as to how the many transformations of the family benefits scheme influence families' optimal childbearing and in what direction, all other things being equal. Based on the model, we also attempt to offer an explanation to in-depth questions such as which benefits affect the timing of childbearing and which benefits foster a quick return to the labour market; which benefits best contribute to childbearing by families with various degrees of education; how does the optimal childbearing strategy of families with different degrees of education differ under a specific benefits system. We look at three various family support regimes: the regime prevailing in 2006–2010, the one in 2011–2013 and the new legislative package on support and paid leave for mothers with small children adopted in 2014 (referred to as GYED-extra).

In the following, this paper is structured as follows: After the introduction, we outline the life cycle model and its resolution in Chapter 2 and provide and calibrate the model's parameters. In Chapter 3, based on the model's simulation results we present how a specific support and tax scheme influences the life cycle decisions on childbearing and female employment of families with different degrees of education, while Chapter 4 summarises the results of the model and presents the final conclusions.

2. The life cycle model of fertility choices under various family benefits schemes

In accordance with the literature, the model features the following general characteristics: It is first made up of a man and a woman who make a decision on their number of children and the timing of childbearing, on the utilisation of their time and incomes during their life cycle. The couple maximises its lifetime utility characterised by constant preference ordering¹ given specific time and budgetary constraints, and given the factors shaping the number of children and the raising of children and the technological constraints shaping the woman's productivity. The man and the woman's human capital increases with age/experience (and, as a result, their labour income), but at a decreasing rate, and the mother's human capital amortises during the time spent at home on maternity leave, in other words if she leaves the labour market (*Bartus et al. 2013*). In terms of the general structure, we supplemented or simplified the known modelling framework in several regards (adjusted to the Hungarian environment) and thus work based on the following specific life cycle model.

A typical family's joint life cycle decision on having children and working is considered realistic at different ages depending on its level of education, but the model investigates a 20-year period uniformly for every level of education. We also assume that the start of work is also the potential start of the option of having children irrespective of the degree of education. We determined the age for starting work based on the age published in the 2012 Hungarian Labour Market Yearbook and corresponding to the starting age for earning included in the 2011 age-income profiles: the examined career begins at age 18 for the unskilled and those with low qualification, at age 20 for those with medium qualification and at age 25 for those with high educational qualification (*Fazekas – Benczúr – Telegdy (edit.) 2012:372, Figure 6.3.4*). In addition, in line with the data, we can also assume that most women in Hungary currently have a child before the age of 45 and most women have their first child before the age of 30 on average (*KSH 2015:53, Table 4.1.3*).

We know, however, that having a child is a lifelong commitment and can be a significant cost/benefit for the family until the child turns 25. It would therefore be warranted for the couple to consider its decision to have children in light of their entire lifetime. However, we rejected this option for two reasons. For one, in line with the literature, we also use the simplifying assumption that borrowing or saving is not possible in either period (*Arroyo – Zhang 1997; Hotz et. al 1997; Francesconi 2002; Del Boca – Sauer 2009; Bick 2010; Keane – Wolpin 2010*). For this reason, we are unable to model either the flow of wealth between nations or long-term

¹ This assumption is important because we are currently not looking at the effect of changes in preference, that is, of changes in value on fertility.

care for the child in this framework. The other reason is that we do not think that someone who is planning to have a child in the near future is able to calculate or thoroughly consider the associated costs over the long term, i.e. over the next 20–25 years. Even if they could, it is improbable that this would substantially change their current decision.

The woman is the household's actual decision-maker, meaning it is she who decides during every period on consumption, childbearing and, if she has a young child, on labour, while the man has a passive role and works during every period (with the exception of families with education of less than eight years of elementary school) (Hotz *et al.* 1997; Keane – Wolpin 2010; Fehr – Ujhelyiova 2011). We assume that in periods when the family has no young children or does not yet have any young children, the woman works full time (except if the mother has education of less than eight years of elementary school, in which case she remains inactive throughout her entire lifetime). However, if there is a small child under the age of three in the family, the woman can decide about returning to work. Furthermore, each family can have three children at most in the model, only one child can be born per period² and the woman dedicates all of the time to child-raising in the years when a child is born. After they are born, the children consume similarly to their parents, and if the mother works, the children require childcare during early childhood (e.g. public crèche, family day-care, private crèche). We do not differentiate between the raising of children in the model,³ which may be a function of how much different parents spend on their child's education and training. For the sake of simplicity, we assume that free pre-school and school is available for every child at the same standard, and therefore we regard the cost of caring for children over the age of three as zero. We do not incorporate it into the model that the costs of raising children may change as a function of the child's age and the parents' education (Bartus *et al.* 2013).

In the model, we distinguish between families based on the parents' highest level of completed education. Investigating based on education is warranted for several reasons. For one, the productivity of an unskilled couple, a couple with low qualification, medium qualification or high qualification and thus their income profile differs over the course of their lifetime (Bartus *et al.* 2013). In addition, several family benefits in Hungary depend on income or are at least tied to employment therefore in conjunction with the different educational levels these benefits offer quite different economic incentives for having children. Accordingly, we view the model separately for decision-makers representing an unskilled group (up to eight grades of elementary school), a group with low qualification (vocational training, vocational school), medium qualification (high school) and high qualification (higher education),

² We assume that the probability of infant mortality is zero.

³ The trade-off between the number of children and the quality of children is addressed in the work of Becker (1993) and Bick (2010).

and symbolise the highest completed level of education with the starting gross salary and productivity parameters. For men, we assume an exogenous productivity profile depending on age and qualification level. By contrast, the productivity profile for women is endogenous, because we factor in that she does not work during the period when she is raising a child and her human capital accumulated up to that point amortises (*Bartus et al. 2013*). For the sake of simplicity, we assume that the individuals making up a household have the same level of education.⁴ We estimated the productivity trajectory parameters in the 2010 and 2011 environment based on 2011 real gross wages and in the 2014 environment based on 2013 real gross wages for men and women and for every educational level.

We solved the basic model for three different support environments for each educational level under the family support regime prevailing in 2010, in 2011 and the new regime introduced in 2014. The eligibility criteria for family benefits have changed substantially since 1 January 2014, and a new system that is more flexible in several points was adopted. In summary, all of the state tools that provide additional (disposable) income for families after they have children fall under the umbrella of family support; this concept therefore includes cash and family benefits⁵ and the family tax incentive system (according to the definition given by *Ignits – Kapitány 2006*).

To the extent allowed by an abstract model, we attempted to realistically incorporate the eligibility rules for benefits under the old and new regimes, the amount of benefits, the prevailing taxation rules and the income trajectories by educational level. However, we ignored other transfers as they are not closely tied to our main topic. In summary, we made the following assumptions in our model:

- We distinguish for educational levels: unskilled (up to eight grades of elementary school), low qualification (vocational training, vocational school), medium qualification (high school) and high qualification (higher education).
- The father works during every period and throughout his time if he has at least a low level of education. Unqualified fathers only spend a portion of their time working.
- The mother does not work until the child is one year old, can decide on working when the child is between one and three years old and works full time once the child is older than three. Unqualified women are an exception to this and spend all of their time at home irrespective of age or number of children.

⁴ There is empirical evidence that women with higher qualification typically choose a partner with higher qualification (*Becker 1981; Bartus et al. 2013*).

⁵ Cash benefits include the prenatal allowance and the childcare allowance, while family benefits include the maternity allowance, the family allowance, childcare benefits and the child-raising allowance.

- Daytime childcare for young children costs money.
- We consider the period between 18 and 37 years of age as the childbearing period for unskilled women and women with low education, between 20 and 39 years for women with medium education and between 25 and 44 years for women with high education.

In addition, we consider the following “fertility and labour life cycles” as the baseline and feel that these are an accurate reflection of reality for an average family with high, medium, low or no education:⁶

- A typical unskilled woman will have her first child at the age of 20, and we assume that she will remain inactive for the rest of her life.
- A typical woman with low education will have her first child at the age of 27, have two children consecutively and spend five years on average off the labour market.
- A typical woman with medium education will have one child at the age of 29 and spend three years off the labour market.
- A typical woman with high qualification (higher education) will have one child at the age of 31 and spend three years off the labour market.

In the following subchapters, we outline the costs and benefits of a family after they have children; we present the theoretical model that we built and its resolution, and finally we provide and calibrate the model’s parameters.

2.1. Costs and benefits of raising children

Having and raising children is associated with the following direct and indirect costs for families in the model:⁷

- the consumption of children (direct)
- costs for daytime childcare for young children if the mother works (direct)
- the mother’s lost current income during her years spent at home (indirect)
- the mother’s human capital loss reflected in lower lifetime income (indirect)
- the one-off fixed cost of the mother’s return to the labour market which increases with the number of children (representing the utility of household work) (indirect)

⁶ We must emphasise that when interpreting the results, these assumptions will serve as a baseline but will not influence the results.

⁷ András Gábos and Iván Gál Róbert factor similar components into the cost of raising children (*Gábos – Gál – Keller 2007*).

From another perspective, having children is beneficial for families for the following reasons:

- The child is a source of joy. The instantaneous utility function is therefore positively dependent on the number of children.
- Families are eligible for numerous transfers if they have children (family support and social security cash benefits).
- Parents are also eligible for the family tax incentive if they have children.

2.2. The household's decision-making problem

2.2.1. Lifetime utility

The maximisable lifetime utility function takes on the following form in the model:

$$V = \sum_{t=1}^T \beta^{t-1} u(c_{it}, n_{it}), \quad (1)$$

where $0 < \beta < 1$ is the impatience parameter, $u(.,.)$ is the instantaneous utility function⁸ which is the unit of consumption expressed in money (c_{it}) and the additively separable function of the current number of children (n_{it}):

$$u(c_{it}, n_{it}) = \frac{(1+c_{it})^{1-\sigma} - 1}{1-\sigma} + \Omega \frac{(1+n_{it})^{1-\varepsilon} - 1}{1-\varepsilon}, \quad (2)$$

where σ is the reciprocal of the intertemporal elasticity of substitution ($\sigma > 0, \sigma \neq 1$), Ω shows the weight of the number of children within decision-making preferences and ε measures how the utility function value reacts to changes in the number of children. In the lower index, i refers to the type of household based on educational level while T is the length of the examined life cycle, and t pertains to the period under review. We assume that the man and the woman have the same level of completed education which may be one of the following four: unskilled, low, medium or high education, respectively shown as $i=0,1,2,3$. Irrespective of educational level, the life cycle decision to have children is investigated over a 20-year period, although this applies to the period between 18 and 37 years of age for unskilled individuals and those with low qualification, between 20 and 39 years of age for individuals with medium education and between 25 and 44 years of age for individuals with high education.

2.2.2. Changes in the number of children over time

$m_{it} \in \{0,1\}$ is the number of children to be born during the period under review, which is equal to 0 if no children are born at the time t and 1 if a child is born. If

⁸ Bick (2010) and Jones (2010) used a form of function similar to ours. Bick and others included women's leisure time in the momentary utility function, but in our case, it is equal to zero.

$m_{it} = 1$, then the number of children, n_{it} increases by one. Only one child can be born per year, that is

$$n_{it} = n_{it-1} + m_{it}, \quad (3)$$

assuming that the children outlive their parents.

2.2.3. The budget constraint

We assume that neither saving nor borrowing is possible and so the family spends all of the income after tax from the man's work (indicated in the upper index as m) and the woman's work (indicated in the upper index as f) and the transfers granted based on their children on consumption and if necessary for the children's daytime care and the other costs of starting work. Accordingly, the budgetary constraints for a family of type i in the year t under the 2010 regime will be:

$$\bar{w}e_{it}^m L_{it}^m + \bar{w}e_{it}^f L_{it}^f - TAX_t + TR_t = c_{it} (1.7 + 0.5n_{it}) + pL_{it}^f (m_{it-1} + m_{it-2}) + k(1 - L_{it-1}^f) L_{it}^f \quad (4)$$

where the man and the woman's gross wage is the real wage per efficiency unit (\bar{w}) obtained as a multiple of productivity (e_{it}^m, e_{it}^f) and hours worked (L_{it}^m, L_{it}^f). The family pays tax during every period, the sum of which following deduction of the family tax allowance is TAX_t . All tax burdens are incurred differently under the various regimes due to changes in the family tax incentive and tax rules (see below). The benefits granted based on children further boost revenues, marked as TR_t during the period t (see below). The family must use its sources of income to cover all of its outlays, including its total consumption expressed as a multiple of a unit of consumption (c_{it}) and the OECD scale.⁹ Another expense is incurred if the mother returns to work when her child is between one and three years of age ($m_{it-1} = 1$ or $m_{it-2} = 1$), in which case daytime childcare must be arranged; the unit cost of this expressed in hours worked is p . The mother's time spent at home is not only useful in terms of childcare, but also for performing household chores. This is symbolised by the cost k , which is the one-off cost of the mother returning to the labour market. The value of the parameter k increases in parallel to the number of children.

Productivity profile. The man's productivity during his lifetime follows an exogenous trajectory depending on his age and education (Hotz et al. 1997; Attanasio et al. 2008; Keane – Wolpin 2010; Fehr – Ujhelyiova 2011). We describe developments in productivity as a function of the highest completed education (symbolised by the parameters α_i^m, γ_i^m), the amount of work experience (t) and the productivity level of the previous period (e_{it-1}^m), similarly to the 2008 paper of Attanasio et. al.:

$$e_{it}^m = e_{it-1}^m e^{\alpha_i^m + \gamma_i^m t}, \quad \text{where } \alpha_i^m > 0, \gamma_i^m < 0, \quad (5)$$

⁹ The OECD equivalence scale is the most prevalent in the literature, which assumes the consumption of the first adult among those sharing a household as one unit, the consumption of every additional adult as 0.7 units and the consumption of every child as 0.5 units (OECD 1982).

where the function's parameters differ by educational level (i). We also assume that the age of the man and the woman and their highest level of completed education are the same. The woman's productivity profile is defined endogenously because it depends on the labour supply decision also besides the qualification and the work experience (Attanasio *et al.* 2008). If the woman works during the period under review ($L_{it}^f = 1$), then her knowledge acquired until then will increase relative to the previous period at the same rate as the man's, while if she does not work ($L_{it}^f = 0$), then her productivity up to that point amortises at a rate of δ , that is:

$$e_{it}^f = e_{it-1}^f e^{((\alpha_i^f + \gamma_i^f)L_{it}^f - \delta(1-L_{it}^f))}, \text{ where } \alpha_i^f > 0, \gamma_i^f < 0. \quad (6)$$

Amount of taxes under the 2010, 2011 and 2014 regimes. In 2010, personal income tax (τ^1) was due based on the consolidated tax base (gross wage plus 27 per cent, also referred to as super gross) while social contributions (pension contribution (τ^2), health insurance contribution (τ^3), labour market contribution (τ^4)) were charged to the gross wage (NAV 2016). There was also an opportunity for tax credits which we factored into the model. We specify the accurate parameters in *Subchapter 2.4*. The family tax incentive (indicated as tax^1 per child) was granted as a tax rebate if a family had three or more children (Blaskó 2009). The total tax burden decreased by the family tax incentive can be calculated for 2010 as follows:

$$TAX_{it}^{2010} = (\tau^1 \cdot 1.27 + \tau^2 + \tau^3 + \tau^4)(\bar{w}e_{it}^m L_i^m + \bar{w}e_{it}^f L_i^f) - n_{it} \cdot tax^1. \quad (7)$$

By 2011 the tax regime had become a single-rate regime, but super grossing and tax crediting were still employed. Contributions were still due based on the gross wage (NAV 2016). Starting from 2011, the family tax incentive for families with children – the rate per child of which depends on the number of children and is represented by $tax^2(n_{it})$ – is an item which reduces the tax base and is deducted from the super gross wage (NEFMI 2011). Accordingly, the total tax burden decreased by the family tax incentive can be calculated for 2011 as follows:

$$TAX_{it}^{2011} = \tau^1 \cdot (1.27(\bar{w}e_{it}^m L_i^m + \bar{w}e_{it}^f L_i^f) - n_{it} \cdot tax^2(n_{it})) + (\tau^2 + \tau^3 + \tau^4)(\bar{w}e_{it}^m L_i^m + \bar{w}e_{it}^f L_i^f) \quad (8)$$

By 2014, super grossing was abolished, and thus personal income tax and contributions were also due based on the gross wage (NAV 2016). The family tax incentive remained as an item which reduced the tax base, but the tax incentive was expanded to encompass a family contribution incentive from 1 January 2014. Accordingly, if the family is unable to apply the full family incentive due to insufficient taxable income, it had the option to decrease its health insurance contribution, and later its pension contribution, by 16 per cent of the unapplied sum (or a portion thereof); we factored in the specific conditions of this regulation in solving the model (EMMI 2014). Total tax payable can thus be expressed based on TAX_{it}^{2014} as follows:

$$TAX_{it}^{2014} = \tau^1 \cdot (\bar{w}e_{it}^m L_i^m + \bar{w}e_{it}^f L_{it}^f - n_{it} \cdot tax^2(n_{it})) + (\tau^2 + \tau^3 + \tau^4) (\bar{w}e_{it}^m L_i^m + \bar{w}e_{it}^f L_{it}^f) - 0.16 \cdot (n_{it} \cdot tax^2(n_{it}) - \bar{w}e_{it}^m L_i^m + \bar{w}e_{it}^f L_{it}^f). \quad (9)$$

Size of transfers under the 2010 and 2011 regimes. If a child is born in a family ($m_{it}=1$), the mother first becomes entitled to a *maternity allowance* (TR^0) as a guaranteed right (MÁK 2014). Until the child turns one, the mother receives a *prenatal allowance*¹⁰ (tygás, represented by TR_t^1), which amounts to 70 per cent of the mother's real income earned in the year preceding the birth of the child and is only subject to the deduction of advance personal income tax (τ^1):

$$TR_t^1 = (1 - \tau^1) \cdot 0.7 \bar{w}_t e_{it-1}^f L_{it-1}^f \quad (10)$$

(OEP 2014). However, if the mother has another child and does not return to the labour market between the two births, the most she can received after the second child is a prenatal allowance of \bar{TR}^1 . Afterwards, the mother is eligible for the *childcare benefit* (gyed, represented by TR_t^2) until the child turns two, which also rises to 70 per cent of the mother's real income earned in the year preceding the birth of the child, which is also subject to a pension contribution payment obligation (τ^2) besides the advance personal income tax and its amount is maximised (\bar{TR}^2), that is:

$$TR_t^2 = \min \left\{ (1 - \tau^1 - \tau^2) \cdot 0.7 \bar{w}_t e_{it-2}^f L_{it-2}^f; \bar{TR}^2 \right\} \quad (11)$$

(OEP 2014). During parental leave, the mother is eligible for *childcare benefits* (gyes) as a guaranteed right, which is a fixed sum and is only subject to the deduction of a pension contribution (represented, after the deduction of contributions, by TR^3). If the mother returns to work after the child turns one, she loses eligibility for the childcare allowance, but will become eligible for childcare benefits alongside her earned income¹¹ (MÁK 2014). Families with three children receive childcare benefits during their entire life cycle, not only until the children turn three,¹² which we took into consideration in our programming. The family receives a family allowance for every child from the time of their birth ($TR^4(n_{it})$), which is a fixed sum and its amount depends on the number of children (MÁK 2014).

The amount of transfers under the 2010 and 2011 regimes ($TR_t^{2010/2011}$) are determined as:

¹⁰ In reality, the prenatal allowance is only granted in the first 24 weeks and from then on, the mother can apply for childcare allowance until the child turns one. For the sake of simplicity, we nevertheless assume in the model that the mother receives a prenatal allowance until the child turns one and childcare allowance when the child is between one and two years old.

¹¹ Here, we diverged from the actual policy, because under the 2011 regime, mothers could work up to 30 hours per week while receiving childcare benefits instead of full time.

¹² In reality, this benefit is known as the child-raising allowance (gyet), which is in fact the same amount as the childcare benefit.

$$TR_t^{2010/2011} = TR^0 m_{it} + TR_t^1 m_{it} + TR_t^2 m_{it-1} (1 - L_{it}^f) (1 - m_{it}) + TR^3 m_{it-2} (1 - m_{it-1}) (1 - m_{it}) + TR^4 (n_{it}) n_{it} \quad (12)$$

In solving the model, we also factored in that the childcare allowance is only granted to the mother if she was insured before giving birth ($L_{it-2}^f = 1$, or $TR_{t-2}^1 = 1$, or $TR_t^2 = 1$) and does not work when the child is between one and two years old. If the mother did not have valid insurance, she will still receive childcare benefits until her child turns two as follows:

$$TR_t^{2010/2011} = TR^0 m_{it} + TR^3 m_{it} + TR^3 m_{it-1} (1 - m_{it}) + TR^3 m_{it-2} (1 - m_{it-1}) (1 - m_{it}) + TR^4 (n_{it}) n_{it} \quad (13)$$

If there are several children under the age of three, the mother is only entitled to one benefit among the prenatal allowance, the childcare allowance and the childcare benefit under this regime.

Size of transfers under the 2014 regime. In the context of the extra childcare allowance (GYED-extra), the same benefits are available from 2014 and the amount of support (with the exception of the gyed maximum) and calculation method are the same. However, eligibility for these benefits changed significantly in two regards: one is that mothers can work full time while receiving the childcare allowance and the other is that the mother is eligible for several benefits if she had multiple young children. Accordingly, we calculate the amount of benefits under this family benefits scheme (TR_t^{2004}) differently than earlier:

$$TR_t^{2014} = TR^0 m_{it} + TR_t^1 m_{it} + TR_t^2 m_{it-1} + TR^3 m_{it-2} + TR^4 (n_{it}) n_{it}. \quad (14)$$

If the mother does not gain eligibility for the prenatal allowance and the childcare allowance, she can only receive childcare benefits during the first two years:

$$TR_t^{2014} = TR^0 m_{it} + TR^3 m_{it} + TR^3 m_{it-1} + TR^3 m_{it-2} + TR^4 (n_{it}) n_{it}. \quad (15)$$

2.2.4. Time constraint

One usual assumption in the literature is that the man works full time during every period, that is $L_i^m = 1$ (working time is normalised to one) (Hotz *et al.* 1997). We, however, only deem this assumption acceptable taking account of the stylised facts for a representative family of low, medium or high educational level ($i = 1, 2, 3$). By contrast, an unskilled breadwinner only spends 15 per cent of his time working during every period, that is, $L_0^m = 0.15$. Among women, we distinguish between unskilled and qualified women.¹³

¹³ See footnote 17 for an explanation of the choice of parameters.

In the model, we assume that the unskilled woman does not work irrespective of whether or not she has a child, that is, $L_{0t}^f = 0$ during every period. If, however, the woman has at least a low educational level ($i = 1, 2, 3$), we use several limiting assumptions in her case. The woman's labour supply L_{it}^f can take on two discrete values, zero and one. If the family does not yet have or no longer has any young children, the woman works full time, that is:

$$L_{it}^f = 1, \text{ if } m_{it}, m_{it-1}, m_{it-2} = 0 \text{ and if } i = 1, 2, 3 \quad (16)$$

We also assume that the woman does not work in the year of having the child, that is

$$L_{it}^f = 0, \text{ if } m_{it} = 1 \text{ and if } i = 1, 2, 3. \quad (17)$$

However, if the mother's child is between one and three years old, she must make a decision on working (or having more children), considering the associated costs and benefits. In this case, the woman has two options: either working full time or raising the child full time:

$$L_{it}^f = \{0, 1\}, \text{ if } m_{it-1} = 1 \text{ or } m_{it-2} = 1 \text{ and if } i = 1, 2, 3. \quad (18)$$

Although the change is radical between the two options,¹⁴ this assumption is not far removed from reality, as the part-time employment of women in Hungary is still low.

2.3. Resolving the household's decision-making problem

Resolving the household's dynamic problem can be expressed in the following manner. The household maximises its lifetime utility with the constraints presented in the previous chapter. The woman's decision-making problem to be represented using the following Belman equation:

$$\begin{aligned} V_{it}^{L^m} &= \max u(c_{it}, n_{it}) + \beta V_{it+1}^{L^m}, \quad \forall t < T_i \\ V_{it}^{L^m} &= \max u(c_{it}, n_{it}), \quad t = T_i \end{aligned} \quad (19)$$

where $V_{it}^{L^m}$ is the value function. We resolve the model recursively, progressing backwards from the last period according to the dynamic programming model with the help of a MATLAB software package. In the model, the woman must make decisions throughout the life cycle on consumption (c_{it}), childbearing (m_{it}) and labour (L_{it}^f) in possession of the parameters and exogenous status variables (L_i^m, e_{it}^m). We also assume that the couple's decision is characterised by perfect foresight and that the family benefits scheme and tax rules under review (2006–2010, 2011–2013 or 2014) will remain in place throughout their lifetime. It is based on this that we are seeking the "best $m_{it} - L_{it}^f$ series of combinations" throughout the life cycle that

¹⁴ While the mother receives childcare benefits, the employer has an obligation to take her back on the workforce with weekly working time of 30 hours at the mother's request.

maximises the lifetime utility of the household of type i under the specific family benefits scheme. When expressing a series of combinations, we factor in that one family can have up to three children. We then calculate the optimal fertility and labour strategy separately for 12 different scenarios – four educational levels and three family benefits schemes – that yields the *highest lifetime utility*.

2.4. Selecting parameter values

We set the model's parameters in terms of tax and social contribution rates and family benefits and the family tax incentive based on their actual value for all three periods and regimes. In the model, we regard one as HUF 100,000 and compare the other parameters expressed in money to this. *The Appendix summarises the parameter values.*

Under the 2010 regime, in the two-rate tax system, the personal income tax rate levied on labour income was 17 per cent for annual income of up to HUF 5 million ($\tau^1=0.17$) alongside a certain tax credit¹⁵ and for income of over HUF 5 million, 17 per cent for the portion of up to HUF 5 million and 32 per cent for the portion over HUF 5 million (in this case, $\tau^1=0.32$). The contributions were: 9.5 per cent pension contribution ($\tau^2=0.095$), 6 per cent health insurance contribution ($\tau^3=0.06$) and 1.5 per cent labour market contribution ($\tau^4=0.015$) (*NAV 2016*). The family tax incentive was HUF 4,000 per child ($tax^1=0.04$) in 2010 (*SZMM 2010*).

In 2011, the tax regime was a single-rate regime with a 16 per cent personal income tax rate (that is, $\tau^1=0.16$), but super grossing and tax crediting¹⁶ were still employed. Among contributions, only the pension contribution changed to 10 per cent ($\tau^2=0.1$), while the value of τ^3 , τ^4 remained unchanged (*NAV 2016*). From 2011, the family tax incentive granted if a family had children was HUF 62,500 if there were one or two children, and HUF 205,000 if there were three or more children ($tax^2(n_{it})=0.625/2.05$) (*NEFMI 2011*).

In 2011, the personal income tax rate levied on gross wages was also 16 per cent ($\tau^1=0.16$). Another change was the increase in the health insurance contribution to 7 per cent ($\tau^3=0.07$), while all other contributions remain at their 2011 level. The family tax incentive remained unchanged between 2011 and 2014, but the tax incentive was expanded to encompass a family contribution incentive from 1 January 2014 (*NAV 2016; EMMI 2014*).

¹⁵ The tax credit was 17 per cent of the wage (but capped at HUF 15,100) and can be fully used if annual income is HUF 3,188,000; the tax credit then decreases for higher incomes. Annual incomes in excess of HUF 4,698,000 are not eligible for the tax credit (*NAV 2016*). We took into account the relevant threshold value for the tax credit.

¹⁶ The monthly amount of the tax credit in 2011 was 12,100 at most, which can be used if annual income was 2,750,000, while the rate of the tax credit decreased for higher incomes. Annual incomes in excess of HUF 3,960,000 are not eligible for the tax credit (*NAV 2016*).

Because the man works all the time if his educational level is $i = 1, 2, 3$ and working time is normalised to one, $L_i^m = 1$ if $i = 1, 2, 3$. For the unskilled, this is $L_0^m = 0.15$.¹⁷ For the sake of simplicity and easier comparison, we assume that real wage applied to the efficiency unit $\bar{w} = 1$. In the model, we assume that daytime care for young children (in a public crèche, family day-care, private crèche or a babysitter) represents a cost of p per unit of time for parents irrespective of their educational level if the mother returns to the labour market. We compare this value to the monthly cost of day-care for an average family,¹⁸ so we estimated a cost of HUF 50,000 in 2011 and a cost of HUF 60,000 in 2014 ($p = 0.5/0.6$). There is a shortage of capacity in public crèches and the existing crèches are overcrowded. If we take a closer look at the statistical data, it can be seen that in recent years 33–35 per cent of children in crèches were over three years old,¹⁹ which means that the children of mothers who would like to return to the labour market while receiving the childcare allowance or childcare benefits are partially crowded out. In addition, the family also faces a cost of k that increases with the number of children if the mother resumes work: this cost reflects the utility of the household work performed during the mother's time spent at home. We do not differentiate this cost by educational level because we feel that it is not a function of education, but instead it differs individually based on how much and what level of household work is deemed useful for whom.²⁰

We derived the baseline values of productivity and the parameter values based on actual cross-sectional data for 2011 for the 2010 and 2011 versions of the model and based on data for 2013 for the 2014 model in the following manner: In the model, the temporal productivity profile coincides with the developments in the gross wage over time, thanks to the $\bar{w} = 1$ and $L_i^m = 1$ assumptions, which allows us to estimate productivity parameters based on actual real earnings data. For every gender and for all four levels of education, the baseline productivity level at the beginning of the life cycle $(e_{i,1}^f, e_{i,1}^m)$ is based on the average starting gross real wage in 2011 and in 2013 for the 2014 version.²¹ For men, productivity follows an exogenous

¹⁷ In the reference year 2011 the employment rate of males with education of less than eight years of elementary school was 9.7 per cent, while for those with education at most eight years of elementary school 28.7 per cent, for females the values were 4.7 and 17.8 per cent (*KSH 2016*). Based on this, we set the employment of males with education at most eight years of elementary school as 15 per cent, while for females it was 0.

¹⁸ We must dispel the myth that public crèche places can be accessed free of charge. Prior to 2011, payment could be requested for covering the raw material costs of meals and after 2011, a certain portion of overheads can be claimed from parents. Since 1 January 2012, the public operator can collect payment from parents for the daytime care of children. These fees apply over and above meal fees (*Makay – Blaskó 2012*).

¹⁹ Calculated based on *KSH 2014*.

²⁰ For instance, the importance or conversely, the substitutability of homemade meals for a family does not depend on educational level. For more on the calculation of the cost of household work, see *Gábor et al. 2007*.

²¹ *Fazekas – Benczúr – Telegdy (edit.) 2012:372, Figure 6.3.4.* and *Fazekas – Varga (edit.) 2015:238, Figure 6.3.4.*

trajectory. We determined the value of productivity parameters (α_i^m, γ_i^m) in such a manner that the exogenous trajectory of productivity most closely follows men's age-income profile by educational level in 2011 and in 2013 for the 2014 version.²²

Women's productivity trajectory, which plays a role for women with low, medium and high qualification, is determined endogenously, because if the woman has a child and remains off the labour market, then her knowledge acquired until then amortises (δ) and depreciates. We use the values of *Fehr and Ujhelyiova (2011)* for the amortisation rate (δ), i.e. we assume 1 per cent for women with low qualification and 2 per cent for women with medium and high qualification. We first took into account when and how many children a typical woman of a given level of education has and how long she remains off the labour market. The estimation procedure is then identical to the procedure used for men. For women, we determined the value of productivity parameters (α_i^m, γ_i^m) in a manner that the "typical woman's"²³ productivity trajectory most closely follows the age-income profile by educational level in 2011 and in 2013 given the baseline productivity values and the amortisation rate.²⁴

The sum of family benefits for one month provides the monthly amounts used in the model. Accordingly, the value of the maternity allowance is $TR^0=0.053$, as we distributed HUF 64,125 proportionately over 12 months. The amount of the other benefits – the maximum prenatal allowance (\overline{TR}^1) and childcare allowance (\overline{TR}^2), childcare benefits (TR^3), child-raising allowance (TR^3), family allowance (TR^4) – is defined by law; we used the relevant annual amounts for these items. For instance, if the net monthly child care benefit is HUF 25,600, we used $TR^3=0.265$ in the model. The values of the benefits provided under the family policy regimes under review are constant between 2008 and 2014.

The value of the impatience parameter is normally set between 0.9 and 1 for life cycle models (*Attanasio et al. 2008; Bick 2010; Fehr – Ujhelyiova 2011*). Similarly to the model of *Attanasio et. al. (2008)*, we used 0.98 as the value of β . For the parameters of the utility function, we also used the standard values found in the literature when possible. The reciprocal of the intertemporal elasticity of substitution of consumption, σ , ranges between 1.5 and 2, while ε , which measures how much the utility function reacts to changes in the number of children, is around 1.5. We used the values of *Bick (2010)*: therefore we fixed σ at 1.98 and ε at 1.39. Finally, in knowledge of the foregoing, we calibrated Ω , the weight of the number of children within instantaneous utility so that – in line with the stylised facts – highly qualified families optimally had one child at the age of 31 in the 2010 environment,

²² See: *Fazekas – Benczúr – Telegdy (edit.) 2012:372/Figure 6.3.4. and Fazekas – Varga (edit.) 2015:238, Figure 6.3.4.*

²³ See the list in *Chapter 2.*

²⁴ See: *Fazekas – Benczúr – Telegdy (edit.) 2012:372/Figure 6.3.4. and Fazekas – Varga (edit.) 2015:238, Figure 6.3.4.*

while those with low qualification optimally had their child as young as possible. We uniformly use these values for the parameters associated with utility for all three household types/educational levels and all support environments. The *Appendix* provides a tabulated summary of the specific parameter values.

The majority of the taxation/labour market/productivity/family support parameters are estimated based on actual data, and therefore we did not subject them to robustness testing. An exception to this is the parameter representing daytime care for small children (p), in which a ± 20 per cent divergence did not change our conclusions; in other words, our results can be regarded as robust for these parameter values. This is mainly due to the fact that for families of low and medium educational level, this cost is not incurred,²⁵ while for highly qualified families, this expenditure item is so small compared to income that a ± 20 per cent divergence in the cost does not impact their decision. We also performed a sensitivity test for ± 5 , 10 and 20 per cent divergences in the parameter Ω (the weight of the number of children in instantaneous utility). We address our results and interpret the tables in the following chapter.

3. Simulation results

Solving the model gives us an answer to the *optimal number of children* that a couple with a specific level of education should have during their lifetime and *when they should optimally have them* under the current benefits system, and when the mother should return to the labour market in order to maximise her lifetime utility. It should be noted once again that in the model we make the optimal strategy of having children conditional exclusively on the *family support and taxation environment and the level of education*, that is, we ignore all other parameters (such as changes in values or standards, the spread of new forms of relationships, cultural and biological factors and housing status) which, in reality, also shape the number of children and the timing of childbearing. Because other influencing factors are ignored, the model is only partially able to reflect the observable differences in the fertility patterns of families with differing characteristics. In summary, *the results show the optimal number of children that a given family type should have and when they should optimally have them when considering the financial aspects only.*

The decision on the right strategy is made by comparing the various temporal costs and benefits of having children. Childbirth in a family is, on the one hand, a value in and of itself, i.e. it represents a positive benefit throughout the family's lifetime. On the other hand, the parents become entitled to various benefits and allowances for many years (income effect), and therefore family benefits positively impact consumption and life cycle utility by increasing lifetime income. If the amount of benefits increases in parallel with the number of children, or if they become

²⁵ For families of low and medium educational level, we assume that the mother spends three years at home with her children and therefore, there is no need for daytime childcare.

available only subject to a certain number of children, then the income effect is stronger (e.g. the system of tax allowances). But if the amount of benefit per child does not increase or increases only marginally in parallel with the total number of children, then the correlation is weaker (e.g. family allowance) (Gál 2011).

From another perspective, having children represents substantial expenses for the family. These expenses grow as the number of children increases. One of the parents – in reality most frequently, and also in the model, the mother – temporarily leaves the labour market after the child is born, meaning that she does not receive any income and the mother’s expertise acquired until then is eroded during parental leave which will, in turn, setback her entire future life cycle earnings (substitution effect) (Jones *et al.* 2010; Gál 2011; Bartus *et al.* 2013), negatively impacting the family’s life cycle consumption and thereby also its life cycle utility. As the number of children increases, the value of lost income from work and the mother’s human capital amortises to a greater extent (Bartus *et al.* 2013). Moreover, with age and experience, the wage of those with a higher level of education increases much more dynamically than that of their peers with a lower level of education, and for this reason, they are likely to delay their intention to have children (postponement) in order to minimise the impact of the alternative cost of having children on their life cycle income²⁶ (Bartus *et al.* 2013). Therefore, the substitution effect is greater overall in families with a higher level of education. However, this is compensated by the fact that the earning party in the family will have a higher lifetime income throughout his lifetime than a father with a lower educational level. The fact that the children also consume, just as the parents, and during early childhood they need day-care (direct cost) if the mother returns to the labour market, also represent additional costs. However, the model fails to consider that fact that the cost of child-rearing increases in parallel with the age of the children.

Overall, the optimal number of children and the optimal time of having them is obtained so that the family reaches the maximum lifetime utility resulting from the above-mentioned costs and benefits arising during the life cycle. It is of crucial importance as to what extent the amount of benefits and tax allowances received for the children can compensate the family’s lost lifetime income and the direct costs associated with child-rearing. The costs and the benefits associated with having children differ for decision-makers with different educational levels and in the different family support and taxation environments, and for this reason, the optimal outcome will be different for each of them. In our analysis, we compare the values of the calculated life cycle utility in the different scenarios. We always take as a starting point the “typical fertility- employment life cycles”²⁷ which we deem to best reflect reality in the case of an average family with high, medium and low educational level. We always compare the values of lifetime utility with the initial

²⁶ Aspects of economies of scale.

²⁷ We presented the typical fertility-employment life cycles in *Chapter 2*.

case in such a way that, *ceteris paribus*, we modify some of our initial conditions. In the simulation, *the final number of children, the mother's optimal age, the timing of having children and the length of parental leave* are examined. First, we collectively examine the fertility strategy of individuals with low, medium and high educational level; we then separately analyse the change in the willingness to have children of unskilled couples.

When resolving the model, we took into consideration the current support and taxation environment for every regime, while we estimated the gross earnings based on the gross income paths based on education in 2011 and 2013 in the case of the support system in force in 2010/2011 and in 2014, respectively. However, in its current form, the model is unable to reflect the changes in real values (not adjusted by inflation), and therefore we should not draw conclusions in the case of the different regimes from the comparison of the absolute values of obtained life cycle utility. The model can, however, provide answers as to which family support environment incentivises more children, earlier employment or, in the case of more children, a faster timing to have children. In the following subchapter, we present the simulation results and the conclusions we can draw from them in the different cases.

3.1. Optimal number of children

Firstly, the main focus of the analysis is to find out how *many actual children* is optimal for the different family types during their lifetime in the different support and taxation environment in 2010, 2011 and 2014. For the purposes of the analysis, we considered the mother's time spent with her child/children as constant. The preconditions and the results of the simulation are presented in *Table 1*.

Table 1												
Value of lifetime utility of the various types of families in the examined three benefits systems, in the cases of zero, one, two and three children												
Educational level	High				Medium				Low			
	zero	one	two	three	zero	one	two	three	zero	one	two	three
Number of children	zero	one	two	three	zero	one	two	three	zero	one	two	three
Age	–	31	31, 33	31,33,35	–	29	29, 31	29,31,33	–	27	27,30	27,30,32
Leave (years)	–	two	two	two	–	three	three	three	–	three	three	three
2010	11.841	11.921	11.74	11.634	9.662	9.573	9.484	9.499	8.767	8.722	8.222	8.62
2011	12.021	12.154	12.02	12.148	9.589	9.591	9.568	9.703	8.675	8.747	8.68	8.769
2014	12.337	12.462	12.321	12.474	9.678	9.668	9.659	10.051	8.714	8.755	8.684	8.947

Note: The educational level variable denotes the mother's and the father's highest completed educational qualification, while the number of children designates the final number of children in the family, the age indicates the mother's age at the time her child/children were born, and the years of leave variable denotes the number of years the mother spends at home with her children on maternity leave (gyes/gyed).

Interpretation of the results can be facilitated if we first think over the underlying effect mechanisms. As the parameters of the utility function evolve independently from both the educational level and the family benefits system, having one child for any family type in any environment represents the same additional benefit. However, the financial costs of having children depends on the level of education and the support and taxation environment. Therefore, it is of key importance what percentage of the costs can be compensated by benefits and tax allowances after the birth of a child during the life cycle, and as such, what will be the consumption per unit in a family, as this is the other factor that influences utility. Compared to not having any children, upon having the first, the second and often the third child consumption per unit will be increasingly lower during the life cycle which means that the benefits and allowances only partially cover the costs. However, the joy associated with having children intensifies as the number of children increases. Lifecycle utility measures the total value of the current utility taken at present value, which stems from whether the damage caused by a decline in consumption due to having children or the joy takes precedence.

Accordingly, we can draw the following conclusions from the results of *Table 1*. The utility of families with one child is higher in every case than that of families with two children, but it is not always higher than that of families without children. The tax allowance system in force between 2006 and 2010 available only for families with three or more children was so beneficial for families with low and medium educational level that it also encouraged families with only two children to have a third child. However, this was not the case for families with higher education having only two children. This was attributable to the fact that by having the third child, the life cycle consumption decreased only marginally or might have even increased compared to having only two children, although its rate still fell short of the life standard of families without children or only with one child. Overall, however, the joy associated with having children was able to compensate the negative effect of the lower consumption cycle for families with a lower level of education. Only in the case of families with higher education can we observe that families with one child were better off than those without children; in their case, this was attributable to the high value of the prenatal allowance and childcare allowance (tygás and gyed) they received.

Thanks to the tax allowances which were extended in 2011 to families with one or two children, the situation improved in the sense that having the first child put these families in a better position than families without a child. Moreover, the tax allowance, which was substantially augmented for three children, reduced the cost of child-rearing for families of every educational level to such an extent that they could overall achieve higher lifetime utility with three children than with only two children, because by having the third child, their life cycle consumption

did not decline, but on the contrary actually increased compared to having only two children. Although its value still lagged behind the living standard of families without children or only with one child, the joy associated with having children was able to compensate the reduced utility stemming from a lower consumption cycle. This is why we obtained the results according to which families with low and medium educational level with three children were even *better off* than having only one child. Our results also proved to be robust when the Ω parameter was changed as well.

The child benefit extra (GYED-extra) introduced in 2014 brought about additional improvement, as the benefits available at the expense of contributions further reduced the cost of having children, and so the system encouraged couples of every educational level to have three children. However, not even the measures of GYED-extra could improve the fact that “it is not worthwhile” to have a second child after the first one, as the position of families with two children, no matter what level of education, was worse than that of families with one child or without children. This is related to the fact that the birth of the second child represents a bigger burden for families (reduction of utility during the life cycle), even taking account of the supports and allowances, than the amount of joy they gain (increase in utility during the life cycle). This may mainly restrain the willingness to have children of those earning lower wages.

The described results can be regarded as robust within the ± 20 per cent interval of the value of parameter Ω in the case of families with low educational level, and only within the ± 10 per cent interval in the case of families with medium educational level. That is, in their case, a greater increase in the weight of the number of children tilts the scale in favour of families with two children compared to single-child families, but only in 2011 and 2014. We must handle the results for 2014 with caution for the group with higher education as they react sensitively to the change in Ω : if this is reduced by only 5 per cent, we can observe that three children do not provide a higher lifetime utility than a single child; there is only an improvement in comparison with childlessness and the two-children family model.

3.2. Increasing age of mothers

In the following, we analyse for families with one and two children whether the change in the support and taxation environment influences the optimal age of the mother in the case of couples with different educational levels. For the purposes of the analysis, we considered the mother’s time spent with her child/children as constant, in the case of both one child and two children.

Table 2
Value of lifetime utility of the various types of one-child families in the examined three benefits systems considered with advancing maternal age

Educational level	High			Medium			Low		
	one	one	one	one	one	one	one	one	one
Number of children	one	one	one	one	one	one	one	one	one
Age	27	31	35	25	29	34	24	27	30
Leave (years)	two	two	two	three	three	three	three	three	three
2010	11.899	11.921	11.907	9.547	9.573	9.571	8.736	8.722	8.703
2011	12.148	12.154	12.128	9.609	9.591	9.543	8.799	8.747	8.713
2014	12.455	12.462	12.436	9.685	9.668	9.628	8.774	8.755	8.726

Note: See Table 1.

According to Tables 2 and 3, under the family policy regime in 2011 and 2014, those who became parents at a younger age could achieve a higher lifetime utility (or the improvement was greater) than those having children at a later age, as in 2006, but only in the case of couples with low and medium educational levels. This stems from the fact that starting from 2011 in the case of those with lower education/wages, the family tax allowance also extended to families with one and two children, and the support available from 2014 in the form of contribution allowances was able compensate for the lost income at a younger age still. The described results are also considered robust when the value of parameter Ω changes by ± 10 per cent. However, the value differences are not significant among the groups with either educational level.

Table 3
Value of lifetime utility of the various types of two-child families in the examined three benefits systems considered with advancing maternal age

Educational level	High		Medium		Low	
	two	two	two	two	two	two
Number of children	two	two	two	two	two	two
Age	27,29	31,33	25,27	29,31	23,26	27,30
Leave (years)	two	two	three	three	three	three
2010	11.659	11.74	9.409	9.484	8.211	8.222
2011	11.971	12.02	9.574	9.568	8.754	8.68
2014	12.27	12.321	9.671	9.659	8.696	8.684

Note: See Table 1.

The legislative changes in 2011 and 2014 did not trigger a turning point for families with higher education. They can obtain higher lifetime utility in all three examined support environments if they have children above the age of 30 and not below.

A 5 per cent decrease in the parameter Ω shifts the optimal timing for having children in families with low and medium educational levels to an older age, but in their case, the results are also not sensitive to a 20 per cent increase in the

parameter Ω . By contrast, those who are highly qualified would have been better off having children at an increasingly younger age in 2011 and 2014 if parameter Ω increased by 5 per cent, but if the parameter decreases by 20 per cent, then the presented correlations do not change.

3.3. Scheduling the time of having children

In this sub-section, we compare the different scheduling of having children of families with three children. For the sake of the analysis, we fixed the mother's age in the case of the first child and only changed it for the second and third child, and we also considered the mother's time spent with her children at home as constant in every case.

Table 4									
Value of lifetime utility of the various types of three children families in the examined three benefits systems in the case of different timing									
Educational level	High			Medium			Low		
Number of children	three	three	three	three	three	three	three	three	three
Age Leave (years)	31,33,35 two	31,33,37 two	31,35,39 two	29,31,33 three	29,31,35 three	29,33,37 three	27,30,32 three	27,30,34 three	27,32,36 three
2010	11.63	11.613	11.624	9.50	9.35	9.28	8.62	8.46	8.48
2011	12.15	12.08	12.05	9.70	9.54	9.40	8.77	8.62	8.60
2014	12.47	12.41	12.39	10.05	9.73	9.52	9.95	9.48	9.46

Note: See Table 1.

Based on economic rationality, when having several children, families are better off financially if the children are born with the least age difference possible since this enables the mother to potentially stay away from work for a shorter period of time (less loss of human capital and also loss of life cycle income), and the family is able to use the tax allowance available for the three children sooner, or at least part of it.

The family policy regime between 2006 and 2010 did not significantly influence the timing of having children for highly qualified couples, but in certain cases a greater age difference may have been more beneficial: it is true that the tax allowance for families with three children became available only from a later period, but this was compensated by the fact that the mother could re-enter the labour market between children, and thanks to this, she received more pregnancy and maternity aid (tgyás) after the second and the third child (see first line of Table 4). But in the case of families with lower education, a smaller age difference between children clearly increases lifetime utility.

Highly educated couples who therefore earned high wages were incentivised for *fast scheduling* mainly by the tax allowance regime introduced in 2011, as they were already able to access a large part or the entire amount of the tax allowance under the 2011 regime. But the value of life cycle utility does not significantly change

under either regime when modifying the timing of having children. By contrast, the measures of GYED-extra clearly incentivised couples with lower earnings (those with low and medium educational level) for *faster timing of children bearing*. This is explained by the fact that the family tax allowance, also extended as a contribution allowance, and the rule that the family could be entitled in parallel to family benefits for several young children could compensate the costs associated with having children to a greater extent. The described results are robust for the ± 20 per cent change of the value of the parameter Ω for couples of every educational level. In this case, the optimal outcomes were not influenced by the pillar of the measures introduced in 2014 based on which the mother would be entitled to receive child benefit (gyed/gyes) even if she returned early to the labour market, because we assume in their case that the mother stays home for three years after the birth of every child. This enables us to separately analyse the components of GYED-extra.

3.4. Increasing parental leave

In the following, we compare the fertility strategies prevailing in the case when the mother’s parental leave is increasingly long (one or three years). We present the analysis only for highly qualified couples with one, two or three children. The results obtained for couples with low and medium educational level are presented only in the essay part. In the simulation, we left the mother’s age constant throughout the period examined.

Table 5				
Value of lifetime utility of highly qualified families in the three examined benefits systems, assuming increasing parental leave				
Educational level		High		
Support environment		2010	2011	2014
Number of children Age Leave (years)	one 31 one	12.03	12.266	12.603
	one 31 three	11.816	12.052	12.355
Number of children Age Leave (years)	two 31,34 one	11.912	12.188	12.541
	two 31,34 three	11.576	11.864	12.145
Number of children Age Leave (years)	three 31,34,37 one	11.834	12.28	12.678
	three 31,34,37 three	11.439	11.94	12.239

Note: See Table 1.

For highly educated families, we obtained the rationally expected result that the less time the mother spent at home with her child/children, the better (in financial terms) for the family under all support schemes, provided that the mother had a job. In other words, higher lifetime utility can be achieved if the mother only spends one year at home with each of her children after they are born instead of two or three years.

Based on the table's results, it is also apparent that there is a stronger incentive for the mother's early return to the labour market in 2014 compared to 2010 and 2011 thanks to the fact that the mother does not lose the childcare allowance if she resumes work after her child turns one. The improvement between 2011 and 2014 is significant for both one, two and three children if measured by the percentage change in lifetime utility. In other words the childcare allowance extra (GYED-extra) measures foster a faster labour market return of mothers, which claim holds true for families with low and medium qualification. The described conclusions are robust for changes in parameter Ω , as here, we assumed that the number of children and the constraints on childbearing are fixed.

3.5. Childbearing among unskilled families

According to the stylised data and thus also in the model, unskilled families behave differently, so we investigate them separately. We symbolise the unskilled couple using the following conditions: the mother remains an active throughout her lifetime irrespective of the number of children and the father also only works a part of the time. The following table summarises the value of lifetime utility if the family has one, two or three children paced differently.

Table 6									
The value of lifetime utility of unskilled families under the examined three benefits systems, in case of one, two and three children									
Educational level	Uneducated								
Number of children	one	one	one	two	two	two	three	three	three
Age	18	23	28	18,20	18,22	18,24	18,20,22	18,20,24	18,20,26
2010	5.065	4.558	4.248	5.446	5.325	5.248	6.302	6.232	6.155
2011	5.156	4.525	4.064	5.717	5.705	5.633	7.958	7.907	7.889
2014	5.067	4.374	3.886	7.812	7.671	7.588	8.059	8.014	7.996

Note: See Table 1.

Changes in family benefits schemes did not affect the optimal number of children in their case, nor did it significantly impact the timing of childbearing. By and large, the optimal strategy for them is to have as many children as possible (three children is better than one or two), as this entitles the mother to more benefits. In addition,

it is best to have the children as young as possible and spaced close together, as this allows the family to take advantage of the benefits derived from the tax and contribution allowances applied to the father's income earned during his part-time work and to extend the eligibility period of the family allowance.

The results also show that the rise in the value of lifetime utility is more significant in 2011 compared to 2010, thanks to the tax incentive by having a second child after the first and a third child after the second. The measure, extended as a contribution incentive in 2014, yields a *greater improvement* than earlier by having a second child after the first, measured at a change in the value of lifetime utility. By contrast, families are unable to benefit from the high tax and contribution incentive granted for the third child to a greater extent, at which point “only” the childcare benefits and the family allowance provide an additional incentive.

4. Results and conclusions

In light of the results from the *life cycle model*, we are able to determine *when and how many children* different types of families, i.e. families of different educational levels should have, and *when the mother should return to the labour market* to achieve the maximum lifetime utility given the specific family support and tax regime. In other words, the model gives us an answer to how the multiple transformations of the family benefits scheme influences families' optimal childbearing and in what direction, with all other things being equal. Due to the differences in employment, we analysed the behaviour of families with low, medium and high qualification separately from the behaviour of qualified families. As a result, the first four points pertain to the first three family types while the last point addresses the fertility strategy of unskilled families.

1. The tax allowance system in force between 2006 and 2010 available only for families with three or more children was so beneficial for families with low and medium educational level that it also encouraged families with only two children to have a third child. However, this was not the case for families with higher education with only two children. From 2011, having the first child put these families in a better position than being childless. Furthermore, due to the significantly increased tax incentive for families granted for three children, all couples of every educational level could achieve higher lifetime utility than with two children. However, from 2014, we observed that having three children is clearly the most beneficial for families compared to having one child/being childless. However, a major impeding factor is that *it is not “worthwhile”* to have a second child after the first one in either support and taxation environment.
2. Among the family support tools introduced, the family tax incentive scheme introduced in 2011 has the greatest impact on optimal maternal age, as it makes

it worthwhile to have children at a *younger* age for those with low and medium education.

3. Those with high qualification and thus high income *were most incentivised to have children at a young age by the tax incentive* adopted in 2011 while among lower-income families (medium and low qualification), the greatest incentive to having children early) were *certain elements of the childcare allowance extra*.
4. There was *growing incentive for the mother's early return* to the labour market after having a child from 2010 to 2011, and then to 2014 for all three educational levels.
5. For *unskilled families*, the optimal strategy under all three support regimes is to *have as many children as possible spaced closely together and at the earliest possible age*. The measures of the childcare allowance extra shaped the optimal outcome in that having a second child after the first one resulted in a *greater improvement in lifetime utility than earlier*.

Although we laid the foundations of the model, we do not feel that in its current state it is able to fully factor in and assess the impact of all economic and institutional factors shaping fertility choices. But building and stimulating the model brings us closer to understanding which support elements foster and contribute to childbearing *from a financial perspective*.

References

- Arroyo, C. – Zhang, J. (1997): *Dynamic Microeconomic Models of Fertility Choice: A Survey*. Journal of Population Economics, 10: 23–65. <https://doi.org/10.1007/s001480050030>
- Attanasio, O. – Low, H. – Sanchez-Marcos, V. (2008): *Explaining Changes in Female Labor Supply in a Life-Cycle Model*. American Economic Review, 98(4): 1517–1552. <https://doi.org/10.1257/aer.98.4.1517>
- Bartus, T. – Murinkó, L. – Szalma, I. – Szél, B. (2013): *The Effect of Education on Second Births in Hungary: A Test of the Time-Squeeze, Self-Selection, and Partner-Effect Hypotheses*. Demographic Research, 28(1): 1–32.
- Becker, G.S. (1981): *A Treatise on Family*. Cambridge, MA: Harvard University Press.
- Bick, A. (2010): *The Quantitative Role of Child Care for Female Labor Force Participation and Fertility*. MPRA Paper No. 25474.
- Blaskó, Zs. (2009): *Családtámogatás, gyermeknevelés, munkavállalás (Family support, child-rearing, employment)*. In: Monostori, J. – Őri, P. – Molnár, E. S. – Spéder, Zs. (editor): *Demográfiai Portré 2009. Jelentés a magyar népesség helyzetéről (Report on the situation*

- of the Hungarian population), Hungarian Demographic Research Institute of the Hungarian Central Statistical Office, Budapest. 41–53.
- Del Boca, D. – Sauer, R. M. (2009): *Life Cycle Employment and Fertility across Institutional Environments*. *European Economic Review* 53: 274–292. <https://doi.org/10.1016/j.euroecorev.2008.06.001>
- EMMI (2014): *Family benefit, family contribution benefit 2014*. Ministry of Human Capacities. <http://emmiugyfelszolgalat.gov.hu/kereses?searchStr=csal%C3%A1di+ad%C3%B3kedvezm%C3%A9ny>. Downloaded: 10 September 2014
- Fazekas, K. – Benczúr, P. – Telegdy, Á. (edit.) (2012): *Munkaerőpiaci Tükör 2012 (Hungarian Labour Market, 2012)*. HAS Institute of Economic Centre for Economic and Regional Studies, Országos Foglalkoztatási Közhasznú Nonprofit Kft. Budapest. www.bpdata.eu/mpt/2012hua06_04. Downloaded: 12 October 2014.
- Fazekas, K. – Varga, J. (edit.) (2015): *Munkaerőpiaci Tükör 2014 (Hungarian Labour Market 2014)*. HAS Institute of Economic Centre for Economic and Regional Studies. Budapest. http://econ.core.hu/file/download/mt_2014_hun/egyben.pdf. Downloaded: 21 January 2016.
- Fehr, H. – Ujhelyiova, D. (2011): *Fertility, Female Labor Supply, and Family Policy*. CESINFO Working Paper No. 3455.
- Francesconi, M. (2002): *A Joint Dynamic Model of Fertility and Work of Married Women*. *Journal of Labor Economics*, 20(2): 336–380.
- Gábos, A. (2005): *A magyar termékenységi rendszer termékenységi hatásai (The impacts on fertility of the Hungarian fertility system)*. PhD dissertation, Corvinus University of Budapest, Sociology PhD programme. Budapest, 2005.
- Gábos, A. – Gál, R.I. – Keller, T. (2007): *A gyermeknevelés költsége és a társadalmi kompenzáció (The cost of child-rearing and social compensation)*. In: Mészáros, J. (editor): *Családi jövedelemadózás*. Budapest, Barankovics Alapítvány.
- Gál, R.I. (2011): *Fenntartható életpálya-finanszírozás*. NFFT Műhelytanulmányok No 4.
- Hotz, V.J. – Klerman, J.A. – Willis, R.J. (1997): *The Economics of Fertility in Developed Countries: A Survey*. Chapter 7. In: *Handbook of Population and Family Economics*. Volume A.: 275–347, Elsevier.
- Ignits, Gy. – Kapitány, B. (2006): *A családtámogatások alakulása: célok és eszközök (Evolution of family allowances: objectives and means)*. *Demográfia*, 49(4): 383–401.

- Jones, L.E. – Schoonbrodt, A. – Tertilt, M. (2010): *Fertility Theories: Can they Explain the Negative Fertility-Income Relationship?* In: Shoven, J.B. (ed.): *Demography and the Economy*. University of Chicago Press: 43–100.
- Keane, M.P. – Wolpin, K.I. (2010): *The Role of Labor and Marriage Markets. Preference Heterogeneity and the Welfare system in the Life Cycle Decisions of Black, Hispanic and White Women*. *International Economic Review*, 51(3): 851–892. <https://doi.org/10.1111/j.1468-2354.2010.00604.x>
- KSH (2014): Hungarian Central Statistical Office: Tables (STADAT). Annual time-series data. Social safety-net. Table 2.4.10 http://www.ksh.hu/docs/hun/xstadat/xstadat_eves/i_fsg005a.html. Downloaded: 10 September 2014.
- KSH (2015): *Annals of Demography 2014*. Hungarian Central Statistical Office. Budapest.
- KSH (2016): Hungarian Central Statistical Office: Tables (STADAT). Annual time-series data. Labour market. Table 2.1.44 http://www.ksh.hu/docs/hun/xstadat/xstadat_eves/i_qlf045.html. Downloaded: 5 January 2016.
- Makay, Zs. – Blaskó, Zs. (2012): *Családtámogatás, gyermeknevelés, munkavállalás (Family support, child-rearing, employment)*. In: Óri, P. – Spéder, Zs. (edit.) (2012): *Demográfiai Portré 2012*. Hungarian Demographic Research Institute of the Hungarian Central Statistical Office. Budapest: 45–56.
- MÁK (2014): *Családtámogatási ellátások (Family support allowances)*. Hungarian Treasury. <http://www.allamkincstar.gov.hu/maganszemelyek/csaladtamogatas>. Downloaded: 2 September 2014.
- NAV (2016): National Tax and Customs Administration: *Tax rates, contribution rates: Tax tables 2010-2016; Payable contributions 2010-2016*. https://www.nav.gov.hu/nav/szolgaltatasok/adokulcsok_jarulekmertekek. Downloaded: 13 January 2016.
- NEFMI (2011): *Családi adókedvezmény (Family tax allowances)*. Ministry of National Resources, Social and Information Department, Budapest.
- OECD (1982): *The OECD List of Social Indicators*. OECD. PARIS. 1982.
- OEP (2014): *Az Egészségbiztosítási Alapból támogatott ellátások (Benefits funded from the National Health Insurance Fund)*. National Health Insurance Fund. http://www.oep.hu/portal/page?_pageid=34.66136161&_dad=portal&_schema=PORTAL. Downloaded: 30 September 2014.
- SZMM (2010): *Tájékoztató a gyermekek után járó ellátásokról (Information on the benefits for dependent children)*. Ministry of Social Affairs and Labour. Household Information and Advisory Department. Budapest, 2010.

Appendix: Summary table of parameter values

Description of parameter	Parameter value
Taxation parameters	
Total taxes and contributions on gross wage-related income	$\tau=0.345$
Personal income tax rate	$\tau^1=0.17/0.32/0.16$
Pension contribution	$\tau^2=0.095/0.1$
Health insurance contribution	$\tau^3=0.06/0.07$
Labour market contribution	$\tau^4=0.015$
Rate of family tax allowance per child in 2010 (only in case of three children)	$tax^1=0.04$
Rate of family tax allowance per child in case of one, two or three children after 2011	$tax^1(n_{it})=0.625/2.05$
Labour market parameters	
Time spent working by an unskilled man	$L_0^m = 0.15$
Time spent working by a man ($i=1,2,3$)	$L_i^m = 1$
Time spent working by an unskilled woman	$L_0^f = 0$
Time spent working by a woman ($i=1,2,3$)	$L_i^f = 0/1$
Real wage for efficiency unit	$\bar{w} = 1$
Cost of day-care of young children	$p = 0.5/0.6$
Fixed cost of re-entering the labour market in case of one, two and three children	$k = 0.5/0.7/0.8$
Productivity parameters	
Productivity of a highly qualified man during the first period (2011/2013)	$e_{3,1}^m = 2.172 / 2.498$
Productivity of a medium skilled man during the first period (2011/2013)	$e_{2,1}^m = 1.207 / 1.374$
Productivity of a low skilled man during the first period (2011/2013)	$e_{1,1}^m = 1.009 / 1.004$
Productivity of an unskilled man during the first period (2011/2013)	$e_{0,1}^m = 1.013 / 9.45$
Productivity of a highly qualified woman during the first period (2011/2013)	$e_{3,1}^f = 1.869 / 1.933$
Productivity of a medium skilled woman during the first period (2011/2013)	$e_{2,1}^f = 1.078 / 1.205$
Productivity of a low skilled woman during the first period (2011/2013)	$e_{1,1}^f = 0.927 / 0.901$
Productivity parameters of a highly qualified man (2011/2013)	$\alpha_3^m = 0.1226 / 0.097$ $\gamma_3^m = -0.00693 / -0.0052$

Description of parameter	Parameter value
Productivity parameters of a medium skilled man (2011/2013)	$\alpha_2^m = 0.0636 / 0.058$ $\gamma_2^m = -0.00301 / -0.003$
Productivity parameters of a low skilled man (2011/2013)	$\alpha_1^m = 0.052 / 0.0766$ $\gamma_1^m = -0.0027 / -0.00454$
Productivity parameters of an unskilled man (2011/2013)	$\alpha_0^m = 0.01855 / 0.037$ $\gamma_0^m = -0.00097 / -0.00185$
Productivity parameters of a highly qualified woman (2011/2013)	$\alpha_3^f = 0.1 / 0.104$ $\gamma_3^f = -0.0062 / -0.0062$
Productivity parameters of a medium skilled woman (2011/2013)	$\alpha_2^f = 0.072 / 0.0705$ $\gamma_2^f = -0.00375 / -0.00375$
Productivity parameters of a low skilled woman (2011/2013)	$\alpha_1^f = 0.048 / 0.084$ $\gamma_1^f = -0.0021 / -0.0046$
Amortization rate of low and medium skilled and highly qualified persons	$\delta = 0.01/0.02/0.02$
Family support allowances	
Maternity support	$TR^0 = 0.053$
Maximum net value of the prenatal allowance (tygás)	$\overline{TR}^1 = 1.3$
Maximum net value of the childcare allowance (gyed)	$\overline{TR}^2 = 0.808$
Childcare benefit/child-raising allowance amount (gyes/gyet)	$TR^3 = 0.256$
Amount of family allowance per child in case of one, two and three children	$TR^4 = 0.122/0.133/0.16$
Parameters related to utility	
Impatience parameter	$\beta = 0.98$
The weight of the number of children within current utility	$\Omega = 0.123$
Reciprocal of the intertemporal elasticity of substitution of consumption	$\sigma = 1.98$
To what extent does the utility function react to the change in the number of children	$\epsilon = 1.39$

Possible Paths for GDP Per Capita – Simulation with a Demographic Growth Model

Éva Berde – Izabella Kuncz

To begin our paper, we point out the importance of demographic growth models by highlighting the conceptual framework of overlapping generations. We define the formulas in our own model based on Lee – Mason (2010), modifying the original framework in several respects. We present the exogenous fertility and survival rates in different demographic scenarios, and then we derive the simulation paths for GDP per capita from these. A word of caution regarding our results: the simple structure of our model disregards several factors potentially influencing growth. When concentrating solely on the impact of changes in fertility and mortality rates, our simulation results suggest that a drop in the total fertility rate, even to slightly below the replacement level, and even in the context of a relatively high survival rate, may increase GDP per capita. However, in the case of an extremely low total fertility rate and a high survival rate, an economic downturn can be expected.

Journal of Economic Literature (JEL) codes: E23, J11, O40

Keywords: overlapping generations model, total fertility rate, survival rate, simulation path, GDP per capita

1. Introduction

The exploration of the regularities in economic growth and the projection of potential future paths has excited economists since the emergence of economic thinking.¹ Nevertheless, the significance of population numbers and structures

*Éva Berde is a professor at the Corvinus University of Budapest. E-mail: eva.berde@uni-corvinus.hu
Izabella Kuncz is an assistant lecturer at the Corvinus University of Budapest.
E-mail: izabella.kuncz@uni-corvinus.hu*

The study was financed by the Pallas Athena Domus Animae Foundation. The views expressed in this article are solely those of the authors.

The Hungarian manuscript was received on 26 June 2017.

DOI: <http://doi.org/10.25201/FER.16.4.3657>

¹ The first such efforts that can be cited appeared with the representatives of the classical school of economics. *Smith (1776)* not only presented the benefits of the division of labour, but also mentioned that the productivity growth that can be achieved through the division of labour increases economic prosperity, and facilitates the future development of the economy.

only started to be taken into account in growth models much later.² From the second half of the 1980s, as humanity entered the fourth stage of demographic transition, demographics determined economic developments to such an extent that demographic variables could no longer be left out of growth models.

In the fourth stage of demographic transition,³ the mortality rate steadily declines in almost all developed countries of the world, and in parallel with this, the total fertility rate (TFR) is even lower than earlier. The total fertility rate has dropped to below the replacement level several times, which portends serious problems for the future. Fulfilling the needs of the ageing population, and, more generally, maintaining the operation of the economy may prove to be highly difficult under such circumstances. While the total fertility rate of more developed regions drops, huge net population growth characterises certain less developed areas, due to a diminishing mortality rate combined with a high fertility rate, which, however, is still lower than earlier. In these countries, the future decrease in the total fertility rate and the further drop in the mortality rate are expected, and therefore it is likely that the share of the working-age population will soar temporarily. These trends may provide an economic stimulus to the countries concerned, and the beneficial effect of the so-called first demographic dividend⁴ can take hold (see, for example, *Bloom et al. 2003, Mason 2005*). Ultimately, even the ranking in the economic clout of the countries may change.

Out of the growth models, the changing demographic trends can be best captured by the so-called overlapping generations (OLG) model. In the study by *Diamond (1965)*, which laid the foundations of OLG models, changes in the birth rate were still exogenous, however, it already differentiated between two age groups of the population. This framework was expanded by *Auerbach – Kotlikoff (1987)* to include further cohorts, and then the modelling tools were enhanced by *Barro – Becker (1989)*. The article by Barro–Becker also treats the birth rate endogenously: therefore, the total fertility rate is quantified as a result of the processes in the model, and it has an impact on the development of the model's variables.

Our paper draws conclusions with respect to the potential paths of economic growth using one of the overlapping generations models. The basic model we used, *Lee – Mason (2010)*, uses a special approach as compared to the usual OLG models. Still, it is based on Diamond's idea, i.e. the overlapping generations, and

² *Malthus (1798)* concentrated specifically on the negative impact of excessive population growth, but did not link this issue to the potential growth in productivity. Demographic factors were first taken into account in growth models by *Solow (1956)*. The historical sequence of the demographic approaches to growth models is summarised in *Table A1* in the Appendix.

³ For more on the demographic transition, see, for example, *Van de Kaa (2010), Frejka (2016)*.

⁴ The first demographic dividend shows the difference between the growth rate of the weighted number of workers and the growth rate of a similarly weighted number of consumers. If the number of the so-called effective workers increases more (or falls less) than the number of effective consumers, that contributes to economic growth.

the neoclassical growth model framework, in which economic growth is simulated as a function of the total fertility rate and the mortality rate. One simplification as compared to Diamond's model is that this does not contain consumer optimisation, and one added, novel feature is that the expansion of human capital is among the explanatory variables of the model. Its core concept is the trade-off between quantity and quality as formulated by Becker (*Becker 1960, Becker – Lewis 1973, Willis 1973, Barro – Becker 1989, Becker et al. 1990, Galor – Weil 1999*), according to which adults spend more on their offspring when they have less children, i.e. they provide them with more human capital. As a result, when children grow up, their productivity increases, and therefore the impact of the lower total fertility rate can be offset. Our paper shows that the total fertility rate can drop to a critical level where the opportunity for growth is threatened, and society is unable to maintain the earlier level of GDP per capita. The study is structured as follows: after the introduction, in the second part, we describe Becker's quantity–quality trade-off, and the corresponding theories. In Part 3, we will present our own model, and show in which aspects the framework developed by us differs, and to what extent, from *Lee – Mason (2010)*. In Part 4, we use artificially constructed exogenous fertility and mortality rates, and simulate paths for GDP per capita based on the model of the previous part. We detail the demographic trends that the exogenous total fertility and mortality rates express, and why it is important to be aware of the impact these trends may have on economic growth. Part 5 comprises the summary of our study. Then in the Appendix we provide an overview of the historical development of growth models using demographics in a tabular form, and we also employ a table to show the values of the parameters in our model.

2. Beckerian quantity–quality trade-off

The basic idea behind the Beckerian quantity–quality trade-off is attributed to *Becker (1960)* in the literature. In short, the quantity–quality trade-off means that if there are many offspring in a family, less time and money are devoted to one, while in the case of fewer offspring, the expenses per child increase drastically. Similar approaches to the quantity–quality trade-off are discussed in *Becker – Lewis (1973), Willis (1973)* and *Galor – Weil (1999)*. This issue, especially the global contraction of the total fertility rate, is very topical now. For example, *Lee – Mason (2010)* use statistical analysis to prove that in the countries in the NTA⁵ database, children with fewer siblings receive more human capital investments from their parents; therefore, they can work more efficiently as adults than their peers growing up in larger families. The link between the reduction in the total fertility rate and the increase in human capital investments may explain why production values increase

⁵ NTA: National Transfer Accounts. The data can be found at <http://www.ntaccounts.org/> (last accessed on 20 September 2017).

even if the share of the working-age population diminishes. *Roudi-Fahimi – Kent (2007)* provide an outstanding summary of the studies in connection with this idea.

Even the so-called synthesis theory, which treats the causal relationships in a complex manner, and which does not support the concept of the Beckerian quantity–quality trade-off (e.g. *Adelman 1963, Freedman 1963, Silver 1965, Freedman – Coombs 1966a, Freedman – Coombs 1966b and Easterlin 1973*) concedes that the decrease in the total fertility rate and the growing human capital investments occurred at roughly the same time in history. Further schools of thought other than the synthesis theory that distance themselves from Becker’s ideas explain the reduction in the fertility rate with several factors. These include the increase in female employment, economic crises, people’s love of comfort and the opinion of contemporaries (see, for example, *Kaplan 1994, Black et al. 2005, Ellis 2008, Luci – Thevenon 2010, Sobotka et al. 2011, Colleran et al. 2015, Dang – Rogers 2016*). And despite the apparent facts, the hotly debated *Lawson – Borgerhoff Mulder (2016)* study flatly rejects that there is any causality between human capital investments and the development of the total fertility rate. Nonetheless, these two authors admit that fertility has declined and the human capital invested in children has increased since the mid-19th century. *Guo – Zhang (2017)*, partly in response to the paper by *Lawson – Borgerhoff Mulder (2016)*, show that the reason why certain authors doubt the theory of the quantity–quality trade-off is a misguided interpretation of the facts.

One exception to Becker’s ideas could be that in the countries with an advanced public sector, all children have access to basic schooling and healthcare, regardless of the number of siblings. Yet *Vargha – Donehower (2016)* estimated the value of the so-called invisible transfers,⁶ i.e. the care and attention directly devoted to children by their parents. The two authors have confirmed that the children growing up among fewer siblings receive much more care in the countries with advanced public benefits than their peers living in large families, mainly because of the invisible transfers.

In our article, we base our model on the negative relationship between the number of children and human capital investments, i.e. we accept the causality derived from Becker’s ideas. We examine what would happen if this correlation determined the productivity of an economy, *ceteris paribus*. The impact of material capital and all the other factors is quantified only indirectly, through human capital, but it is indicated that sometimes material capital can be the driver of development. The approach examining the role of human capital *ceteris paribus* is increasingly common in the literature, in view of the drying-up of the first demographic dividend

⁶ For more on the invisible transfers themselves, see *Gál et al. (2016)*.

and the increasingly dominant role of the second demographic dividend⁷ (e.g. Bloom et al. 2003, Mason 2005, Mason et al. 2016).

3. Our model

3.1. Differences between Lee – Mason (2010) and our model

Just as in the case of Lee – Mason (2010), our overlapping generations model is highly stylised. Our basic aim is to determine the impact of the population's age distribution (and indirectly population numbers) on output per capita. Despite its stylised nature, it is important that the model be as realistic as possible.

From this perspective, taking into account four overlapping generations instead of the three, as in Lee – Mason (2010), provides enormous help in achieving better alignment with the facts. With three overlapping generations, those living until the end of the third period spend merely one-third of their life working, and live from the transfers provided by the others for two-thirds of their life. In the case of three generations, there is an unreasonably high number of children in any period in the context of the actual TFR. If we assume that $N_t^{working}$ is the number of workers and F_t is the fertility rate (for one person rather than one woman, which is usually the case in such models), then $F_t N_t^{working}$ children will be born altogether, and the share of the first generation will be disproportionately large, even when F_t is barely over 1 (and the TFR is over 2). In the case of four generations, the people living until the end of the fourth period spend half of their lives working, and if the total number of workers in the two working generations is $N_t^{working}$, while the total fertility rate is F_t , then, assuming that the two working generations are roughly of the same size,⁸ only $F_t 0.5 N_t^{working}$ children will be born. The share of children within society is much more realistic this way, just like the fact that people spend at least half, rather than a third, of their lives on the labour market. Those who die after the third period, work for two-thirds of their lives.

Accordingly, our model assumes two, rather than one, working generations living at the same time: they are the second and third generations in the model. These two generations resemble each other, but only the second generation can produce children. In addition, the income of the third generation is somewhat higher than that of the second. Our model does not differentiate between whether the transfers or the in-kind goods are provided to children (Generation 1) and the old (or pensioners, i.e. Generation 4 in our model) within the family or

⁷ For a short description of the first demographic dividend, see Footnote 4. The second demographic dividend refers to the economic growth fuelled by the stronger human capital investments and the related physical capital investments.

⁸ The roughly same size of the two working generations is merely a stylised and momentary simplification, with the sole aim of highlighting the magnitude.

through redistribution. The overall benefits cover the consumption of children and pensioners, and guarantee the human capital investments in children. These benefits together are referred to as transfers.

Somewhat simplifying the complex correlations in reality, we assume in our model that both working generations divide their income based on the same principles between children's transfer, their own consumption and the elderly's consumption. Therefore, both working generations take part in caring for children,⁹ and the elderly have no savings at all, i.e. their livelihood depends completely on the transfers provided by the second and third generations.

The generations that do not work, i.e. children and the elderly, may receive transfers from the workers, with variable lower and upper limits, depending on certain factors in our model. Thanks to the limits, those who do not work receive at least minimal benefits, even in hard times. Furthermore, the limits also ensure that the workers do not have to spend on those who do not work beyond their means, even if the number of those who do not work is relatively high. The rules in the model that determine the way the income of the two working generations is spent also ensure that the simulation results are determined not only by the proportion of the size of the individual generations in the population, but also indirectly by the total population. Meanwhile, the amount spent on human capital investments from the transfers influences the productivity of the young workers in the next period, and of the older workers in the era coming after that.

Although in their model Lee and Mason quantify the fertility elasticity of human capital, as the embodiment of Beckerian quantity–quality trade-off, using actual statistical data and an econometric equation, the value of elasticity is a constant negative number. We did not regard the value of elasticity to be constant, but to be dependent on the total fertility rate. The details of the calculations are presented in *Chapter 3.3*.

3.2. Equations in the model

As described in the previous part, the economy in the model contains four overlapping generations:¹⁰ children (N_t^1), young workers (N_t^2), older workers (N_t^3) and pensioners (N_t^4). On average, the members of the young working generation produce F_t children in Period t . In the next period, children become economically active as young workers, and young workers become older workers. Finally, the latter's s_t share will reach retired age, and the others die at the end of the third

⁹ According to the model, both working generations finance children's consumption and human capital, but they do not necessarily provide the transfers directly. The distribution in the model is similar to a system in which children receive the amount through central redistribution.

¹⁰ The population number of the generation where $i = 1, 2, 3, 4$ in Period t is denoted with N_t^i .

period. The demographic transitions in the model are described in the following equations:

$$N_t^1 = F_t \cdot N_t^2 \quad (1a)$$

$$N_t^2 = N_{t-1}^1 \quad (1b)$$

$$N_t^3 = N_{t-1}^2 \quad (1c)$$

$$N_t^4 = s_t \cdot N_{t-1}^3 \quad (1d)$$

Total population size in Period t : $N_t = N_t^1 + N_t^2 + N_t^3 + N_t^4$.

As indicated earlier, in our model F_t denotes the fertility rate for one person rather than one woman in Period t . Therefore, assuming an equal number of men and women in a stylised approach, our simulation used half of the TFR value actually possible. In our model, no one dies until the end of the third period, and those surviving the third period die only at the end of the fourth period, but then they do so for sure. Therefore, in our model, s_t is half of the statistically measurable survival rate in the third period. The development of F_t and s_t is also key in the model. The two exogenous ratios together determine the size and structure of the population in each period.

Only two working generations perform income-generating activities, and they receive wages for their work. The wage of young workers (W_t^2) depends on the amount of human capital they possess (H_t), which was accumulated in the previous period:

$$W_t^2 = g(H_t), \quad (2a)$$

where $g'(H_t) > 0$ and $g''(H_t) < 0$. The wages of older workers are proportionately higher than that of younger ones, which is indicated by the parameter φ , the value of which used in our calculations can be seen in *Table A2* of the Appendix.

$$W_t^3 = f(W_t^2) = \varphi \cdot W_t^2, \quad (2b)$$

where $\varphi > 1$.

Human capital is invested in children by the two working generations, spending h_t of their income for this purpose. Equation (3) shows the amount of human capital of the young workers entering working age in Period t . In line with the assumptions in *Lee – Mason (2010)*, Equation (3) presumes that everyone receives the human capital investments in childhood, and that this will determine their productivity when they become workers:

$$H_t = h(F_{t-1}) \cdot (W_{t-1}^2 + W_{t-1}^3). \quad (3)$$

Since only the two working generations have income, they finance their own consumption as well as that of the two dependant generations, and the human capital investments in children. Accordingly, the budget constraint is as follows:

$$W_t^2 \cdot N_t^2 + W_t^3 \cdot N_t^3 \geq C_t^1 \cdot N_t^1 + C_t^2 \cdot N_t^2 + C_t^3 \cdot N_t^3 + C_t^4 \cdot N_t^4 + H_{t+1} \cdot N_t^1. \quad (4)$$

In line with the above, the amount of human capital investments financed from income depends on the total fertility rate, based on the following correlation:

$$H_t = \alpha \cdot F_{t-1}^{\beta_{t-1}} \cdot (W_{t-1}^2 + W_{t-1}^3), \quad (5)$$

where $\alpha > 0$ is the human capital investment ratio with one unit of F (replacement level). β_{t-1} is the elasticity of human capital investments as a function of F_{t-1} , which will be later referred to as fertility elasticity. This was fixed by *Lee – Mason (2010)* at an average value, but we made it dependent on the total fertility rate. The estimation of β_{t-1} is presented in *Chapter 3.3*.

The wage of the two working generations is

$$W_t^2 = \gamma \cdot H_t^\delta = \gamma \cdot (\alpha \cdot F_{t-1}^{\beta_{t-1}} \cdot (W_{t-1}^2 + W_{t-1}^3))^\delta, \quad (6a)$$

where $0 < \delta < 1$, $\gamma > 0$ and substituting Equation (6a) in Equation (2b):

$$W_t^3 = \varphi \cdot \gamma \cdot (\alpha \cdot F_{t-1}^{\beta_{t-1}} \cdot (W_{t-1}^2 + W_{t-1}^3))^\delta. \quad (6b)$$

We would like to point out parameter γ in Equation (6b), the value of which was fixed at 1 in our model calculations, similar to *Lee – Mason (2010)* (see *Table A2* in the Appendix). If γ is greater than 1, this represents a sort of growth factor in the model. This growth factor can express the expansion in the efficiency of either human capital or physical capital. Since we currently aim to monitor the potential changes in fertility and mortality rates *ceteris paribus*, we utilised only one direct efficiency-increasing option in the model, the impact of human capital investments on the efficiency in production. Nonetheless, we allow for the option of the parameter γ being greater than 1 when our simulation is rerun. In this manner, the growth in the efficiency of physical capital can be modelled discretely.

Correlation (4) regarding the budget shows that the different generations can spend more money only at each other's expense. Therefore, we had to incorporate limits in our model that prevent the potentially large first or fourth generations from burning through the income generated by the workers, leaving nothing to those who produced it. In addition, a similar limit had to be applied to prevent the opposite, i.e. we had to ensure that the two working generations provide at least minimal transfers to the first and fourth generations, even if there are many people in the

latter groups in a given period. Equations (7a) and (8a) below place caps and floors on the consumption of the elderly and children, respectively, using Ψ_t defined in Equation (7b), μ_t determined in Equation (8b) and v_t established in Equation (8c).¹¹ The consumption of the fourth generation is:

$$C_t^4 = \Psi_t \cdot \left(W_t^2 \cdot \frac{N_t^2}{N_t^4} + W_t^3 \cdot \frac{N_t^3}{N_t^4} - H_{t+1} \cdot \frac{N_t^1}{N_t^4} \right), \quad (7a)$$

where

$$\Psi_t = \min \left(0.25; \frac{N_t^4}{N_t}; 1.1 \cdot \frac{N_{t-1}^4}{N_{t-1}} \right). \quad (7b)$$

The consumption of the first generation is:

$$C_t^1 = \mu_t \left(W_t^2 \cdot \frac{N_t^2}{N_t^1} - \alpha \cdot F_t^{\beta_1} \cdot W_t^2 \right) + v_t \left(W_t^3 \cdot \frac{N_t^3}{N_t^1} - \alpha \cdot F_t^{\beta_1} \cdot W_t^3 \right), \quad (8a)$$

where

$$\mu_t = \min \left(0.25; \frac{N_t^1}{N_t}; 1.1 \cdot \frac{N_{t-1}^1}{N_{t-1}} \right) \quad (8b)$$

$$v_t = \min \left(0.25; \frac{N_t^1}{N_t}; 1.1 \cdot \frac{N_{t-1}^1}{N_{t-1}} \right). \quad (8c)$$

The income left over after the payment of the transfers to the first and fourth generations is spent by younger and older workers on their own consumption:

$$C_t^2 = (1 - \mu_t - \Psi_t) \left(W_t^2 - \alpha \cdot F_t^{\beta_1} \cdot W_t^2 \cdot \frac{N_t^1}{N_t^2} \right) \quad (9a)$$

$$C_t^3 = (1 - v_t - \Psi_t) \left(W_t^3 - \alpha \cdot F_t^{\beta_1} \cdot W_t^3 \cdot \frac{N_t^1}{N_t^3} \right). \quad (9b)$$

The formula of GDP per capita, which in our model means income per person, should be expressed separately. This can be seen in Equation (10). This is because our simulation paths determine the development of GDP per capita.

$$\frac{GDP_t}{N_t} = \frac{W_t^2 \cdot N_t^2 + W_t^3 \cdot N_t^3}{N_t}. \quad (10)$$

¹¹ The above-mentioned payment limits were incorporated in the model using the ratios Ψ_t , μ_t and v_t . If the four generations are of the same size, and the available income (remaining after the payment of human capital) is distributed evenly among them, all generations receive 25 per cent of the income. Otherwise, in the case of generations of a relatively smaller size, the proportion of spending on their consumption changes in line with their share within the population, but it may rise by 10 per cent at most in a period to avoid extreme growth.

Using the previous equations of the model, wage dynamics, i.e. the ratio of future and current wages, can be described as follows:

$$\frac{W_{t+1}^2}{W_t^2} = \frac{\gamma \cdot (\alpha \cdot F_t^{\beta_t} \cdot (W_t^2 + W_t^3))^\delta}{W_t^2} = \gamma \cdot (\alpha \cdot F_t^{\beta_t})^\delta \cdot \frac{(W_t^2 + \varphi \cdot W_t^2)^\delta}{W_t^2} = \gamma \cdot (\alpha \cdot F_t^{\beta_t})^\delta \cdot (1 + \varphi)^\delta \cdot W_t^{2\delta-1} \quad (11a)$$

$$\frac{W_{t+1}^3}{W_t^3} = \frac{\varphi \cdot W_{t+1}^2}{\varphi \cdot W_t^2} \quad (11b)$$

In a steady state, wages are constant, and the wage of young workers is

$$W_t^{2*} = \left(\frac{1}{\gamma \cdot \alpha^\delta} \right)^{\frac{1}{\delta-1}} \cdot F_t^{\frac{\beta_t \delta}{1-\delta}} \cdot \left(\frac{1}{(1+\varphi)^\delta} \right)^{\frac{1}{\delta-1}}, \quad (12a)$$

while, when substituting Equation (12a) in (2b), the wage of older workers is

$$W_t^{3*} = \varphi \cdot \left(\frac{1}{\gamma \cdot \alpha^\delta} \right)^{\frac{1}{\delta-1}} \cdot F_t^{\frac{\beta_t \delta}{1-\delta}} \cdot \left(\frac{1}{(1+\varphi)^\delta} \right)^{\frac{1}{\delta-1}}. \quad (12b)$$

3.3. Estimating the fertility elasticity of human capital investments

In *Chapter 3.1*, we already indicated that *Lee – Mason (2010)* determined the value of the elasticity of human capital investments as a function of the total fertility rate to be negative but constant. The two authors used *NTA (2009)* data and found a significant negative correlation between the total fertility rate and human capital investments. However, they relied on this relationship as a generally occurring link, quantifying the elasticity of human capital investments as a function of the total fertility rate based on the average value.

By contrast, we employed elasticity that depends on the value of the total fertility rate. We used the average number of years spent in school (ISCED 1 or higher level studies) as a variable accurately representing the amount of human capital investments from the *UNESCO (2016)* database. The value of the total fertility rate was compiled from *UN (2015)*, and in the end we had data for 98 countries.¹² The countries were divided into two groups, one for the countries with a TFR of over 2.1 (the replacement level), and the other for the countries below that. Using the ordinary least squares (OLS) method, we prepared regression estimates for the two country groups separately. In the case of high-fertility countries, the absolute value of the regression coefficient was significantly higher: $\beta_t = -0.8348$, while it

¹² We were able to use data from far more countries than *Lee – Mason (2010)*, who relied solely on *NTA (2009)*.

was $\beta_t = -0.273$ in the low-fertility country group. After that, assuming a linear relationship¹³ between β_t and the F_t fertility rate, β_t was determined as follows:

$$\beta_t = -0.4072 - 0.761 \cdot \ln F_t. \quad (13)$$

The values determined by Equation (13) fit well on the points represented by the total fertility rate and human capital investments of the 98 countries, and that is why this correlation was chosen in our model for estimating the elasticity of human capital investments as a function of the total fertility rate. If $F_t < 0.588$, β_t would be positive, therefore in this range, Equation (13) conflicts with the theory based on which we quantified it. However, there was no such low total fertility rate in the statistical data, and we did not use such low F_t values in our simulation. Further research is necessary to confirm or reject the relevance of the Beckerian quantity–quality trade-off with respect to TFR values below 1.176. Yet during our model calculations, we never used such a low total fertility rate.

4. The relationship between the total fertility rate, the survival rate and growth based on our model's correlations

The simple structure of our model facilitates simulation testing that examines the impact of changes in only the total fertility rate and the survival rate on economic growth, independent of other factors. These simulation calculations show the impact of the two aforementioned demographic indicators on the development of GDP per capita, all other factors being fixed. In other words, only the categories determining the size of the population¹⁴ change in our model, while everything else is constant.

Assuming simple correlations has both benefits and drawbacks. The most important benefits in this case are:

- We can see the effects exerted by the simultaneous change in the total fertility rate and the mortality rate *ceteris paribus*.
- We can analyse how the different paths of the total fertility rate and the mortality rate as compared to each other change per capita output.
- We can highlight the cases when the expected demographic developments lead to a dangerous economic situation.

¹³ It was assumed that the points determined by the average total fertility rate and human capital investments (or, more precisely, its logarithm), can be connected with a straight line. The mathematical statement of this line can be seen in Equation (13).

¹⁴ There is no migration in our model.

- If demographic developments influence economic growth favourably, even temporarily, the projection helps the economy in utilising the positive trends as long as possible.

The drawbacks of the simple approach are the following:

- Several factors can influence GDP per capita that cannot be derived from the development of the total fertility rate and the survival rate.
- Our forecasts can stoke unreasonable fears if the negative impact of demographic factors can be offset.
- Positive predictions can falsely suggest that nothing has to be done for economic growth, as the demographic drivers take care of that “on their own”.

Now we will present the GDP per capita values on the various simulated paths. During our simulations, both the total fertility rate (F) and the survival rate (s) were provided exogenously.¹⁵

The exogenous time series for F and s were determined in a way so that there would be cases when a high fertility rate nosedives, when the moderately high fertility rate stay roughly the same, and when the low fertility rate continues to decline. In parallel with this, the survival rate stagnates or grows in all cases. Therefore, we sought to find out how the various possible paths of the total fertility rate influence GDP per capita *ceteris paribus*, while survival rates do not drop. Of course, our model can be used to conduct simulations using several other fertility and mortality time series, although we mainly focused on the various, hypothetical and opposite paths of the total fertility rate, while assuming that the mortality rate values do not deteriorate at all.

We started our simulation from a steady state, from which the economy shifts due to the change in the total fertility rate and the survival rate. The values of the parameters used during the simulation can be seen in *Table A2* of the Appendix. Similar to *Lee – Mason (2010)*, the initial wage in the steady state is the wage value in the context of the given parameters and exogenous variables. After the total fertility rate and the survival rate start changing in each period, the economy shifts from this steady state, and later, when F and s do not change anymore, it converges towards a new steady state. Similar to *Lee – Mason (2010)*, our study focused on the path towards the new steady state, i.e. the time of transition.

¹⁵ As a reminder: F is the fertility rate per person, and half of the expected TFR, while s is the survival rate, and half of the survival rate for 60 years of age supported statistically, since in the model those living to be 60 die at the age of 80.

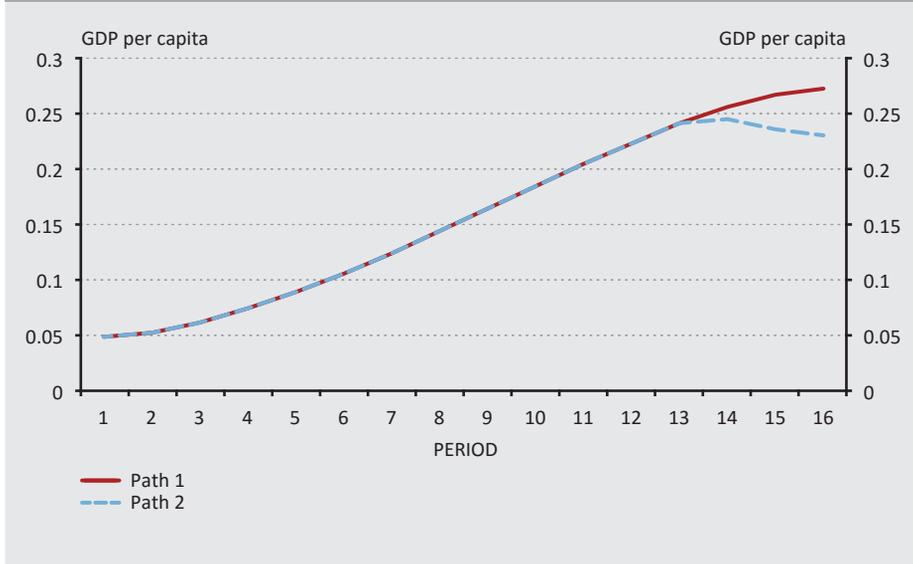
Several other authors have used a modelling technique in which the stationary condition of the economy was upset, and it was examined how the path leading to the new steady state developed. For example, *Cipriani (2014)* used a steady state of an OLG model to demonstrate the expected impact of rising life expectancy in the context of first an exogenous, then an endogenous total fertility rate. He wanted to find out how much pension will be paid to the members of the older generation. *Becker et al. (1990)* also examined the steady state and its stability using endogenous fertility and human capital. The model by *Barro – Becker (1989)* administered hypothetical shocks to the steady state, and monitored the values of the variables in the model. *Kalemli-Ozcan et al. (2000)* analysed the impact of the change in the mortality rate in the context of variable and constant educational attainment, and this model also started out from a steady state. To mention some Hungarian examples, *Simonovits (2009)* and *Varga (2014)* upset the steady state of their models, which were very similar to each other in many respects, and examined how the parametric pension reforms affected the sustainability of the pension system.

We also assumed in our model that the economy was in a stationary state in the 0th period before modelling. Then the total fertility rates and mortality rates changed from period to period. We sought to find out how the changes shape GDP per capita. *Table 1* contains the values of s and F used in the first and second simulation calculations in the 16 periods under review, while *Figure 1* presents the development of GDP per capita corresponding to these (s,F) pairs.

The growth in GDP per capita along the first path clearly shows that if the initially very high total fertility rate continuously declines, even sinking well below the replacement level by the end of the period under review, GDP per capita may increase even in the context of a high survival rate. The survival rate increases steadily and at a relatively fast pace for 13 periods along the first path, and then during the final three periods, it stabilises at the high level achieved until then. GDP per capita grows even in these final three periods, although more moderately than before.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Path 1	F	3.50	3.15	2.84	2.56	2.31	2.08	1.87	1.69	1.52	1.37	1.23	1.11	1.00	0.90	0.81	0.73
	s	0.20	0.22	0.23	0.25	0.27	0.29	0.32	0.34	0.37	0.40	0.43	0.47	0.48	0.48	0.48	0.48
Path 2	F	3.50	3.15	2.84	2.56	2.31	2.08	1.87	1.69	1.52	1.37	1.23	1.11	1.00	1.05	1.10	1.16
	s	0.20	0.22	0.23	0.25	0.27	0.29	0.32	0.34	0.37	0.40	0.43	0.47	0.48	0.48	0.48	0.48

Figure 1
GDP per capita values for Growth Paths 1 and 2 according to Table 1



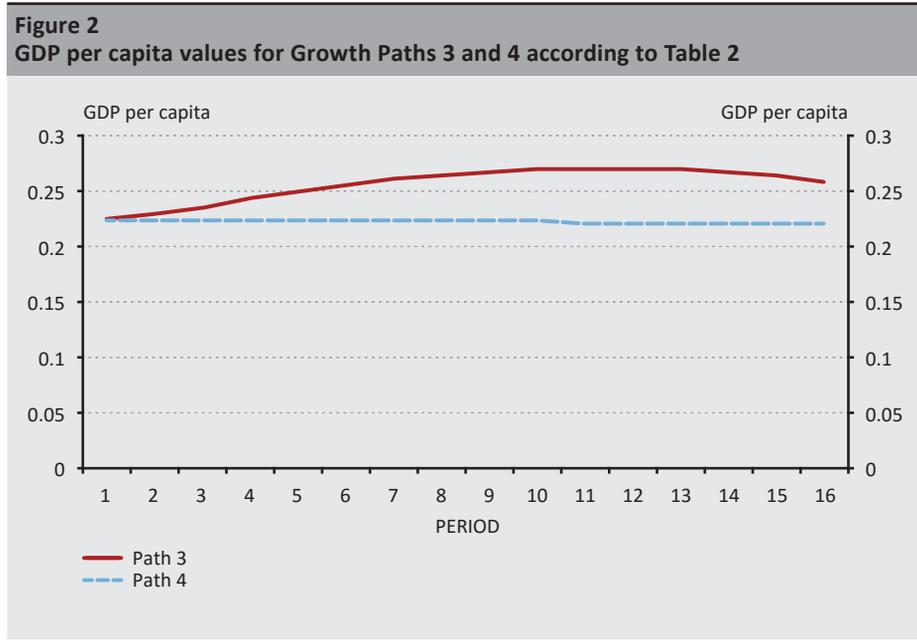
The simulation was performed on Path 2 using the same survival rates as in Path 1, and the total fertility rates were also almost the same (*Figure 1*). The total fertility rate starts out from the same high level in the first period and then falls to the replacement level, just like along Path 1. However, upon reaching the replacement level, it starts increasing slightly instead of continuing on its downward path. In parallel with the hike in the total fertility rate, GDP per capita drops instead of growing as before. Therefore, our model suggests that if we use only the total fertility rate and the mortality (survival) rates as exogenous variables, the continuous contraction of the total fertility rate to any low level is favourable from an economic growth perspective under all conditions. Yet the simulation results along Paths 3 and 4 (*Table 2*) highlight the fact that the drop in the total fertility rate can reach a drastic level in our model where it entails a decrease in GDP per capita.

Table 2
Total fertility and survival rates along the third and fourth simulation paths

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Path 3	F	1.20	1.14	1.09	1.04	0.99	0.94	0.90	0.85	0.81	0.77	0.74	0.70	0.67	0.64	0.61	0.59
	s	0.40	0.41	0.41	0.42	0.42	0.43	0.44	0.44	0.45	0.45	0.46	0.47	0.47	0.48	0.49	0.49
Path 4	F	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20
	s	0.40	0.41	0.41	0.42	0.42	0.43	0.44	0.44	0.45	0.45	0.46	0.47	0.47	0.48	0.49	0.49

If the already high survival rate steadily increases, in an extreme case assuming for the last period that almost all 60-year-olds will live until retirement age, and thus

also live to be 80 in our model, GDP per capita falls in the context of an excessively contracting total fertility rate. Using the same hypothetical survival rate for our simulation calculations, but employing constant total fertility rates slightly above the replacement level, GDP per capita exhibits only a minimal decline. Nevertheless it seems that the development of GDP per capita is most affected in our model by the growth or decrease in the total fertility rate. This is also suggested by the simulation along Paths 5 and 6 (Figure 3).

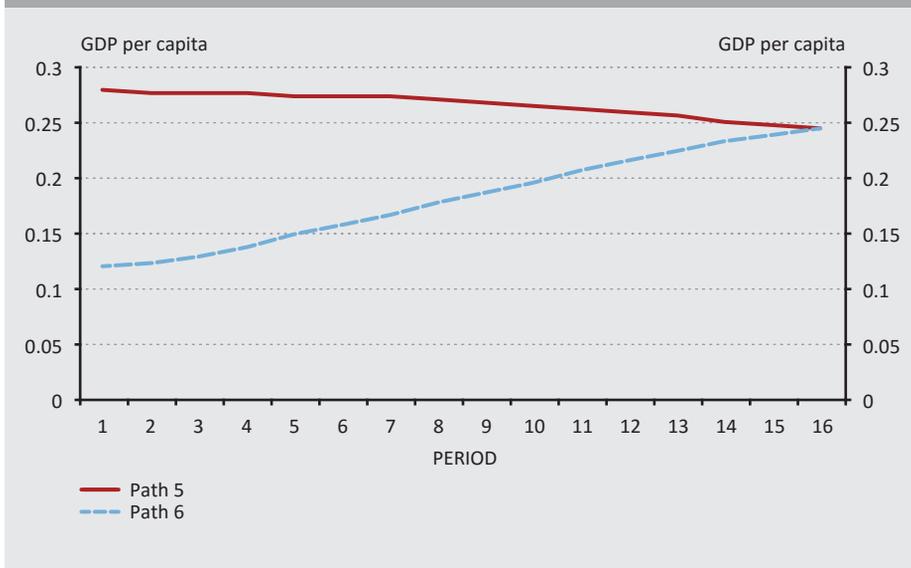


The exogenous total fertility rates and survival rates of Paths 5 and 6 were provided based on a special consideration. The survival rates are the same in both versions, starting out from a moderately high level, and growing relatively slowly but steadily, and only the total fertility rates are different. Along Path 5, the total fertility rates that start out from below the replacement level increase continuously, right until the last period, where they are slightly over the replacement level. However, along Path 6, the total fertility rates that start out from a level more than twice the replacement level drop continuously. As a result of the above, the GDP per capita values along Paths 5 and 6 are quite different from each other. They continuously decline along Path 5, while they steadily grow along Path 6. The initial value is very low on Path 6, but after a consistent rise, it is only infinitesimally lower in the last period than the corresponding value on Path 5. The levels achieved on the paths should be compared to the values along the same path – one only need to think of the hurdles described at the beginning of the chapter – however, the contrasting development of the GDP values in the first and last periods clearly shows the opposing dynamics of the two paths.

Table 3
Total fertility and survival rates along the fifth and sixth simulation paths

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Path 5	F	0.65	0.67	0.69	0.72	0.74	0.77	0.79	0.82	0.85	0.88	0.91	0.94	0.97	1.00	1.04	1.05
	s	0.35	0.36	0.36	0.37	0.37	0.38	0.38	0.39	0.39	0.40	0.41	0.41	0.42	0.42	0.43	0.44
Path 6	F	2.10	2.00	1.90	1.81	1.73	1.65	1.57	1.49	1.42	1.35	1.29	1.23	1.17	1.11	1.06	1.05
	s	0.35	0.36	0.36	0.37	0.37	0.38	0.38	0.39	0.39	0.40	0.41	0.41	0.42	0.42	0.43	0.44

Figure 3
GDP per capita values for Growth Paths 5 and 6 according to Table 3



5. Summary

Our study summarised how, after the initially simple approaches, the various types of models analysing economic growth took into account the changes in the size and composition of the population. From the perspective of our present article, the simple OLG model presented in *Lee – Mason (2010)* is key. We performed our simulation calculations with a model that can be considered an enhanced version of *Lee – Mason (2010)*.

Our simulation results clearly prove that we can draw many valuable conclusions even if, similar to the model by Lee and Mason, we assume that the only factors that determine the path of the OLG model are the total fertility rate and the survival rate. In such cases, the implicit level of physical capital is fixed, and the increasing efficiency of physical capital materialises in the expansion of the efficiency of

human capital investments. Therefore, output per capita in itself does not provide meaningful information, however, the comparison of the levels by periods taught us many lessons.

The general trends observed in the model included the following:

- The change in fertility influenced GDP per capita much more than the shifts in the survival rate.
- The survival rate had a substantial impact when the total fertility rate was well below the replacement level. When the total fertility rate per person (for both men and women) was slightly below 0.6 after declining steadily from period to period, while the survival rate continuously increased, GDP per capita started to dip. In such cases, the survival rate, which increased and approximated 1 (meaning an s value converging towards 0.5 in our model at 60–80 years of age), prevented the growth of GDP per capita.
- When comparing two simulation calculations in which the survival rates were the same in each period, we mostly saw that the increasing and more steeply increasing GDP per capita values were on the path along which the total fertility rates declined or declined more, even to below the replacement level. However, it was important not to let the total fertility rates fall as low as we described in the previous paragraph.
- When on one of two model paths containing the same survival rates the total fertility rate increased from period to period from a low initial value, and it declined from an initial high level on the other path, the first path indicated a steadily decreasing output per capita, while the other exhibited a consistently expanding value, even if the total fertility rates were the same in the last period. These two simulations showed the most clearly that such models “reward” the reduction in population and “punish” population growth.

No hasty conclusions should be drawn from the above. For example we cannot say that a drop in fertility benefits economic growth, since the combined effect of several other factors also needs to be taken into account in connection with this issue. We should also not forget that if the Beckerian quantity–quality trade-off does not take hold, for example the parents do not provide substantial human capital investments to a low number of children, all the conclusions of our model instantly become unrealistic. Nonetheless, based on the simulation calculations of our model, in most cases it can be assumed that the gradual decrease of the total fertility rate to slightly below the replacement level has a beneficial effect on the development of GDP per capita.

Appendix

A.1. Demographics in growth models

Since ours is a demographic growth model, the role of demographic variables in the development of growth models should be summarised. This can be seen *Table A1*.

Table A1			
The role of population in the various growth models			
Type of model	Creators of the model	The way population is taken into account	Size of the population, and its role in growth
Classical	Thomas Robert Malthus	Population grows more rapidly than the amount of food.	Excessive population growth leads to lower prosperity.
Keynesian	Roy F. Harrod, Evsey Domar	Exogenous savings ratio, no consumer optimisation.	Economic growth does not depend on it.
Neoclassical	Robert M. Solow, Trevor W. Swan	There is no consumer optimisation, the labour force and productivity of the population influences economic output.	The change in exogenous population growth influences the development of income per capita only in the period of convergence towards the equilibrium growth path, and does not do so in the long term.
Neoclassical	Frank P. Ramsey, David Cass, Tjalling C. Koopmans	Households decide on their consumption and savings path, maximising their own utility.	Short-term dynamics diverge from the Solow–Swan model, but the increase in income per capita does not depend on exogenous population growth in the long run.
Endogenous growth	Kenneth J. Arrow, Paul M. Romer, Robert E. Lucas, Sergio Rebelo	The development of workers' productivity is endogenous.	Using certain parametrisation, the externality and R&D based models exhibit a positive relationship between the exogenous population growth rate and the increase in GDP per capita.
Overlapping generations	Paul A. Samuelson, Peter A. Diamond, Alan J. Auerbach, Laurence J. Kotlikoff	There are several overlapping generations, maximising their lifetime utility.	Population growth is exogenous, but the composition of the population may vary, and there are transactions across generations.
Endogenous fertility	Gary S. Becker, Robert J. Barro	Utility-maximising households also decide about the number of offspring.	Fertility develops in line with the optimal consumer decision.
Growth models based on national transfer accounts	Ronald Lee, Andrew Mason	Consumption and income indicators by age.	The impact of the first and second demographic dividend on growth depends on the composition of the population.

A.2. Parameter values used during the simulation

Table A2 Parameters of the model	
Value	Source
$\alpha = 0.075$	Authors' calculation based on the NTA (2016) database ¹
$\gamma = 1$	Lee – Mason (2010)
$\delta = 0.33$	Mankiw et al. (1992), Lee – Mason (2010)
$\varphi = 1.128$	Authors' calculation based on the NTA (2016) database ²

¹ The values of all the countries, 19 in all, for which the NTA (2016) contained complete data were taken into account (Austria, Brazil, Costa Rica, Finland, Hungary, India, Indonesia, Japan, Kenya, Mexico, Nigeria, Philippines, Slovenia, South Korea, Spain, Sweden, Taiwan, Thailand, United States). We found that the human capital investment rate (the share of education spending per child within the income of a 21–40-year old and that of a 41–60-year-old) was 0.075 on average.

² The information from the 19 countries with complete data were taken into account (NTA 2016) here as well. We saw that the earned income of the workers in the age group of 41–60 was higher by 12.8 per cent on average than that of their peers in the age group of 21–40.

References

- Adelman, I. (1963): *An Econometric Analysis of Population Growth*. The American Economic Review, 53(3): 314–339.
- Auerbach, A.J. – Kotlikoff, L.J. (1987): *Dynamic fiscal policy*. Cambridge: Cambridge University Press.
- Barro, R.J. – Becker, G.S. (1989): *Fertility Choice in a Model of Economic Growth*. *Econometrica*, 57(2): 481–501. <https://doi.org/10.2307/1912563>
- Becker, G.S. (1960): *An Economic Analysis of Fertility*. In: National Bureau of Economic Research, (Ed): *Demographic and Economic Change in Developed Countries*, pp. 209–240.
- Becker, G.S. – Lewis, H.G. (1973): *On the Interaction between the Quantity and Quality of Children*. *Journal of Political Economy*, 81(2): S279–S288. <https://doi.org/10.1086/260166>
- Becker, G.S. – Murphy, K.M.– Tamura, R. (1990): *Human Capital, Fertility, and Economic Growth*. *Journal of Political Economy*, 98(5): S12–37. <https://doi.org/10.1086/261723>
- Black, S.E. – Devereux, P.J. – Salvanes, K.G. (2005): The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *The Quarterly Journal of Economics*, 120(2): 669–700. <https://doi.org/10.1093/qje/120.2.669>
- Bloom, D. – Canning, D. – Sevilla, J. (2003): *The demographic dividend: A new perspective on the economic consequences of population change*. Rand Corporation.

- Cipriani, G.P. (2014): *Population aging and PAYG pensions in the OLG model*. Journal of Population Economics, 27(1): 251–256. <https://doi.org/10.1007/s00148-013-0465-9>
- Colleran, H. – Jasienska, G. – Nenko, I. – Galbarczyk, A. – Mace, R. (2015): *Fertility decline and the changing dynamics of wealth, status and inequality*. Proceedings of the Royal Society B: Biological Sciences, 282(1806).
- Dang, H.-A.H. – Rogers, F.H. (2016): *The Decision to Invest in Child Quality over Quantity: Household Size and Household Investment in Education in Vietnam*. The World Bank Economic Review, 30(1): 104–142.
- Diamond, P.A. (1965): *National Debt in a Neoclassical Growth Model*. The American Economic Review, 55(5): 1126–1150.
- Easterlin, R.J. (1973): *Does Money Buy Happiness?* The Public Interest, 30: 3–10.
- Ellis, J. (2008): *Culture, fertility, and son preference*. PhD thesis. The London School of Economics and Political Science (LSE).
- Freedman, D.S. (1963): *The Relation of Economic Status to Fertility*. The American Economic Review, 53(3): 414–426.
- Freedman, R. – Coombs, L. (1966a): *Childspacing and Family Economic Position*. American Sociological Review, 31(5): 631–648. <https://doi.org/10.2307/2091855>
- Freedman, R. – Coombs, L. (1966b): *Economic Considerations in Family Growth Decisions*. Population Studies, 20(2): 197–222. <https://doi.org/10.2307/2172982>
- Frejka, T. (2016): *The demographic transition revisited: a cohort perspective*. Max Planck Institute for Demographic Research, Rostock, Germany, working paper WP–2016–012.
- Gál, R.I. – Vanhuyse, P. – Vargha, L. (2016): *Pro-elderly welfare states within pro-child societies: Incorporating family cash and time into intergenerational transfers analysis*. Center for Economic Institutions, Working Paper No 2016–6.
- Galor, O. – Weil, D.N. (1999): *From Malthusian Stagnation to Modern Growth*. The American Economic Review, 89(2): 150–154. <https://doi.org/10.1257/aer.89.2.150>
- Guo, R. – Yi, J. – Zhang, J. (2017): *Family Size, Birth Order, and Tests of the Quantity–Quality Model*. Journal of Comparative Economics, 45(2): 219–224. <https://doi.org/10.1016/j.jce.2016.09.006>
- Kalemli-Ozcan, S. – Ryder, H.E. – Weil, D.N. (2000): *Mortality decline, human capital investment, and economic growth*. Journal of Development Economics, 62(1): 1–23. [https://doi.org/10.1016/S0304-3878\(00\)00073-0](https://doi.org/10.1016/S0304-3878(00)00073-0)

- Kaplan, H. (1994): *Evolutionary and Wealth Flows Theories of Fertility: Empirical Tests and New Models*. Population and Development Review, 20(4): 753–791. <https://doi.org/10.2307/2137661>
- Lawson, D.W. – Borgerhoff Mulder, M. (2016): *The Offspring Quantity–Quality Trade-off and Human Fertility Variation*. Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1692).
- Lee, R. – Mason, A. (2010): *Fertility, Human Capital and Economic Growth over the Demographic Transition*. European Journal of Population / Revue Européenne de Démographie, 26(2): 159–182. <https://doi.org/10.1007/s10680-009-9186-x>
- Luci, A. – Thevenon, O. (2010): *Does economic development drive the fertility rebound in OECD countries?* Working paper No. hal-00520948, HAL.
- Malthus, T. (1798): *An Essay on the Principle of Population, as it Affects the Future Improvement of Society with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers*. J. Johnson in St. Paul’s church-yard. London.
- Mankiw, N.G. – Romer, D. – Weil, D.N. (1992): *A Contribution to the Empirics of Economic Growth*. The Quarterly Journal of Economics, 107(2): 407–437. <https://doi.org/10.2307/2118477>
- Mason, A. (2005): *Demographic transition and demographic dividends in developed and developing countries*. In United Nations expert group meeting on social and economic implications of changing population age structures, vol. 31. Mexico City.
- Mason, A. – Lee, R. – Jiang, J.X. (2016): *Demographic dividends, human capital, and saving*. The Journal of the Economics of Ageing, 7: 106–122. <https://doi.org/10.1016/j.jeoa.2016.02.004>
- NTA (2009): National Transfer Accounts, <http://ntaccounts.org>. Downloaded: 1 June 2014.
- NTA (2016): National Transfer Accounts, <http://ntaccounts.org>. Downloaded: 10 June 2016.
- Roudi-Fahimi, F. – Kent, M.M. (2007): *Challenges and Opportunities – The Population of the Middle East and North Africa*. Population Bulletin, 62(2).
- Silver, M. (1965): *Births, Marriages, and Business Cycles in the United States*. Journal of Political Economy, 73(3): 237–255. <https://doi.org/10.1086/259013>
- Simonovits, A. (2009): *Néességöregedés, tb-nyugdíj és megtakarítás – parametrikus nyugdíjreformok (Population aging, the public pension system and savings: Parametric pension reforms)*. Közgazdasági Szemle, 56(4): 297–321.
- Smith, A. (1776): *An inquiry into the wealth of nations*. London: Strahan and Cadell.

- Sobotka, T. – Skirbekk, V. – Philipov, D. (2011): *Economic Recession and Fertility in the Developed World*. Population and Development Review, 37(2): 267–306. <https://doi.org/10.1111/j.1728-4457.2011.00411.x>
- Solow, R.M. (1956): *A Contribution to the Theory of Economic Growth*. The Quarterly Journal of Economics, 70(1): 65–94. <https://doi.org/10.2307/1884513>
- UNESCO (2016): *Education: Mean Years of Schooling*. <http://data.uis.unesco.org/Index.aspx?queryid=242>. Accessed: 20 May 2016.
- UN (2015): United Nations, Department of Economic and Social Affairs, Population Division: *World Population Prospects: The 2015 Revision, DVD Edition*.
- Van de Kaa, D.J. (2010): *Demographic transition*. Encyclopedia of Life Support Systems, 1, pp. 65–103.
- Varga, G. (2014). *Demográfiai átmenet, gazdasági növekedés és a nyugdíjrendszer fenntarthatósága (Demographic transition, economic growth and sustainability of the pension system)*. Közgazdasági Szemle 61(11): 1279–1318.
- Vargha, L. – Donehower, G. (2016): *The Quantity-Quality Tradeoff: A Cross-Country Comparison of Market and Nonmarket Investments per Child in Relation to Fertility*. Paper presented at the European Population Conference, 1–3 September 2016, Mainz.
- Willis, R.J. (1973): *A New Approach to the Economic Theory of Fertility Behavior*. Journal of Political Economy, 81(2), pp. S14–S64. <https://doi.org/10.1086/260152>

Fiscal Policy and the Business Cycle

György Molnár – Gábor Dániel Soós – Balázs Világi

After the 2007–2008 financial crisis and the subsequent prolonged, deep recession, the question of whether fiscal policy can be applied to stimulate the economy came into focus, as conventional tools of monetary policy became ineffective. Accordingly, the related research also received a new boost. The purpose of our paper is to present the new research results, which mainly focus on whether fiscal policy can be applied to influence business cycles. Although there is still much debate, the claim that the effects of fiscal policy are not constant, but rather depend on the state of the business cycle, has been supported by numerous theoretical models and empirical results: in prolonged, deep recessions, and especially in the case of a liquidity trap, the multiplier of government expenditures is high, while in the case of booms it is relatively low.

Journal of Economic Literature (JEL) codes: E62, E52, E21, E12

Keywords: fiscal policy, expenditure multiplier, business cycles, liquidity trap, debt-constrained households

1. Introduction

Over the last eighty years the position of economics on the applicability of fiscal policy has changed significantly. Owing to the works of Keynes, it became the prevailing opinion in the 1930s that fiscal policy is suitable to boost the economy during times of recession, and from a more general approach, to influence business cycles. In the 1970s new classical economics formulated the exact opposite of this view, and the doubts about the fiscal policy gradually reshaped economic thinking.

After the 2007–2008 financial crisis and the deep, prolonged recession that followed it, fiscal policy came into focus again, as the traditional tools of monetary policy reached their constraints. As a result, the related theoretical and empirical research also received a new boost, and the purpose of our paper is to present these new research results.

György Molnár was an analyst at the Magyar Nemzeti Bank at the time of writing the study.

E-mail: gyorgy.molnar@ogresearch.com

Gábor Dániel Soós is a senior economic analyst at the Magyar Nemzeti Bank. E-mail: soosg@mnb.hu

Balázs Világi is a Head of Department at the Magyar Nemzeti Bank. E-mail: vilagib@mnb.hu

The Hungarian manuscript was received on 20 December 2016.

DOI: <http://doi.org/10.25201/FER.16.4.5885>

Although there is still much debate about the applicability of fiscal policy, the view that the impacts of fiscal policy are not constant, but rather depend on the state of the business cycle, has become increasingly common. Fiscal expenditures are capable of influencing the movement of real GDP, to a higher extent during recessions and to a lesser extent during booms.

Since our paper mainly focuses on the presentation of new research, and the new research efforts primarily deal with the circumstances under which fiscal policy is suitable to affect the business cycles, in this study we do not consider the long-run effects of fiscal policy and only touch on the problem of government debt sustainability. We do not deny the importance of these problems, but also with a view to volume constraints, we have intentionally narrowed the subject of our paper and focused primarily on the new research which has not yet found its way to the broader community of professionals.

The study is structured as follows: In *Section 2* we review how the theory of fiscal policy has changed since Keynes and the current position of economics on this subject. In *Section 3* we discuss the problem of the empirical identification of the effects of fiscal policy and present empirical results. *Section 4* deals with the implication of the above in economic policy, and *Section 5* concludes.

2. Theory

The subject of the paper is the assessment of whether fiscal policy is capable of influencing business cycles. Therefore, we disregarded other issues belonging to fiscal policy, such as the role of the state in the creation of public goods or how the investments of the state influence the long-term growth potential of the economy.

The degree of efficiency at which fiscal policy is capable of influencing business cycles is characterised by the magnitude of the so-called expenditure multiplier. The expenditure multiplier measures how many units real GDP will be raised by a raise of one unit of real government expenditure.

If the value of the multiplier is zero, then each unit of the increase in public spending will reduce private consumption or investments by the same amount. In other words, public spending will crowd out private spending. In extreme cases the multiplier may even be a negative number, in such cases one unit of public spending will crowd out more than one unit of private expenditure. In such cases fiscal spending will not increase and might even reduce real GDP; therefore if this happens, the fiscal policy is entirely ineffective.

If the expenditure multiplier is between 0 and 1, then government expenditures will only partially crowd out private expenditures, and an increase in public spending by one unit will reduce private consumption and investments by less than one unit. In this case, it is possible to increase real GDP by fiscal spending. However, in this case in parallel with the increase in GDP social welfare could decrease, if the government spending crowds out a significant volume of private consumption, or it may decrease long-run growth, if it crowds out important private investments. Therefore, in such a case the application of fiscal policy to boost the economy requires substantial caution, and the benefits and costs should be carefully assessed.

If the value of the multiplier is 1, then fiscal policy no longer crowds out private expenditures at all, and if it is higher than 1, then government expenditures not only do not crowd out private spending, they even generate additional private income. In such cases fiscal policy should clearly be used to boost the economy. In the remaining part of the study, when we use the term *fiscal policy is effective*, we refer to cases when the value of the expenditure multiplier is higher than 1.

Our paper primarily focuses on the magnitude of the expenditure multiplier, since after the crisis the main debated matter was whether the increase in government spending is able to mitigate the recession. In addition to the expenditure multiplier, the literature deals with the negative multiplier effects of various taxes in detail. We only assess this matter partially, primarily in the context of the extent to which various forms of funding for government expenditures could reduce the effect of the expenditure multiplier.

2.1. The Keynesian theory of fiscal policy

Today, Keynesian theory is an essential part of macroeconomic textbooks. For example, the system of IS-LM curves and its implications for fiscal policy are dealt with in detail in *Chapter 2* of the textbook of *Benczes and Kutasi (2010)*. Precisely because of that, following overview is not intended to convey new knowledge; our purpose is merely to prepare the presentation of new theories discussed later. The purpose of the following section is not to provide a formal presentation of the IS-LM curves, but rather to clarify the key assumptions of the Keynesian theory from which its results applying to fiscal policy are derived.

Before the 1930s it was the prevailing opinion that fiscal policy has no effect on GDP, since government spending crowds out private expenditures.¹ After the recession of 1929–1933, it was the works of Keynes that changed the majority opinion on that matter.

¹ Nevertheless, classical theory did not completely reject the idea of accumulating deficits; in extraordinary cases, such as during wartime, it was considered acceptable.

According to Keynes, one unit of real expenditure by the government raises real GDP by more than one unit. The arguments of Keynes are well known from introductory level macroeconomics textbooks. One unit of government spending will increase GDP by one unit² in the first round, but this is not yet the end of the process. The increase in GDP will raise the incomes of households, and since according to Keynes, consumption moves closely in parallel with current incomes, the consumption of households and consequently, aggregate demand will also increase, which will further boost GDP, as a result of which the consumption of households will increase, and so forth (see *Figure 1*).

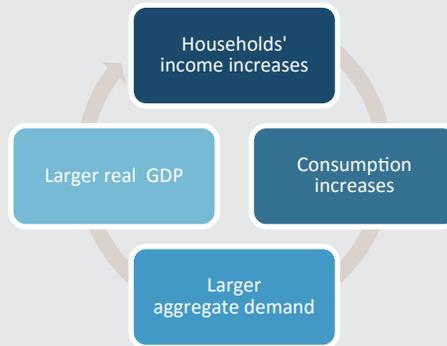
The relationship between the current income and consumption of households is captured by the *marginal propensity to consume*. This shows how many units of consumption are increased by one unit of current income. If we mark marginal propensity to consume with c , $0 < c < 1$, then according to Keynes one unit of government expenditure will increase GDP by $1/(1-c)$. We call this latter formula the Keynesian expenditure multiplier, since c is positive and lower than 1, the Keynesian multiplier is higher than 1, furthermore, the higher the marginal propensity to consume, the higher the value of the multiplier.

The co-movement of current income and consumption is of key importance in Keynesian theory, but we also need other conditions to demonstrate that fiscal policy is capable of efficiently increasing economic activity.

Essentially, Keynesian theory is the theory of aggregate demand, and assumes that the development of aggregate demand determines GDP. In other words, supply does not respond to increasing aggregate demand by raising prices, but rather by quantitative expansion. To express this in the language of macroeconomics, the aggregate supply curve is horizontal: therefore, as a result of a shift in the aggregate demand curve, in the new equilibrium output will be higher while price level will remain unchanged.

² To make the essence of this argument more accessible, following the simplification generally applied in the literature of macroeconomics, in this paper we presume all along that government expenditures include government consumption and government investments. We disregard household and corporate transfers, which are also part of government expenditures in practice.

Figure 1
The expenditure multiplier and current income



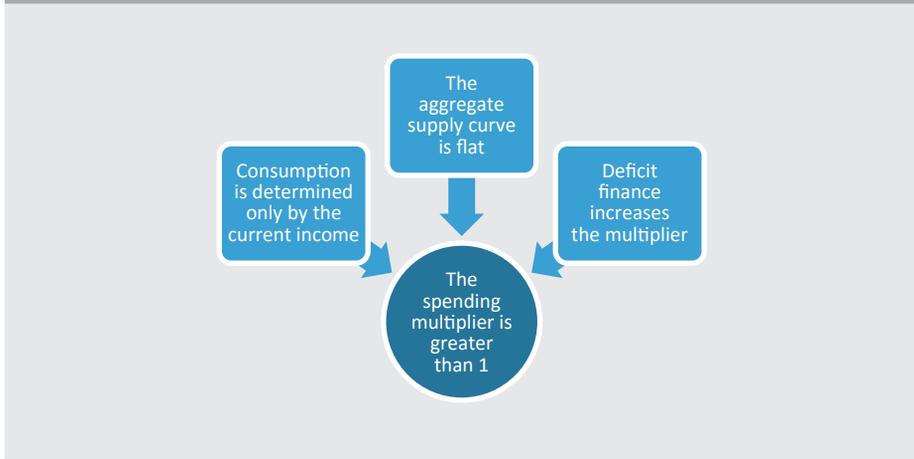
Since Keynesian theory was created as a reaction to the Great Depression of the 1930s, it was a natural assumption that supply adjustment is primarily quantitative, since the depression economy of this era was, in fact, characterised by the fact that both prices and wages were inelastic, with a significant amount of idle capacities.

In addition to supply assumptions, the assumptions about the methods of financing government expenditures also affect the rate of the multiplier. The rate of the expenditure multiplier is affected by the methods used to finance expenditures. The multiplier only reaches its maximum value of $1/(1-c)$ if the government finances budgetary expansion from debt. If the expenditures are financed partly from tax, it will reduce the rate of the multiplier, since the disposable incomes of households will be lower, and therefore consumption growth and aggregate demand will be lower. It can be acknowledged that if the expansion of government expenditures is fully financed from taxes, then the rate of the multiplier will be reduced to 1. *Figure 2* summarises the underlying factors of the Keynesian expenditure multiplier theory.

Finally, a common misunderstanding related to Keynesian theory should be clarified. According to the theory, the rate of the multiplier is independent of whether the government spends the fiscal expenditures on useless things, or on useful purposes, such as infrastructure investments or the development of human capital. Keynesian theory was not intended to explain factors determining the long-term growth of the economy, but rather meant to find a way for recovery from the depression when aggregate demand is insufficient and monetary policy is no longer capable of stimulating the economy. The key to the Keynesian theory is that government expenditure generates additional private incomes, and from this point of view it makes no difference what the targets of the initial government expenditures are

that induce private demand. Naturally, the target of the government expenditures may have an impact on the long-term growth of the economy, since efficient public spending may increase potential output for the future. However, the Keynesian theory does not consider this question, as the subject of its assessment is the short-term effects of fiscal policy.

Figure 2
The Keynesian expenditure multiplier



2.2. The criticism of Keynesian theory

As discussed in the previous section, from the aspect of the conclusions of Keynesian theory one fundamental assumption is that there is a close relationship between consumption and current incomes. However, subsequently severe criticism was formulated about the consumption theory of Keynes, with the most important challenges posed by the permanent income theory of *Milton Friedman (1957)* and the life cycle hypothesis of *Franco Modigliani (1966)*. Both authors established empirically and theoretically that when making consumption-saving decisions, households consider not only the development of their current income but also the development of their incomes in the longer term.

To put it in another way, households *smooth their consumption*, i.e. fluctuation in their income is not fully followed by fluctuation in their consumption. This is also supported by the data, as the volatility of aggregate consumption is lower than the volatility of the GDP. Smoothing consumption specifically means that when households temporarily have lower incomes, they reduce their savings and try to maintain the level of their consumption. And if their income rises temporarily, then they will not spend their total increase in income on current consumption, but rather save part of it so that they can increase their consumption in the future as well.

It follows from the above that the marginal propensity to consume is lower than had been presumed by Keynes. As noted, the rate of the Keynesian expenditure multiplier is $1/(1-c)$, if marginal propensity to consume, i.e. if c is low, then the value of the multiplier will also be close to 1.

Discarding the Keynesian consumption function also has further consequences. The Keynesian theory of consumption is basically static: current consumption is only affected by current income. By contrast, subsequent consumption models are dynamic: the household develops its optimal consumption path in consideration of a longer time horizon. Accordingly, in addition to income, the real interest rate also has a significant influence on consumption-saving decisions. Since monetary policy is capable of influencing the real interest rate, it therefore has a significant impact on both consumption and aggregate demand. As a result, monetary policy can neutralise the effect of fiscal policy. Let us assume that the decisions of the central bank can be described by the following interest rule: in the case of inflation or an increase in real GDP the central bank raises the interest rate. If fiscal expenditures increase, inflation and output will also be higher, as a result, the central bank will raise the interest rate, which will result in the increase in the real interest rate, which will in turn, reduce consumption and aggregate demand. As result, the efficiency of fiscal policy will decrease, and the expenditure multiplier might be lower than 1.

From the 1970s, Keynesian theory has been subjected to even more radical criticism from the new classical school, associated with the name of Robert Lucas. This school of economic thinkers questioned the Keynesian opinion on the behaviour of supply. In their opinion, prices and wages respond flexibly to changes in aggregate demand, and therefore the adjustment of supply is basically not quantitative. In their opinion, the aggregate supply curve is basically vertical: it only departs from vertical temporarily, if firms face unforeseen surprises (*Lucas 1973; Sargent and Wallace 1975*). As a result, if the aggregate demand curve shifts due to a predictable fiscal expansion then it will have no quantitative effect and real GDP will not change; by contrast, the price level will increase. In other words, in such a case the value of the expenditure multiplier will be zero. If the fiscal expansion takes firms by surprise, then supply will respond quantitatively as well, for a temporary period, but even in this case the growth of GDP will be much lower than the magnitude assumed by Keynes.

The views of the new classical theory on the behaviour of aggregate supply can be formally summarised by the new classical Phillips curve, which is discussed, for example, in Chapter 4 of *Benczes and Kutasi (2010)*.

The other important element of new classical criticism was the Ricardian equivalence, which was first formulated by *Barro (1977)*. It claims that if government spending is financed from deficit, it has exactly the same effect as that of tax

financing. This is supported by the following argument: According to the theory of permanent income, the consumption decisions of households are not determined by current income, but by the present value of expected lifetime income. If the consumers are appropriately rational, then they will also exactly understand that a current government deficit will imply higher taxes in the future, since government debt will have to be repaid in the future. In other words, if the fiscal expansion is not financed from current taxes, then it will be financed from future taxes. It is clear that the present value of the tax increase in the future is just as high as the current tax increase; therefore, in both cases the present value of the lifetime income will be reduced by the same extent, and accordingly, current consumption will decrease by the same amount.

It is important to underline that the Ricardian equivalence has the consequence that the rate of the expenditure multiplier is independent of the way the government expenditures are financed. However, this does not imply in itself that the value of the expenditure multiplier should be less than 1. Ricardian equivalence only claims that whatever the rate of the multiplier when the expenses are fully financed from taxes, it is not possible to increase this rate further. Therefore, the Ricardian equivalence will only corroborate the inefficiency of fiscal policy if the multiplier of expenditures financed from taxes is less than 1. For lack of the Ricardian equivalence, in such cases it could be argued that in the case of deficit financing this value could even be higher than 1, but if we accept the Ricardian equivalence, this cannot happen. Therefore, the Ricardian equivalence can only be used to support a sceptical position towards fiscal policy if we supplement it with other arguments, for example, those related to the behaviour of supply.

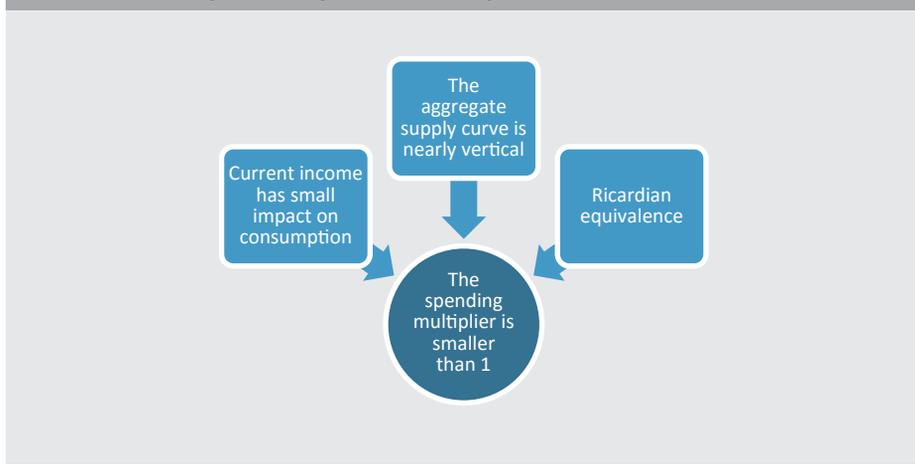
The critiques above are summarised in *Figure 3*. They were fundamentally intended to demonstrate that fiscal policy is not efficient and that fiscal spending of one unit will increase real GDP by less than one unit, because it crowds out private expenditures. But there are also counter-arguments related to the practical implementation of fiscal policy, which we review in the following section.

One frequent counter-argument is that monetary policy serves to smooth the volatility of the economy much more efficiently than does fiscal policy. This argument posits that monetary policy is capable of responding quickly and efficiently to changes in the economy: the decision-making body of a central bank can meet at any time and make a decision on the policy rate within a few hours, while the discretionary fiscal policy supplementing the automatic fiscal stabilisers³ is the result of a slow and sluggish legislative process, and therefore it is only able to respond to developments in the economy with long delays.

³ We refer to automatic fiscal stabilisers when the budget balance changes automatically whenever the business cycle enters a new stage.

Furthermore, it not only takes a long time for a decision to be made, there is also a long time between the decision and the implementation of the fiscal project. Therefore, it can take several years until the effect of fiscal policy is realised, whereas in the case of monetary policy new interest rates can be implemented practically instantly after an interest rate decision. All of this can have the consequence that if we try to boost the economy by fiscal policy during a recession, the recession may well be over by the time results start to appear. This was considered an especially important problem in the two decades preceding the crisis, at the time of the so-called *Great Moderation*, when the depth of business cycles' volatility and the length of recessions decreased significantly in developed economies.⁴

Figure 3
Criticism of the Keynesian expenditure multiplier



2.3. Post-crisis view

So far we have summarised the opinion of economics on the efficiency of fiscal policy before the crisis. These results are known to the wider professional community. Our purpose with the review was to prepare the discussion of new research results and put these in context. Much of this research applies the stochastic dynamic framework of modern macroeconomics, and because this framework is technically demanding they still have not found their way into introductory textbooks or to the wider professional community. In the following section we attempt to present these results in an accessible way.

⁴ The idea that discretionary economic policy should not be pursued, owing to delayed effects, among others, can be traced back to Milton Friedman.

The crisis of 2007–08 and the recession that followed it forced the profession to reconsider its views on the role of fiscal policy. As a result, the profession now has much more complex opinions on the possibility of applying fiscal policy than previously.

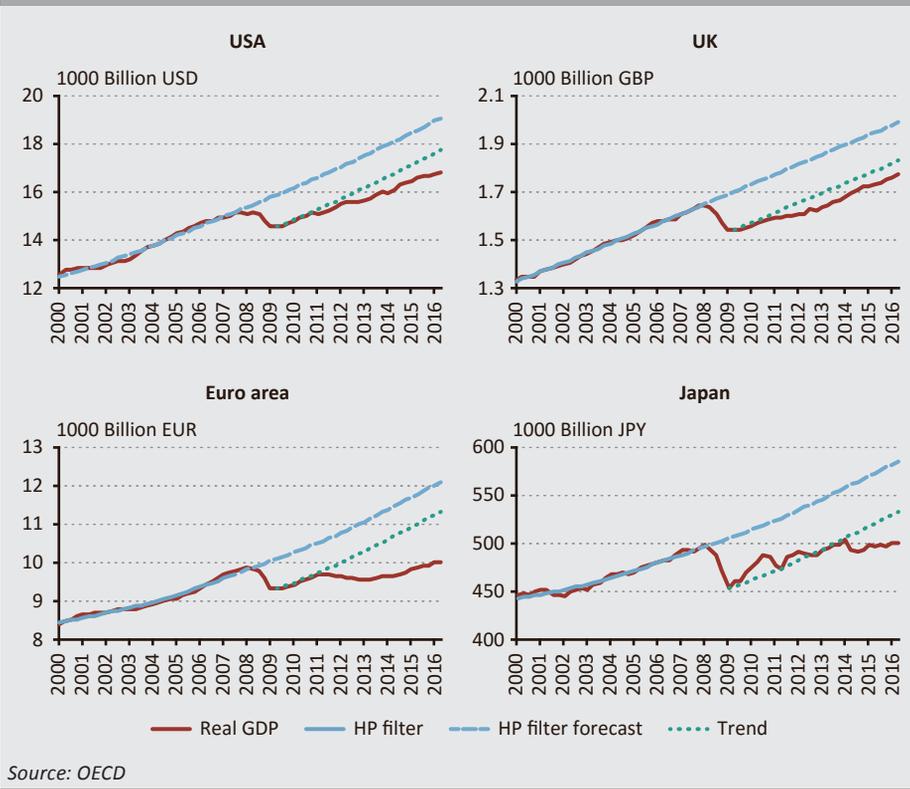
It is not like the pendulum has swung back from one extreme view into the other direction. The Keynesian theory is no longer acceptable today in its original form, but it has become clear that many of its elements can be once again relevant, under certain conditions. It has become clear that the applicability of fiscal policy is not a timeless property, but rather something that depends on the state of the business cycle. In this part, we overview this process of reconsideration.

First, we detail how the opinion of the profession has changed on the practical applicability of fiscal policy. On the one hand, it has been proven that monetary policy is not efficient in every situation. Keynes realised back in the 1930s that in a severe recession and in a low interest environment monetary policy becomes ineffective; this is what he called a liquidity trap. The liquidity trap is a rare phenomenon, which is why the profession did not take it seriously, although the long recession of Japan in the 1990s could have been a warning sign. However, after the crisis both the euro area and the United States found themselves in a liquidity trap, when the natural rate of interest, i.e. the real interest rate consistent with the natural rate of output, became negative, and monetary policy could not replicate this negative real interest rate, since – given the low inflationary environment – it would have required a significantly negative nominal interest rate, which is not feasible. Although it was possible to simulate the economy to some extent using unconventional tools, monetary policy basically reached its limits. In such a situation, the application of fiscal policy as an alternative tool was naturally raised.

Starting from the 1980s, the rate of economic volatility was lower than ever before. As mentioned, this period is called the Great Moderation in the literature. By contrast, since the Great Depression of the 1930s there has not been such a deep, prolonged recession in the developed world as the one that followed the year 2008. *Figure 4* shows the recession that unfolded in four developed economies. As illustrated in the figure, the economic downturn following the crisis was persistent and deep in all four countries. In addition, the post-crisis situation can be characterised by the concept of hysteresis: the economy grows on a lower trend line than before the crisis, or the growth rate even becomes lower than previously.⁵ Based on the above, it can be concluded that the argument formulated at the time of the Great Moderation, i.e. that owing to the short duration of recessions fiscal policy should not stimulate the economy, does not apply.

⁵ See the study of *Blanchard et al. (2014)* on hysteresis.

Figure 4
The recession of developed economies after 2008



Revaluation of the extreme views of new classical economics on the aggregate supply curve began long before the crisis. This process was not motivated by thinking about fiscal policy, but rather by that view of new classical macroeconomics that monetary policy has no effect on the real economy, which ran counter to both the empirical results of academic research and to the practical experiences of central bankers.

From the 1990s, in most macroeconomic models the behaviour of supply has been described by the New Keynesian Phillips curve (see its formal presentation in the advanced textbook of *Galí (2008)*). The New Keynesian Phillips curve is associated with the new classical Phillips curve discussed in the previous section only in name and is based on an entirely different concept. While in new classical theory the transitional frictions of supply can be explained by the imperfect information of

firms and households, the New Keynesian Phillips curve is based on *Calvo's (1983)* theory of sticky prices.⁶

The New Keynesian Phillips curve describes the behaviour of supply in a much more balanced way than the new classical school does. As a result of the assumption of sticky prices, some firms respond to an increase in demand by raising their prices, while another parts respond with quantitative expansion. As a result, the aggregate supply curve is rising, not horizontal, as in the original Keynesian theory, or vertical, as in the new classical models. Therefore, if aggregate demand shifts as a result of a fiscal expansion, then it will cause both the price level and real GDP to increase. The flatter the curve, the higher the quantitative shift is. Compared to the original Keynesian approach, the GDP effect and the expenditure multiplier will be lower, but if the slope of the New Keynesian Phillips curve is not too high, then the expenditure multiplier is higher than 1. This conclusion is radically different from the arguments of the new classical school, which state that fiscal expansion will only have a real impact if it is introduced by surprise, but even that impact will be transitional and negligible.

It is obvious from the above that the lower the slope of the New Keynesian Phillips curve, the higher the expenditure multiplier will be. In recessions, when there is low capacity utilisation, the significance of price adjustment is reduced, and consequently, the slope of the curve will also be lower. All of this supports the argument that in crises, if a fiscal expansion is capable of efficiently increasing aggregate demand, then much of it will be ultimately reflected in GDP growth, and the inflationary effect will be minimal.

There is also an extensive body of empirical literature that deals with the process that during recent decades the slope of the Phillips curve had significantly decreased long before the crisis, which is discussed in the review of *Szentmihályi and Világi (2015)*. The flattening of the Phillips curve started long before the crisis, but during the crisis this phenomenon intensified, because of the downward rigidity of wages and low capacity utilisation.

As discussed in the previous section, if consumption is affected not only by income, but also by the real interest rate, then it neutralises fiscal expansion to a significant extent. On the other hand, if during a crisis the economy falls into a liquidity trap, then monetary policy will reach the zero lower bound of nominal interest rates and the interest rates of the central bank will stabilise at around zero for a longer time. In such a case, monetary policy does not neutralise fiscal expansion. *Christiano et al. (2011)*, *Woodford (2011)*, and *Erceg and Linde (2016)* demonstrate that in a liquidity trap, if the nominal interest rates are unchanged, the expenditure multiplier will

⁶ See also the study of *Szentmihályi and Világi (2015)* on the new classical and New Keynesian Phillips curve.

be significantly higher than 1 even if the behaviour of households is described not by the Keynesian, but rather by the permanent income theory. In contrast with the static approach of Keynes, these papers present dynamic models, which shed light on channels increasing the multiplier that are lacking from the original Keynesian theory: according to these models, expected fiscal expansion will increase inflation expectations, as a result of which the real interest rate will decrease, which, in turn, raises current consumption and aggregate demand, and therefore the magnitude of the expenditure multiplier.

The importance of these studies cited above lies in the fact that these results are based on the dynamic stochastic general equilibrium (DSGE) models of modern macroeconomics. Typically, prior to the crisis the literature on DSGE was sceptical concerning the applicability of fiscal policy. However, the above authors demonstrated that in the case of a liquidity trap even the standard DSGE models imply the efficiency of fiscal policy.

In the following part, we present new results that are also derived using the stochastic, dynamic framework of modern macroeconomics, but extend the standard DGSE models. While in the standard models the only constraint to the indebtedness of households is their expected lifetime income, the studies presented now take into account that in reality, and especially at the time of crises, debt constraints are much stricter than that.

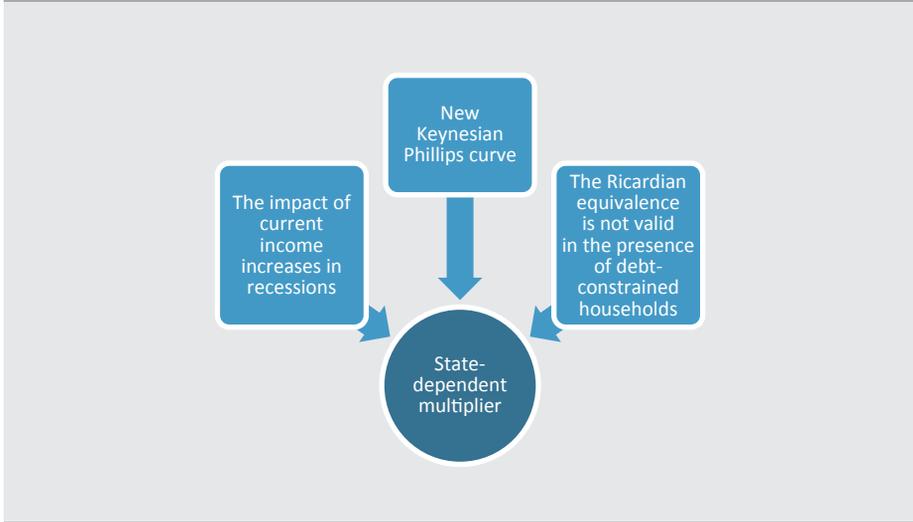
The studies of *Eggertsson and Krugman (2012)*, and *Eggertsson and Mehrotra (2014)* reconsider the relationship between current income and consumption: the authors point out that after financial crises, since the debt constraints become more stringent, households are not capable of smoothing their consumption, and therefore the permanent income theory does not work in the case of indebted households. Consequently, in a post-crisis recession the effect of current income on consumption will increase, i.e. the Keynesian consumption function becomes a good approximation of the behaviour of aggregate consumption. Based on that, the authors demonstrated that in a deep recession and liquidity trap the behaviour of the economy will become similar to that outlined in the original Keynesian theory, and therefore in such a case fiscal policy is effective and the value of the expenditure multiplier is higher than 1. Furthermore, the above studies highlight that if consumers with debt constraints exist, then the Ricardian equivalence will certainly not be valid, accordingly, it is possible to increase the magnitude of the multiplier if government expenditures are financed not from taxes, but rather by increasing the deficit.

Summarising the above points, although a significant portion of the criticism of the Keynesian theory of fiscal policy is justified, in deep, post-crisis recessions the major conclusions of Keynes are still valid. In such a case the economy falls into a

liquidity trap, which renders monetary policy ineffective; on the other hand, the effects of fiscal expansion will be strengthened, partly because these effects are not neutralised by the inefficient monetary policy, and partly because owing to the indebted consumers, consumption will be correlated with current incomes to a higher extent than in a normal situation. Furthermore, in a recession, due to slack, the downwardly rigid wages, the reaction of supply is also primarily quantitative.

Figure 5

The post-crisis position: a state-dependent expenditure multiplier



If we compare the above arguments to the conclusions of the criticisms of Keynesian theory, it is safe to say that – according to the current position of macroeconomic theory – the efficiency of fiscal policy and the expenditure multiplier are state-dependent. During booms the expenditure multiplier is lower than 1, and therefore fiscal policy is not effective, but there is no need to apply it since the business cycle can be appropriately controlled by monetary policy. By contrast, in recessions after crises the value of the expenditure multiplier is higher than 1, and therefore in such situations fiscal policy is an efficient tool to stimulate the economy and is necessary in this case, because the conventional tools of monetary policy become ineffective. The above train of thought is summarised in *Figure 5*.

2.4. Small open economies

The theoretical framework presented so far has been created fundamentally for the analysis of large economies, only open to a small degree, such as the USA. In order for this framework to be applicable to small, open economies, we must add some more considerations to the analysis presented so far. In a small, open economy

there are two important factors which have not yet been discussed so far, which influence the rate of the multiplier.

On the one hand, in an open economy if domestic production needs imported inputs, then part of the incomes primarily and additionally generated by fiscal expenditures will “leak” abroad. If the domestic economy is large enough, then the income generated abroad is capable of increasing foreign GDP significantly, which could increase foreign demand for domestic products, and this could partly compensate for the leaked income. However, if the economy is small, then this compensatory effect does not exist, because in this case the income leaking from the domestic economy is not sufficient to exert a meaningful influence on the foreign economy. Therefore, in the case of a small open economy the rate of the expenditure multiplier decreases owing to the import needs of the economy. For example, if we apply the original argument of Keynes to a small open economy, then the value of the expenditure multiplier with deficit financing will be determined by the following formula: $1/(1-c+m)$, where c means marginal propensity to consume, and m stands for marginal propensity to import. If the demand for imports of the economy, i.e. the value of m is high, then the rate of the multiplier will be significantly lower than in the case of a closed economy.

On the other hand, the magnitude of the multiplier is not independent of the exchange rate regime. The effects of fiscal policy in the case of small open economies were first analysed by *Fleming (1962)* and *Mundell (1963)*. Chapter 2 of *Benczes and Kutasi (2010)* contains a formal presentation and discussion of the Mundell–Fleming model.

According to the Mundell–Fleming model, in the case of a fixed exchange rate regime, fiscal expansion will increase output more efficiently. The reasoning is the following: if fiscal policy raises demand, then interest rates will rise, and the higher interest rate will put the domestic currency under pressure for appreciation. In the case of a fixed exchange rate regime, the central bank will respond to this with intervention, to mitigate (or terminate) the appreciation by easing the monetary stance; therefore, ultimately monetary policy will support the boosting of the economy. By contrast, in the case of a flexible exchange rate regime, the central bank does not respond exclusively to the movement of the exchange rate, and if it does not boost the economy or only boosts it to a smaller extent compared to the fixed exchange rate regime, then the more appreciated exchange rate will mitigate the increase in output via more restricted export performance, and ultimately the impact of the fiscal spending will be smaller. *Farhi and Werning (2016)* analyse the effects of fiscal expansion in a modern dynamic stochastic model framework in open economies, but their fundamental conclusions are in harmony with the results of Fleming and Mundell.

There is another very important element that is missing from the fiscal theory developed for large economies, which is of key importance for small economies: the issue of the sustainability of the deficits related to fiscal expansion, see, e.g. Chapter 3 of *Benczes and Kutasi (2010)* or the study of *D’Erasmus et al. (2016)*. In the case of some large economies – such as that of the USA or Germany – there is practically no constraint to increasing public debt: even in the case of a relatively high public debt-to-GDP ratio the financial markets regard the funding of the public debt as a risk-free investment. The financing practice of small economies is in sharp contrast with that. In these economies, an increase in public debt will raise the risk premium of the debt, and beyond a certain threshold this premium may even rise at an extreme pace. Furthermore, it is possible for the debt to reach a level when the sources of financing dry up completely.

The above factors have a significant impact on the effectiveness of fiscal policy. When the public debt is appropriately high, a fiscal expansion financed from deficit would increase the public debt and the burdens of financing to such an extent that it would no longer stimulate the economy and instead would have a contractionary effect. If the burdens of financing are extremely high, sooner or later the government will have to introduce such fiscal austerity, which will result in a significant growth sacrifice. How quickly the contraction of the economy occurs depends on how rational the expectations of the relevant economic agents are. But regardless of the timeframe, if we do not regard the problem statically, then with a high public debt the increase of fiscal expenditures will ultimately not increase, but rather lower real GDP.⁷

3. Empirical research

3.1. Identification

The empirical identification of expenditure multipliers is fundamentally difficult,⁸ since their value depends on several other factors, including interactive effects. The fiscal multiplier depends on the state of the business cycle, the expectations of economic agents, the reaction function of economic policy, i.e. in general on several structural parameters, the accurate measurement of which poses a challenge to the analyst. The main problem lies in the fact that it is not sufficient merely to divide the changes in GDP over time by the changes in the fiscal instrument that occurred over the same period, using actual data, in order to find out the change that a change of one unit in the fiscal variable causes in GDP. This is particularly true because the level of activity in the economy is affected by several other factors as well. In

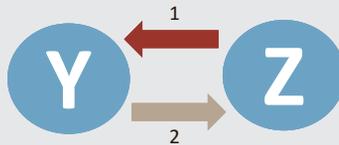
⁷ The study of *Baksa et al. (2014)* discusses in detail in what way expectations affect the multiplier. At the same time, this model does not assess the impacts of fiscal policy in the case of a liquidity trap, and therefore they find low multiplier values.

⁸ This is generally true for other kinds of fiscal multipliers, such as tax multipliers.

the identification, we want to know what the ultimate exclusive effect of the fiscal instrument is on output. A two-way causal relationship may exist between the two variables. On the one hand, changing fiscal expenditures affects GDP, as this is the channel which practically identifies the fiscal multiplier itself (*Figure 6*, arrow no. 1). On the other hand, economic policy may and does also respond⁹ to changes in GDP, for example, exactly by the stimulation or possibly by the restraint of public sector spending (*Figure 6*, arrow no. 2). Therefore, the so-called identification problem lies in whether or not we only capture the first channel. Therefore, in terms of the causal relationships we look for the effect of the one that is exerted by the governmental instrument on economic growth and not vice versa.

As a solution, there are two major and widely used methods in the literature,

Figure 6
A stylised presentation of the problem of identification



Note: Legend: Y means output, Z is the fiscal instrument in question. Channel 1 is the exclusive effect of public sector spending and changes in taxes on the output, while channel 2 is the response of the government to changes in GDP.

which we discuss later on in more detail. One approach identifies the fiscal shocks by a structural vector-autoregressive model (SVAR), and based on that, it becomes possible to determine the multiplier values. This method was first applied by *Blanchard and Perotti (2002)* for estimating the fiscal multipliers. In this case, the condition of identification is usually the requirement that the change in GDP should not affect public sector spending in the given period. This assumption is not only a simplifying methodological condition necessary for the SVAR simulations, but also an approach defined based on economic considerations. The GDP of the given period is not known to start with, and naturally, its exact value will be known only later on, at least one quarter later; on the other hand, owing to institutional features, the steps of economic policy can also not necessarily be immediately ascertained.¹⁰

⁹ Either by direct steps, or through the automatic fiscal stabilisers. We explain this latter in more detail in the part titled *4. Implications on economic policy*.

¹⁰ For example, major changes in taxation or expenditures are preceded by social, professional, parliamentary debates concerning the effects of the fiscal measure, which could, by necessity, generate some subsequent actual response.

The other approach builds on the search for a fiscal variable that is presumably independent of the fluctuation in GDP. The first cases in point concerning such a variable are those of *Barro (1981)*, and then *Ramey and Shapiro (1998)*, in which the authors identified fiscal shocks by relying on figures on military expenditures. They simply used information taken from reports on planned military expenditures as a shock variable. Changes in military expenditures are independent of domestic growth and mainly depend on external factors (a period of wars), therefore from the aspect of identification using that figure is a good choice.

Among the empirical methods used to assess the impacts of other, fiscal shocks, we should also mention the one in which the authors use full econometric models. However, the development of these models is not exclusively aimed at estimating the fiscal impulses and owing to their econometric structures they are suitable for creating simulations, enabling the exploitation of these opportunities.¹¹ In addition, another approach is when the authors use a New Keynesian macroeconomic model, the major parameters of which are estimated and then the fiscal multiplier can be calculated using the estimated parameters, based on simulation.¹² The empirical strength of this latter practice is fundamentally weakened – in addition to the usual estimation properties of the Bayesian estimation, characteristic of complex models – because the applied models result in strong theoretical restrictions in themselves, and therefore they provide much less room for information deriving from the observed data.

The methods are presented in detail, and several methodological, literature references are given in: *Chinn (2013)*, and *Spilimbergo et al. (2009)*. In the following, we will mainly focus on the first two methods and on investigations where the authors specifically analysed the factors affecting the value of the fiscal multipliers.

3.2. Results – an overview of the literature

The study of *Ilzetzki et al. (2011)* contains a detailed analysis of the multiplicative effects of public sector spending, in which over forty countries were assessed in the period between 1960 and 2007. They were merely seeking answers as to what is the impact of the most important economic properties on the effect of the fiscal stimulus. Of the factors determining the rate of the multiplier, the authors assessed the following: level of development (developed or emerging), the exchange rate regime (fixed or flexible), the openness of the economy and indebtedness.

They prepared their estimate using a panel SVAR model, with the structural condition that the GDP of the given quarter has no impact on governmental consumption, only one quarter later.¹³ They assessed both short-term and long-

¹¹ See e.g.: *Horváth et al. (2006)*

¹² See e.g.: *Zubairy (2010)*

¹³ The SVAR model included the following endogenous variables: governmental instrument (e.g. government consumption), base interest rate of the central bank, GDP, current account, real exchange rate. In their estimates, they used four lags uniformly. The various lags did not significantly affect their conclusions.

term effects, and furthermore, whether the difference between the multipliers of public-sector consumption and public investments are significant. We summarise their results in the following table.

	Short-term multiplier		Long-term multiplier	
	Government consumption	Government investment	Government consumption	Government investment
Developed economies	0.37	0.41	0.8	1.15
Developing economies	-0.21	0.57	0.18	0.75
Fixed exchange rate	0.09	0.36	1.5	1.42
Flexible exchange rate	-0.28	0.46	-0.41	0.16
Closed economy	0.02	0.46	1.29	0.7
Open economy	-0.28	0.51	-0.75	-0.23

Source: Ilzetzki et al. (2011)

In general, they conclude that the multiplier of government investments is higher than that of government consumption, but at the same time, in several cases there is no statistically significant difference between the two. In relation to the level of economic development, the multiplier is much higher for developed economies. In emerging economies, the long-term, demand-stimulating effect of governmental consumption is, in fact, not significantly different from zero, while that of government investments is a positive figure and is substantially higher. *Estevão – Samake (2013)* and *Kraay (2012)* estimate similarly lower multiplier values for developing economies.

As for the assessment of various exchange rate regimes, they found that the value of the fiscal multiplier is higher in the case of a fixed exchange rate regime compared to a flexible exchange rate regime. *Born et al. arrived (2013)* at a similar conclusion: in the case of a fixed exchange rate regime they measured a multiplier effect that is 2 or 3 times higher compared to a flexible exchange rate regime. Underlying this conclusion is the intuition that – in the case of a fiscal shock – a flexible exchange rate regime enables faster implementation of mechanisms that initiate economic adjustment.

In assessing the openness of economies, the authors usually found a higher fiscal multiplier for closed economies, in accordance with the results of *Flemming (1962)* and *Mundell (1963)* cited above.

The study of *Ilzetzki et al. (2011)* uses a sample that lasted until 2007, i.e. the period preceding the crisis, and primarily supports the sceptical approach to fiscal policy. Studies that also considered the post-crisis period in the estimations arrived at significantly different results.

Of these, it was the study of *Blanchard and Leigh (2013)*, two researchers at the IMF, that attracted the most attention. They came to the conclusion that the value of the fiscal multiplier is higher than 1 after the crisis, which was much higher than assumed previously. This study stirred up quite a storm when it was published, since up to that point the IMF had taken a rather sceptical approach. In making their estimates, they used the following considerations. They estimated the following equation:

$$\text{forecast error of growth} = \alpha + \beta \cdot \text{forecast of fiscal consolidation} + \epsilon$$

In the equation, they regressed the forecast error of GDP growth from one year to the next by the forecast of the change of fiscal consolidation. The structural balance of the government is contained on the right side. With rational expectations and assuming that an appropriate model was used for the forecasts, the value of β should be zero. The investigation was conducted for the 2010–2011 period, on a sample of 26 countries.¹⁴ They arrived at a significantly negative value as the value of the estimated β . A large number of robustness tests were carried out, with different specifications, and although different results were obtained in several cases, in total, significantly negative values were estimated for this parameter. For example, certain countries were left out from the estimate, the outliers were treated by different statistical procedures, the above equation was extended by various additional explanatory variables (e.g. CDS, fiscal balance, initial debt, debts of households, current account balance, actual fiscal measures instead of planned fiscal steps); in addition, the sample was expanded for the period between 2009 and 2012. Furthermore, the estimate was also performed for the pre-crisis period (1997–2008), and a significantly zero β value was found.

The results were robust for the various specifications and revealed that the current fiscal multiplier values are higher than believed previously. Although the definition of the exact multiplier values that were used for earlier forecasts is unclear, since it fundamentally depends on the structure of fiscal consolidation and on economic conditions, it was a general opinion based on the results of several studies that in respect of developed economies the value of the multiplier is about 0.5 for two years. For example, in IMF (2008) a multiplier value of 0.5 was estimated for 21 developed economies for the period between 1970 and 2007; furthermore, the IMF (2010) also estimated an average value of about 0.5 on average for 15 developed economies for the period between 1979 and 2009. In light of the above, the authors conclude that based on the results of their estimates currently the value of the multiplier is above 1. At the same time, they note that the results should be treated with caution, particularly because fundamentally there is not one single multiplier value for every economy, for every fiscal shock. The value of the multiplier may vary

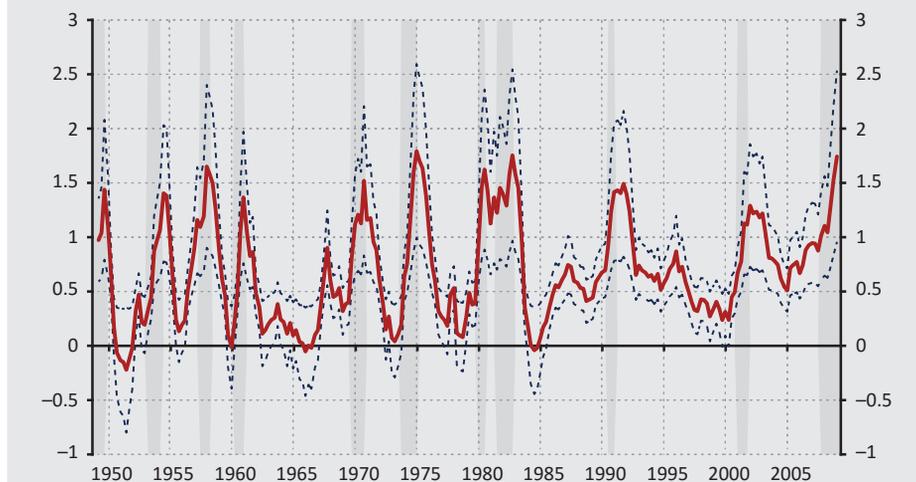
¹⁴ Basically, it was the EU27, however, there were no data on the structural fiscal balance for Estonia, Latvia, Lithuania and Luxembourg in the WEO of April 2010. On the other hand, they added three developed European economies – Iceland, Norway, Switzerland – to the sample and thus carried out the test on a total of 26 countries (27–4+3).

by country and by period of time, and furthermore, as economies recover from a crisis, their multiplier values will presumably return to pre-crisis levels.

Applying the methods outlined above, *Mohlmann and Suyker (2015)* also conducted the estimate for later periods. On the one hand, they refreshed it: they used a version of the database of *Blanchard and Leigh (2013)*, that was updated 3 years later. They also found a similarly significant negative β value for the periods between 2009 and 2010 and 2010 and 2011, but the coefficient is no longer significant for 2011–2012. They estimated a value close to zero for later years.¹⁵

Based on the empirical results reviewed so far, it can be understood that the value of the expenditure multiplier is not constant and that it depends on the state of the business cycle. This is clearly underpinned by the studies of *Auerbach and Gorodnichenko (2012, 2013)*. The authors estimate fiscal multiplier values of about 1.0–1.5 by estimating a regime-changing model for a recession environment, while for a boom the estimate is around 0.0–0.5. Therefore, according to their main conclusion, fiscal expansion is much more efficient in a period of recession compared to a period of expansion.¹⁶ The results are summarised in *Figure 7*.

Figure 7
Development of the expenditure multiplier in the United States



Note: Shaded regions are recessions defined by the NBER. Blue dashed lines are 90 per cent confidence interval.

Source: *Auerbach – Gorodnichenko (2012)*

¹⁵ The authors note that, naturally, this result could also reflect that the forecasters of the IMF have “learned their lessons” from the crisis, and for the later periods indicated they already used higher multipliers compared to the pre-crisis period.

¹⁶ They also assessed to the various fiscal expenditures and typically arrived at higher values for military expenditures and government investments, in contrast to the less powerful demand-stimulating effect of government consumption.

The above results are confirmed by the study of *Huidrom et al. (2016)*, demonstrating that in a period of recession the stimulating effect of fiscal expenditure is significantly higher. While in a period of recession multipliers of around 1.5–2 are obtained for a time horizon of 1–2 years, the values estimated for periods of expansion are statistically typically not different from zero.

Nonetheless, the value of the fiscal multiplier is not only affected meaningfully by the cyclical state of the real economy, the level of indebtedness of the given country may also affect it. According to the study of *Ilzetki et al. (2011)*, the level of indebtedness of the given economy significantly affects the efficiency of fiscal spending. Excessive indebtedness substantially mitigates the growth effect of fiscal expansion. In proportion to GDP, in the case of economies with a level of indebtedness exceeding 60 per cent a negative fiscal multiplier is arrived at, which is a strong indication and also a serious challenge concerning the recovery from the crisis of economies with an excessive level of indebtedness. In their assessment of robustness, the authors found that indebtedness exceeding broadly the threshold of 60 per cent will result in an insignificant multiplier. *Huidrom et al. (2016)* obtained similar results. The authors assessed 34 countries (of which 19 were developed and 15 emerging) for the period between 1980 Q1–2014 Q1, with the application of the so-called IPVAR.¹⁷ model. In respect of the immediate effect, fiscal multiplier becomes insignificant starting from an indebtedness level of roughly 65–70 per cent, and in the case of a low level of indebtedness the long-term value is close to 1.

4. Implications for economic policy

In this section, we provide an overview of what conclusions can be drawn based on the theory and empirical research on the operation of fiscal policy. Our purpose is not to describe an optimal rule of fiscal policy in the strict sense, but rather to consider those basic principles that are characteristics of a proper fiscal policy in our opinion.

As demonstrated in the previous sections, both the theoretical models and the empirical studies confirm that the magnitude of the fiscal expenditure multiplier is state-dependent: during booms it is relatively low, while in recessions, and especially after major crises, it is relatively high, higher than 1.

From that it follows that the proper fiscal policy is also state-dependent. It is not useful or possible to apply fiscal stimulus to the economy over the whole business cycle; at the same time, the claim that strict fiscal policy is required always and under any circumstances is not true either.

One frequent argument against fiscal policy is that it is much more difficult to implement than monetary policy; therefore, it is not recommended to be used

¹⁷ Interacted Panel Vector Autoregressive model.

for smoothing the business cycles.¹⁸ However, after the crisis of 2007–2008 it was proven that in a deep recession monetary policy can quickly reach its limits and become ineffective. In such a case, the only tool to stimulate the economy is fiscal expansion, without the application of which the economy could easily move into a state of persistent stagnation, as is demonstrated by the paper of *Eggertson and Mehrotra (2014)*. In addition, the fiscal multiplier is the highest in such a situation, and therefore fiscal policy is not only desirable, but also efficiently applicable.

At the same time, from the aspect of the sustainability of the deficit the professional literature has been traditionally sceptical about fiscal stimulus. It is believed that even with a high value of the expenditure multiplier it is impossible to implement a meaningful stimulus of the economy, since it will relatively quickly lead to an unsustainable deficit; calculations on this are included, for example, in the study of *Chand (1993)*.

By contrast, the post-crisis research supports the claim that fiscal expansion implemented in a crisis does not necessarily lead to an unmanageable deficit. *Delong and Summers (2012)* demonstrate that if the recession is so deep that it has long-term negative impacts on growth – i.e. in the case of hysteresis – with realistic parameters fiscal expansion can even be self-funding, i.e. the additional tax revenues arising from the long-run growth could cover the deficit used to finance the expansion. The empirical study of *Auerbach and Gorodichenko (2017)* is in accordance with the above results: according to their research on developed economies, projects of fiscal stimulus carried out at the time of crisis did not result in a permanent increase in debt as a percentage of GDP or that of the risk indicators.

On the whole, the above claims do not mean that in small countries that are indebted crisis management based on fiscal stimulus is necessarily a viable option, but the most recent crisis serves to demonstrate exactly that fiscal restriction also does not necessarily work in such cases. In fact, if the value of the multiplier is high, then the decrease in fiscal expenditures by one unit will probably reduce GDP by more than one unit. As a result, debt as a percentage of GDP will not decrease, but will instead rise. Therefore, the result of restriction will be the exact opposite of the original aim. A similar story was observed in Greece after the crisis.

Therefore, if an economy enters into a deep crisis and a liquidity trap in which it has no fiscal space for manoeuvre, then it has fallen into a trap from which it will not be able to escape on its own. Fiscal easing is not feasible, but tightening will also not resolve the situation. *Blanchard et al. (2016)* assessed this problem with a sophisticated DSGE model in the case of the central and peripheral countries of the euro area. According to the calculations of the authors, in such a case the fiscal

¹⁸ Naturally, here we mean such a degree of volatility in business cycles that cannot be sufficiently managed by automatic fiscal stabilisers.

expansion of the central countries may help the peripheral countries. The reason is that the increase in fiscal expenditures of the centre will increase the real GDP of the periphery countries by more than one unit.

Since the lack of space for fiscal manoeuvre in a deep recession and liquidity trap makes recovery from the crisis extremely difficult, it is very important to build up fiscal space during booms. This means that in booms strict fiscal policy is recommended. On the one hand, during such times there is no need for fiscal stimulus, stimulating an overheated economy contradicts the principle of smoothing the cycles. On the other hand, during such times the fiscal multiplier is low, and therefore increasing expenditures would not significantly stimulate the economy, but reducing them would also not cause a substantial loss in output. But the most important aspect is that during periods of booms a tight fiscal policy could serve to build up the fiscal space that is very much needed should a crisis occur.

Based on the arguments presented so far, it is safe to say that counter-cyclical fiscal policy is recommended to be pursued. However, the application of counter-cyclical policy requires great caution. It is important to estimate the state of the business cycle as robustly as possible and to carefully estimate the current value of the expenditure multiplier. The application of automatic fiscal stabilisers could significantly support the implementation of counter-cyclical fiscal policy. We call it an automatic fiscal stabiliser when the budget balance changes automatically whenever the business cycle enters a new stage. For example, in the case of a recession expenditures grow on a stable course, while revenues will fall short owing to the shrinking tax bases. An advantage of automatic stabilisers is that they provide an immediate, symmetric fiscal response to the change in the business cycle, whereas they do not influence budget sustainability over the cycle.¹⁹ The application of automatic fiscal stabilisers does not exclude the application of discretionary steps, for example, in a deep and persistent recession individual economy-stimulating fiscal programmes may also be necessary.

It should be highlighted that the opportunity of applying a counter-cyclical fiscal policy is independent of the steady state size of the public spending in a given economy.²⁰ There are various feasible regimes in the case of developed economies, the average ratio of government spending as a percentage of GDP is relatively low in the United States, for example, while in Sweden it is relatively high. Theoretically, counter-cyclical fiscal policy is feasible in both cases; from this aspect it is the cyclical position of fiscal spending that is important, rather than its average weight. The

¹⁹ As is detailed in the study of *Buti and Gaspar (2015)*, originally the automatic fiscal stabilisers were aimed at redistribution, their business cycle smoothing effect is only a beneficial by-product. Precisely because of that, the matter of automatic fiscal stabilisers should be reconsidered and their business cycle smoothing capability should be improved.

²⁰ Although this has an effect on the size of the automatic fiscal stabilisers, it is independent of the extent of the discretionary decisions that can be made.

average size of fiscal spending does not necessarily affect the fiscal space. If the weight of public expenditure is high, but together with that, the tax burdens are also proportionately high, then public debt remains manageable in terms of size and will not restrict the fiscal space.

Finally, we would like to highlight that the principle of counter-cyclical fiscal policy in itself is not a new result, it was already known before the crisis. The novelty is that since the crisis both the theory and empirical research has confirmed that the size of the fiscal multiplier depends on the state of the business cycle, which provides another strong argument for counter-cyclical fiscal policy.

5. Conclusions

In the study we explored how efficient fiscal policy is as a tool for smoothing business cycles. According to the current economic theory and the most recent empirical research, both the efficiency of fiscal policy and the magnitude of the expenditure multiplier that measures it are state-dependent. During recessions, especially prolonged and deep ones that follow financial crises, when monetary policy becomes ineffective, the magnitude of the expenditure multiplier is high. By contrast, during times of booms the value of the expenditure multiplier is low. In addition to the state of the business cycle, the structure of the economy also affects the magnitude of the expenditure multiplier: in the case of small open economies the value of the multiplier is lower than in similar, but closed economies.

Since the rate of the expenditure multiplier is state-dependent, it is recommended to pursue fiscal policy in a counter-cyclical manner. During times of recessions, when monetary policy is ineffective, fiscal expansion can be used to stimulate the economy in such a manner that no consumption or private investments will be crowded out. By contrast, during times of booms strict fiscal policy should be applied in order to build up the fiscal space that will enable the required fiscal expansion during times of crises.

References

- Auerbach, A.J. – Gorodnichenko, Y. (2012): *Measuring the Output Responses to Fiscal Policy*. American Economic Journal: Economic Policy, 4(2): 1–27. <https://doi.org/10.1257/pol.4.2.1>
- Auerbach, A.J. – Gorodnichenko, Y. (2013): *Fiscal Multipliers in Recessions and Expansions*. In: Alesina, A. – Gavani, F. (ed.): *Fiscal policy after the financial crisis*, University of Chicago Press.
- Auerbach, A.J., – Gorodnichenko, Y. (2017): *Fiscal Stimulus and Fiscal Sustainability*, Kansas City Fed.

- Baksa, D. – Benk, Sz. – Jakab, Z. (2014): *Léteznek "a" költségvetési multiplikátorok? Fiskális és monetáris reakciók, hitelesség és költségvetési multiplikátorok Magyarországon (Does "the" fiscal multiplier exist? Fiscal and monetary reactions, credibility and fiscal multipliers in Hungary)*. *Sigma*, 45(1–2): 57–83.
- Barro, R.J. (1974): *Are government bonds net wealth?* *Journal of Political Economy*, 82(6): 1095–1117. <https://doi.org/10.1086/260266>
- Barro, R.J. (1981): *Output effects of government purchases*, *Journal of Political Economy*, 89: 1086–1121. <https://doi.org/10.1086/261024>
- Batini, N. – Eyraud, L. – Forni, L. – Weber, A. (2014): *Fiscal Multipliers: Size, Determinants, and Use in Macroeconomic Projections*. Fiscal Affairs Department, IMF, Technical Notes and Manuals, September 2014.
- Benczes, I. – Kutasi, G. (2010): *Költségvetési pénzügyek (Fiscal Finances)*. Akadémiai Kiadó.
- Blanchard, O. – Perotti, R. (2002): *An Empirical Characterization of The Dynamic Effects of Changes in Government Spending and Taxes on Output*. *The Quarterly Journal of Economics*, November.
- Blanchard, O. – Leigh, D. (2013): *Are We Underestimating Short-term Fiscal Multipliers?* *American Economic Review*.
- Blanchard, O. – Cerutti, E. – Summers, L. (2014): *Inflation and Activity – Two Explorations, and Their Monetary Policy Implications*. ECB Forum on Central Banking 2014.
- Blanchard, O. – Erceg, C.J. – Linde, J. (2016): *Jump-Starting the Euro Area Recovery: Would a Rise in Core Fiscal Spending Help the Periphery?* Manuscript.
- Born, B. – Jüßen, F. – Müller, G.J. (2013): *Exchange rate regimes and fiscal multipliers*. *Journal of Economic Dynamics & Control*, 37: 446–465. <https://doi.org/10.1016/j.jedc.2012.09.014>
- Buti, M. – Gaspar, V. (2015): *Designing Fiscal Policy for Steady Enduring Growth*, <http://voxeu.org/article/fiscal-policy-enduring-growth>
- Calvo, G. (1983): *Staggered Prices in a Utility Maximizing Framework*. *Journal of Monetary Economics*, 12(3): 383–398. [https://doi.org/10.1016/0304-3932\(83\)90060-0](https://doi.org/10.1016/0304-3932(83)90060-0)
- Chand, S.K. (1993): *Fiscal Impulse Measures and Their Fiscal Impact*. In: Bléjer, M.I. – Cheasty, A. (ed.): *How to Measure Fiscal Deficit*, International Monetary Fund.
- Chinn, M. (2013): *Fiscal Multipliers*. *The New Palgrave, Dictionary of Economics*. https://doi.org/10.1057/978-1-349-95121-5_2859-1
- Christiano, L. – Eichenbaum, M. – Rebelo, S. (2011): *When Is the Government Spending Multiplier Large?* *Journal of Political Economy*, 119(1): 78–121. <https://doi.org/10.1086/659312>

- Delong, J.B. – Summers, L.H. (2012): *Fiscal Policy in a Depressed Economy*. *Brooking Papers on Economic Activity*, Spring: 2332–2374. <https://doi.org/10.1353/eca.2012.0000>
- D’Erasmus, P. – Mendoza, E.G. – Zhang, J. (2016): *What is a Sustainable Public Debt?* In: Taylor, J. B. – Uhlig, H. (ed.): *Handbook of Macroeconomics*, Volume 2, Chapter 32: 2417–2492.
- Eggertsson, G.B. – Krugman, P. (2012): *Debt, Deleveraging and the Liquidity Trap: a Fisher-Minsky-Koo Approach*. *Quarterly Journal of Economics*, 127(3): 1469–1513. <https://doi.org/10.1093/qje/qjs023>
- Eggertsson, G.B. – Mehrota, N.R. (2014): *A Model of Secular Stagnation*, NBER WP, 20574.
- Erceg, C. – Linde, J. (2012): *Is There a Fiscal Free Lunch in a Liquidity Trap?* *International Finance Discussion Papers*, 1003, Board of Governors of the Federal Reserve System.
- Estevão, M. – Samake, I. (2013): *The Economic Effects of Fiscal Consolidation with Debt Feedback*. *IMF Working Paper 13/136*. <https://doi.org/10.5089/9781484304488.001>
- Farhi, E. – Werning, I. (2016): *Fiscal Multipliers: Liquidity Traps and Currency Unions*. In: Taylor, J. B. – Uhlig, H. (ed.): *Handbook of Macroeconomics*, Volume 2, Chapter 31: 2417–2492.
- Friedman, M. (1957): *A Theory of the Consumption Function*. Princeton University Press.
- Fleming, M.J. (1962): *Domestic financial policies under fixed and floating exchange rates*. *IMF Staff Papers*. 9: 369–379.
- Galí, J. (2008): *Monetary Policy, Inflation and the Business Cycle*, Princeton University Press.
- Horváth, Á. – Jakab, M.Z. – P. Kiss, G. – Párkányi, B. (2006): *Myths and Maths: Macroeconomic Effects of Fiscal Adjustments in Hungary*. *Magyar Nemzeti Bank Occasional Papers*, OP 52.
- Huidnrom, R. – Kose, M.A. – Lim, J.J – Ohnsorge, F.L. (2016): *Under The Cloud Of Cycles: Linkages Between Fiscal Multipliers And Positions*. *Policy Research Working Paper*, Wps7724, World Bank Group.
- Ilzetzki, E. – Mendoza, E.G. – Végh, C.A. (2011): *How Big (Small?) are Fiscal Multipliers?* *IMF Working Paper*.
- Keynes, J.M. (1936): *The General Theory of Employment, Interest, and Money*, New York: Macmillan.
- Kraay, A. (2012): *How Large is the Government Spending Multiplier? Evidence from World Bank Lending*, The World Bank.
- Lucas, R.E. (1979): *Some International Evidence on Output-Inflation Trade-offs*. *American Economic Review*, 63(3): 326–334.

- Modigliani, F. (1966): *The Life Cycle Hypothesis of Saving, the Demand for Wealth and the Supply of Capital*. Social Research, 33(2): 160–217.
- Mohlmann, J. – Suyker, W. (2015): *An update of Blanchard’s and Leigh’s estimates in ‘Growth Forecast Errors and Fiscal Multipliers’*. VOX, CEPR’s Policy Portal, 01 December 2015.
- Mundell, R.A. (1963): *Capital mobility and stabilization policy under fixed and flexible exchange rates*. Canadian Journal of Economic and Political Science, 29(4): 475–485. <https://doi.org/10.2307/139336>
- Owyang, M.T. – Ramey, V.A. – Zubairy, S. (2013): *Are Government Spending Multipliers Greater During Periods of Slack? Evidence from 20th Century Historical Data*. Federal Reserve Bank of St. Louis, Research Division, Working Paper 2013-004A.
- Ramey, V.A. – Shapiro, M.D. (1998): *Costly capital reallocation and the effects of government spending*. Carnegie-Rochester Conference Series on Public Policy, 48(1): 145–194. [https://doi.org/10.1016/S0167-2231\(98\)00020-7](https://doi.org/10.1016/S0167-2231(98)00020-7)
- Ramey, V.A. (2011): *Identifying Government Spending Shocks: It’s All in the Timing*. Quarterly Journal of Economics, 126 (1): 1–50. <https://doi.org/10.1093/qje/qjq008>
- Riera-Crichton, D. – Vegh, C.A. – Vuletin, G. (2016): *Tax multipliers: Pitfalls in measurement and identification*. Journal of Monetary Economics, 79: 30–48. <https://doi.org/10.1016/j.jmoneco.2016.03.003>
- Sargent, T. – Wallace, N. (1975): *Rational Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule*. Journal of Political Economy, 83(2): 241–254. <https://doi.org/10.1086/260321>
- Spilimbergo, A. – Symansky, S. – Schindler, M. (2011): *Fiscal Multipliers*. IMF Staff Position Note, SPN/09/11.
- Szentmihályi, Sz. – Világi, B. (2015): *The Phillips curve – history of thought and empirical evidence*. Financial and Economic Review, 14(4): 5–28. <http://english.hitelintezetiszemle.hu/letoltes/1-szentmihalyi-vilagi-en.pdf>
- Woodford, M. (2011): *Simple Analytics of the Government Expenditure Multiplier*. American Economic Journal: Macroeconomics, 3: 1–35. <https://doi.org/10.1257/mac.3.1.1>
- Zubairy, S. (2010): *On Fiscal Multipliers: Estimates from a Medium Scale DSGE Model*. Bank of Canada Working Paper 2010-30, November 2010.

Trade Credit or Bank Credit? – Lessons Learned from Hungarian Firms between 2010 and 2015

Dániel Havran – Péter Kerényi – Attila Víg

This paper addresses the way in which trade credit was used by Hungarian firms in the period between 2010 and 2015. Relying on Burkart and Ellingsen's (2004) theory of trade credit, we use panel data on 14,554 Hungarian firms (including 68 large corporations) to estimate the relationship of trade credit and short-term bank credit. Estimated on sub-samples broken down by profitability, our results only confirm a complementary relationship. We also examine the relationship separately for each category of firm size. We found a complementary relationship for small and microenterprises, whereas the results obtained for large corporations imply a substitution effect. In Hungary, in the period after 2013 accounts payable tended to be increased by financially constrained micro and medium-sized enterprises and mostly held steady by financially unconstrained firms.

Journal of Economic Literature (JEL) codes: G32, C23

Key words: trade credit, financial constraints

1. Introduction

Firms consider trade credit to be an important channel of external funding, particularly in cases when they have no other access to finance. Trade credit is often a substitute for short-term bank credit, while it is also possible that access to bank credit is facilitated by the willingness of suppliers that are more familiar with the firm to finance it. Accordingly, in various stages of the business cycle, certain firms rely on trade credit more heavily, and others less so. In recent years, Hungarian

Dániel Havran is an associate professor at the Corvinus University of Budapest.

E-mail: daniel.havran@uni-corvinus.hu

Péter Kerényi is a PhD student at the Doctoral School of Business Informatics at the Corvinus University of Budapest, and is a scholar on the Pallas Athene Domus Educationis Foundation.

E-mail: peter.kerenyi@uni-corvinus.hu

Attila Víg is a PhD student at the Doctoral School of Quantitative Economics at the Corvinus University of Budapest, and is a scholar on the Pallas Athene Domus Educationis Foundation.

E-mail: attila.vig@uni-corvinus.hu

The authors wish to thank Bisnode Magyarország Kft. for the data provided.

The Hungarian manuscript was received on 9 August 2017.

DOI: <http://doi.org/10.25201/FER.16.4.86121>

firms experienced both a period of recession in the financing environment and the subsequent recovery. This paper is concerned with how accounts payable were used for financing purposes in the period between 2010 and 2015 by Hungarian non-financial and non-utility firms which draw up annual accounts.

For the same period, a number of studies have been produced on the trends in the funding of Hungarian firms, mostly addressing the practice of bank financing or the effect of stimulating credit supply (Funding for Growth Scheme), e.g. *Csubák – Fejes 2015; Bálint – Fellner 2016, 2017; Bokor – Hidasi 2014; Endrész – Harasztosi – Lieli 2015*. In the literature, few analyses have been carried out on the use of trade credit in respect of the past years (*Szűcs 2008; Nábelek 2016*). However, understanding firms' short-term financing practices could vary and complement the overall picture about long-term financing. With our paper, it is on that issue that we wish to contribute to the Hungarian discourse on corporate finance.

Our analysis has both cross-sectional and temporal aspects.

In taking a cross-sectional approach, we seek to explore whether in the period concerned, trade credit was used by firms to complement or substitute bank credit. We rely on *Burkart and Ellingsen's (2004)* model to establish the theoretical framework required to address this issue. To segment firms, we use indices representing firms' financial constraints – *Kaplan–Zingales*¹, *Whited–Wu*², *Cleary*³ – as well as corporate profitability (low, medium and high EBIT/ASSET), in contrast to the empirical work of *Cunningham (2005)*, who used firms' "wealth" (i.e. profit, a proxy for the size of internal or own funds) to measure the presence of financial constraints. The propositions of theory are only supported by our estimates unambiguously for small and microenterprises.

We are also interested in the extent to which the result changes according to the category of firm size. Using Hungarian data, our empirical results mostly confirm that for large and medium-sized enterprises, there is no or only weak substitution between trade credit and bank credit, whereas the relationship is complementary for small and microenterprises. We find that the presence of accounts receivable also leads to increased use of trade credit, and that the effect is the strongest in the case of financially constrained firms.

Owing to the difference in the relationship between these channels of finance, firms respond differently to cyclical changes. In the period under review, the pattern emerging in the use of trade credit may be pro-cyclical for some firms (with the volume of accounts payable falling in crisis and growing in recovery), and counter-

¹ *Kaplan – Zingales (1997)*

² *Whited – Wu (2006)*

³ *Cleary (1999)*

cyclical for others (with days outstanding becoming longer in crisis and shorter in recovery). Could it be argued that firms tend to accumulate payables during times of crisis? Or, conversely, do payables increase when firms use them to finance their growth during times of recovery? Industry character notwithstanding, what are the characteristics of firms in these two groups? What was the level of volatility in each case? Although less reliably in statistical terms, the empirical analysis of cyclicity has found the accumulation of payables in the recovery period to be a feature of financially constrained firms, whereas with financially unconstrained firms days payable outstanding did not increase in the period of recovery. These tendencies could mostly be shown for small and microenterprises.

The rest of this paper is structured as follows. First, an overview is offered on the general financing situation of Hungarian firms in the period between 2010 and 2015, and then a theoretical framework is established as a basis for the empirical analysis. This is followed by the presentation of data and the approach used. The *Results* section presents tables for the estimates obtained, with the corresponding interpretations. A brief summary is provided at the end of the paper.

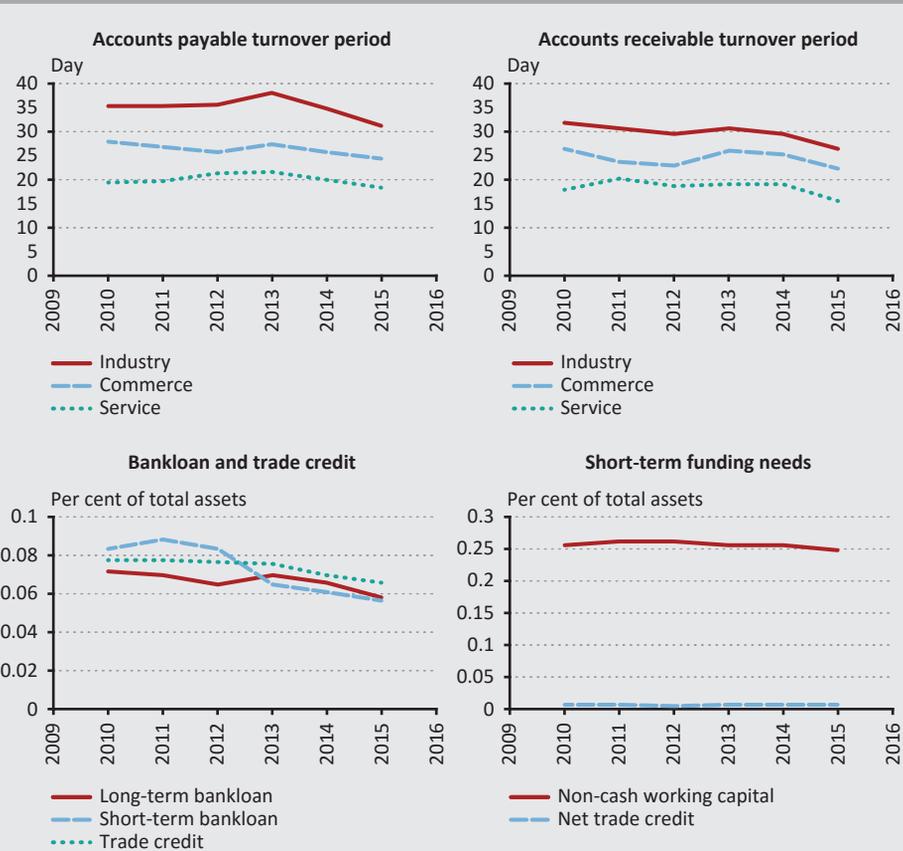
2. Payables as reflected in corporate finance: 2010–2015

For Hungarian firms, recovery from the 2008 crisis was protracted, with the turning point arriving only in 2013/2014 after another dip in 2012. A period characterised by financial constraints and pessimistic expectations was followed by one that offered improved access to credit and a better outlook. In the following, a brief summary is provided of that period in terms of firms' financing practices. Our primary focus will be the MSME segment.

Figure 1 represents some of the trends emerging from the database we processed for the period under review. (The Bisnode database, which contains the financial accounts collected by the tax authority for all firms in scope, will be explained in detail as part of the empirical analysis.) The diagrams show the medians derived from individual ratios.

Figure 1
Main trends in corporate finance, 2010–2015

(median values)



Note: The indicators represented in each diagram were derived as follows: days payable outstanding were derived as the ratio of accounts payable on year-end to material costs, multiplied by 365, and days receivable outstanding were derived as the ratio of accounts receivable on year-end to annual sales, multiplied by 365. Short-term and long-term bank credit outstanding and accounts payable as at the end of the year were related to total assets. Net trade credit refers to accounts receivable less accounts payable, relative to total assets; and non-cash working capital refers to non-cash current assets less accounts payable, relative to total assets.

Source: Calculations based on data from Bisnode Kft.

The settlement of accounts payable took the longest in 2013 in commerce, industry and services alike. After 2013, both accounts payable and receivable days outstanding became shorter. In firms with bank loans outstanding in the period, there was a steady decline in the ratio of short-term loans to total assets. The share of long-term bank loans within the balance sheet also declined, although at a much slower rate. Relative to total assets, net non-cash working capital remained level at around 25–26 per cent, while the median of net trade credit in the sample

stagnated at around 0 per cent. This also implies that accounts receivable frequently correlates with accounts payable, and that out of the components of net working capital, the highest financing requirement is attributable to inventories.

According to *Csubák and Fejes (2014)* and *MNB (2015)*, the volume of lending to MSMEs dropped from its 2008 peak of HUF 3,896.9 billion by almost one third to HUF 2,753.3 billion in 2011, followed by an increase of some HUF 750 billion in 2012. The SAFE reports produced in the same period (*SAFE 2011, 2013, 2014, 2015*) show that for small and medium-sized enterprises, access to external finance was the most difficult in 2011, following which firms found their situations to be steadily improving. At the same time, finding skilled staff became increasingly difficult, while a steadily narrowing group of respondents identified market competition as the key challenge. A summary of the responses is provided in *Table 1*. Questionnaire responses also show that the period between 2013 and 2015 saw a decrease in the ratio of respondents that had their request for financing rejected. As the questionnaires for European countries were administered to relatively small samples in Hungary, the ratios obtained for the attitudes identified must be treated with caution.

Table 1					
Main challenges to Hungarian SMEs					
<i>(SAFE questionnaire surveys)</i>					
Which of the following problems is the most challenging for your firm?	2009	2011	2013	2014	2015
Finding customers	43.50%	24.80%	18.90%	16.22%	26.14%
Competition in the market	1.70%	23.50%	17.10%	14.51%	9.57%
Access to finance	18.70%	22.00%	17.30%	14.06%	11.29%
Costs of labour or production	3.10%	6.10%	11.90%	10.73%	12.37%
Availability of skilled staff	4.20%	6.60%	10.90%	18.36%	21.93%
Regulatory environment	12.40%	10.90%	11.20%	14.68%	9.33%
Other	16.40%	6.10%	12.70%	11.44%	9.37%

Source: Data from the SAFE (2011, 2013, 2014, 2015) questionnaire surveys

The MNB's reports following up on developments in lending (*Bálint – Fellner 2016, 2017*) also confirm the tendency perceived by firms, i.e. that the terms of bank lending to firms significantly improved from 2013 onwards. The study argues that the differential between the lending rate and the cost of funding also dropped, collateral requirements and the required level of creditworthiness were eased, and the available credit lines increased. From June 2013 onwards, subsidised loans (primarily the Funding for Growth Schemes, "FGS") played a major role in facilitating access to finance. While the main focus of FGS 1 was loan replacement, FGS 2, launched in autumn 2013, already brought about improvements in the terms of

access to working capital and investment loans as well (FGS 3 in 2016 already falls outside of the period under review). Analyses are provided in *Módos, Bokor and Hidasi (2014)* and *Bokor, Fellner and Plajner (2014)* on the developments of the period, and in *Endrész, Harasztosi and Lieli (2015)* on the effect of the FGS on investment.

While financing opportunities clearly improved in 2013–2015, developments in the demand for finance appear more gloomy. Demand for finance was predominantly determined by selling opportunities and expectations of growth. The SME Survey of January 2015 by the Institute for Economic and Enterprise Research (“GVI”) (*Nyíró and Hajdu 2015*) shows that while cyclical indicators consistently identify summer 2009 and summer 2012 as the two troughs, expectations improved considerably from 2013. *Juhász and Reszegi (2017)* conclude that – although the change in expectations brought back the sentiment preceding the crisis – firms remained unable to produce growth relative to pre-crisis levels. On balance, demand for finance grew once the effects of the recession were eliminated; however, that meant returning to the previous level without any additional financing requirement resulting from growth. *Bálint and Fellner’s (2017)* report, referred to earlier, also investigated developments in credit demand. The authors explain that while demand for long-term loans dropped significantly in 2010–2013 and then increased again later, no typical trends could be identified in demand for short-term bank loans. As a combined result, between early 2010 and 2013 Q2 the growth rate of long-term corporate loans was negative in each quarter, but the situation improved from 2013 Q3, although without any significant growth yet. Even before the period under review, circular debt was common, especially in construction and among firms with a headcount below 50 employees (*Szűcs 2008*), and this persisted even in 2016. Such debt generally qualified as past due accounts payable, and often also as bad debt from the suppliers’ perspective. According to the GVI survey (*Nábelek 2016*), an outstanding improvement was recorded in payment discipline in 2014.

Of respondents to the *SAFE (2015)* survey, 21.6 per cent had used *short-term bank financing* (credit line, overdraft, credit card) in the preceding six months, and 44 per cent considered this channel to be relevant to their business. Leases were the second most popular form of financing, used by 14.2 per cent of respondents recently, and considered as relevant by 37.7 per cent. Collectively, *subsidised bank loans and grants* were used by 8.9 per cent of respondents over the period, and 29.3 per cent considered them to be relevant. *Non-subsidised bank loans* were used by 6.7 per cent, and other loans by 8.4 per cent of respondents on a six-month time horizon. *Trade credit* was both less used (4.3 per cent of respondents) and considered to be less relevant (13.22 per cent) than the channels mentioned previously. The use of *factoring* was reported by an even lower percentage. According to the MNB Financial Stability Report based on credit institution data

(MNB 2017), finance was used for investments in tangible assets and the purchase of inventories.

Although the cited studies vary as to the precise timing of the turning point due to differences in terms of focus, the periods under review and the availability of samples, on balance access to finance improved from 2013 Q3 onwards. In the banking sector, the volume of loans disbursed increased, but at the level of individual firms there was no significant change in the ratio of long-term bank loans to total assets. Trade credit was most used in 2013. Although trade credit was not listed among the most important sources of external finance, the data suggest that it was an important factor in the operations of domestic firms.

3. Theoretical framework

Trade credit can have a variety of unique motivations. In *Petersen and Rajan (1996)*, the theories seeking to identify the reasons for the use of trade credit are grouped around three motives: (1) special financing advantages (in information acquisition, in controlling the buyer, in salvaging value from existing assets); (2) price discrimination; and (3) transaction costs.

Financing advantages (1) are built on three different arguments. According to the argument based on the first, suppliers – by virtue of their relationship with their buyers – may have more extensive and more accurate information, and better monitoring opportunities compared to other lenders (*Emery 1984; Smith 1987; Freixas 1993; Biais and Gollier 1997*). Arguments for the supplier's bargaining power over buyers hold that the supplier's position enables it to control the buyer, or even hold it at bay, through its ability to decide on future supplies. In explanations arguing with advantages from collateral assets, current assets obtained from the supplier may be used as collateral in the event of a payment default, an opportunity that cannot be taken in the case of cash advances. In the price discrimination approach (2), the phenomenon is captured from the supplier's point of view: *Schwartz (1974)* construes trade credit as the supplier's pricing policy, whereas *Brennan, Maksimovic and Zechner (1988)* argue that suppliers apply price discrimination through their ability to set more flexible payment terms for more price-sensitive buyers. The main proponent of explanations based on transaction costs (3) is *Ferris (1981)*, who argues that trade credit exists because payments could be settled cheaper for both buyers and suppliers if settlement only takes place periodically (e.g. monthly or quarterly).

Burkart and Ellingsen's (2004) model provides a comprehensive theoretical description of the rationale for firms to use trade credit instead of other forms of finance. The model allows firms to use both bank finance and trade credit, but with both forms of finance there is information asymmetry between the firm and the

lender, which imposes a credit constraint. Since under the model the bank knows less about the firm than the supplier does, the firm will face stricter constraints in the course of obtaining external funding. Consequently, due to the closer relationship, the firm can use trade credit even when it no longer has access to bank credit. Another difference is that the firm will get cash in the case of bank credit, and inputs, i.e. current assets that it can use for production, in the case of trade credit. While it is easy for the firm to realise private benefits from the cash, it is more difficult to transform less liquid inputs to private benefits. One possible consequence is that given a high level of information asymmetry, it is worthwhile for a firm to acquire assets by becoming indebted to suppliers, and then to obtain bank credit on grounds of the collateralisation so established. Burkart and Ellingsen identify that scenario as the case where trade credit is able to complement bank credit.

The relationship between the two forms of external funding was investigated empirically on samples of US firms by *Nilsen (2002)*, Canadian firms by *Cunningham (2005)*, and Chinese firms by *Ying, Guo and Yang (2014)*. Most of their findings are consistent with Burkart and Ellingsen's theory. Burkart and Ellingsen viewed trade credit as a source of finance that matured in 1 to 3 months but could be prolonged, and loans provided by commercial banks typically maturing in 1 to 3 years. The primary focus of the theory is therefore on short-term sources of funding. The question of how long-term bank finance is related to trade credit is addressed by *Fazzari and Petersen (1993)* as part of their analysis of US firms' data. The authors observe that investments in fixed assets and net working capital compete with each other for finance, which makes long-term bank credit and trade credit possible substitutes. *Ferrando and Mulier (2013)* find that trade credit may also be an important source of growth.

The decision between trade credit and bank credit may also be strongly influenced by how the firm's buyers are financed. Firms with the capacity to increase their accounts payable will have more room to finance their buyers for price discrimination purposes. Moreover, accounts receivable are less liquid than cash, and additional collateral may also improve the capacity to borrow from banks. For European firms, the components of trade credit were analysed by *García-Teruel and Martínez-Solano (2010)*, whose findings, *inter alia*, confirmed the theory of price discrimination.

Decisions on inventories may also play a part in the use of trade credit, as shown in a theoretical model by *Bougheas, Mateut and Mizzen (2007)*, who also test their argument empirically on a sample of firms operating in the UK. *Chittenden and Bragg (1997)* confirm the relationship between inventory holdings and accounts payable on samples of three countries (Germany, France, UK). As regards cash

holdings, also analysing the data of firms in the UK, *Kling, Paul and Gonis (2014)* find that a higher cash ratio improves access to trade credit.

For our empirical analysis, we invoke the following conclusions by *Burkart and Ellingsen (2004)*:

- 1) With financially unconstrained firms that are still within their bank debt and trade credit capacities, according to the theory complementary, substitution or no relationship between short-term bank credit and trade credit may be possible.
- 2) With financially constrained firms beyond their capacity for bank debt, short-term bank loans and trade credit are:
 - a) substitutes in the case of firms that are still within their capacity for trade debt; and
 - b) complements in the case of firms that are beyond their capacity for trade debt.

The implications can also be extended in terms of business cycles. Where the size of bank credit is pro-cyclical over time in a firm (i.e. showing signs of constraints), trade credit would be:

- counter-cyclical in case 2a; and
- pro-cyclical over time in case 2b.

Therefore, the sensitivity of trade credit to business cycles could vary according to the respective levels of the capacity constraint on trade credit and bank credit. The authors identify the wealth (own funds) available to the firm as the factor that influences whether capacity constraints are reached. Firms with a low level of own funds are constrained in their access to both bank credit and trade credit, those with a medium level of own funds only in their access to bank credit, and firms with a high level of own funds are not constrained in their access to either type of credit.

In our first analysis, we explore how accounts payable are related to short-term and long-term lending. We also address the effect of accounts receivable on accounts payable. Our second question is whether patterns that are similar to the theoretical conclusions are obtained by capturing own funds through firms' average profitability, and the constraints on bank debt in ways commonly reported in the literature (using Kaplan–Zingales, Whited–Wu and Cleary indices). In practice, the capacity for trade debt can be influenced by a number of factors in addition to internal funds. Where a firm has stronger bargaining power over its supplier, it may have much greater capacity for trade debt. Therefore, in the third step we will investigate the extent to which the relationship between bank credit and trade credit varies according to firm size. As an additional fourth step, we profile the behaviour of accounts payable over time.

4. Data and methodology

The data were provided by Bisnode Magyarország Kft. The raw database includes all Hungarian firms that use double-entry bookkeeping and submit annual accounts to the tax authority. Of that population, our analysis includes Hungarian firms that draw up their annual accounts in Hungarian forints and are neither financial firms (NACE 6400–6699) nor utility providers (NACE 3500–3799). The sample does not include firms submitting simplified annual accounts or the type of accounts provided for microenterprises. Although there is a high number of firms of the latter type, their collective size is considerably smaller than the group of firms submitting annual accounts. Consequently, the macro-weight of the dataset used may be considered significant. In cleaning the data, we filtered out firms showing economically nonsensical or erroneous items (such as negative total assets or other negative balance sheet items), and where we found missing data that could not be provided (e.g. missing P/L accounts), we also filtered out the corresponding firm/year observations. Subsequently, the dataset used for analysis was constructed to include firms for which the database contained balance sheets and P/L accounts for at least three consecutive years.

This corresponds to a total of 14,554 firms, comprised of 68 large corporations, 1,332 medium-sized enterprises, 4,354 small enterprises, and 8,800 microenterprises according to the EU sales classification. Except for financial firms and utility providers, which operate using different a business logic, our database thus includes all Hungarian firms on which the tax authority has data of adequate quality. Within the sample, large corporations account for approximately 10 per cent, medium-sized enterprises 46 per cent, small enterprises 37 per cent, and microenterprises 7 per cent of the total sales.

	2010	2011	2012	2013	2014	2015
Commerce	3,190	3,375	3,547	3,215	2,918	1,710
Industry	4,372	4,594	4,773	4,418	4,079	2,502
Services	5,167	5,616	5,840	5,188	4,557	2,507
Total	12,729	13,585	14,160	12,821	11,554	6,719

Note: Authors' calculation on the sample used for analysis

Based on the NACE classification, the sample included 3,665 firms in commerce, 4,896 firms in industry, and 5,993 firms in services (*Table 2*). Except for 2015 (where the annual accounts were not available for all of the firms), the database includes data on an average of 12,000 to 13,000 firms, with a low ratio of new entries and exits.

Table 3

Names of the variables

ACCOUNTS_PAYABLE:	Accounts payable (purchases of goods and services)
ACCOUNTS_RECEIVABLE:	Accounts receivable (goods and services)
ST_LOANS:	Short-term loans
LT_LOANS:	Long-term loans (for investment and development purposes)
ASSET:	Total assets (for regression analysis, its logarithm is used: LOGASSET)
SALES:	Net sales
AP_SALES:	Accounts payable / Net sales
AR_SALES:	Accounts receivable / Net sales
EBIT_SALES:	Earnings before interest and taxes / Net sales
ST_LOANS_SALES:	Short-term loans / Net sales
LT_LOANS_SALES:	Long-term loans / Net sales
LEVERAGE:	(Short-term liabilities + Long-term liabilities) / Total assets
CURRENT:	Current assets / Short-term liabilities
CASH_SALES:	(Cash and cash equivalents) / Net sales
INTEREST_COV:	Interest coverage ratio: Earnings before interest and taxes / (Interest payable and similar charges + Approved dividends and profit sharing)
CAPEX_SALES:	(Fixed assets – Depreciation) / Net sales
SALES_GROWTH:	Sales growth rate as a percentage of the previous year
PROFIT_SALES	Profit after taxes _t / Sales _{t-1}
ROE:	Profit after taxes _t / Shareholders' equity _{t-1}
ROIC:	NOPLAT _t / IC _{t-1} where: NOPLAT: EBIT × (1 – t _{eff}) IC: Fixed assets + Current assets – Short-term liabilities
DIVIDEND_RATIO:	Approved dividends and profit sharing / Profit after taxes
KZ:	<i>Kaplan and Zingales (1997)</i> index for unlisted firms: KZ = -1.002 × CF/ASSETS + 3.3139 × LEVERAGE – 39.368 × DIV – 1.315 × CASH/ASSETS
WW:	<i>Whited-Wu (2006)</i> index: WW = -0.091 × CF/ASSETS – 0.062 × DIVIDEND_PAYER + 0.021 × LEVERAGE – 0.044 × LOGASSET + 0.102 × IND_SALES_GROWTH – 0.035 × SALES_GROWTH where IND_SALES_GROWTH is the industry average sales growth (NACE 2 digits), and DIVIDEND_PAYER is the dummy for firms paying dividends
CLEARY:	<i>Cleary (1999)</i> index: CLEARY = -0.119 × CURRENT – 1.904 × LEVERAGE + 0.001 × INTEREST_COV +1.456 × PROFIT_SALES + 2.035 × SALES_GROWTH – 0.048 × SLACK, where the indicator of financial slack SLACK = (Cash and cash equivalents + 0.5 × Inventories + 0.7 × (Accounts receivable – Short-term liabilities)) / Fixed assets
COST_SALES	Material costs
AP_COST_SALES:	Accounts payable / Material costs

Table 3 lists the short names and descriptions of the variables used in our analysis. In the international literature, it is common (cf. e.g. *Petersen and Rajan 1997, Nilsen 2002, Cunningham 2005, or Bougheas, Mateut and Mizen 2007*) that instead of using the AP/Sales ratio to derive the Days Payable Outstanding ($AP/COST_SALES \times 365$) variable, it is used as a measure of the size of accounts payable, normalised for the size of sales. For the sake of international comparability, we have opted for the latter variable. We measured accounts payable using the “Accounts payable (purchases of goods and services)” line of the balance sheets. To identify short-term loans, we used data from the “Short-term loans” line within the “Short-term liabilities” item of the balance sheet, disregarding the “Short-term advances” item. The item short-term loans may also include repayments due within one year on loans maturing in more than one year; to that extent, a critical approach is needed to our findings concerning short maturities. Although the item chosen does not necessarily include bank loans only, in terms of its function it aggregates short-term funds that are not from suppliers, and as such may be considered as an alternative to trade credit.

With one indicator group (AP_SALES, AR_SALES, CASH_SALES, ST_LOANS_SALES, LT_LOANS_SALES, SALES_GROWTH, AP/COST_SALES) we winsorised the upper 1 per cent, whereas in other cases (EBIT_SALES, PROFIT_SALES, CAPEX_SALES, ROIC) this was done for both the lower and upper 1 per cent. With more special variables we applied different rules, e.g. COVERAGE_RATIO (20%; 70%), INTEREST_COV (20%; 70%), SLACK (5%; 90%), ROE (8%; 90%).

We summarised the descriptive statistics on the most important variables in the database in *Table 4*. In *Panel A*, we considered data for the entire 2010–2015 period, while *Panel B* only contains data for 2011–2015, because with several calculated ratios the denominator contains a value for 2010. In the sample, the median of accounts payable was approximately HUF 16 million, and that of accounts payable was HUF 20.5 million. More than one third of the firms had short-term bank debt, and a smaller percentage had long-term bank debt. The median of total assets was HUF 385 million in the sample. The median of the AP/Sales ratio was 5 per cent, and that of the AR/Sales around 8.4 per cent.

Table 4
Summary statistics

<i>Panel A: Key items</i>								
(HUF thousands)	N	Mean	St. Dev.	Min	Pct(25)	Median	Pct(75)	Max
ACCOUNTS_PAYABLE	71,568	105,172	228,182	0	694	15,916	98,422	2,138,715
ACCOUNTS_RECEIVABLE	71,568	154,412	404,037	0	976	20,588	139,871	16,552,101
ST_LOANS	71,568	75,464	390,532	0	0	0	17,542	15,724,355
LT_LOANS	71,568	26,932	142,189	0	0	0	0	2,146,469
ASSET	71,568	1,254,481	7,441,845	1	51,983	385,357	1,172,100	1,240,000,000
SALES	71,568	1,220,289	2,762,188	101	35,276	332,765	1,300,015	111,000,000

<i>Panel B: Indicators describing operations</i>								
Ratios	N	Mean	St. Dev.	Min	Pct(25)	Median	Pct(75)	Max
AP_SALES	57,014	0.139	0.352	0.000	0.010	0.050	0.127	2.838
AR_SALES	57,014	0.154	0.258	0.000	0.016	0.084	0.182	1.805
EBIT_SALES	57,014	-0.053	0.747	-5.643	0.000	0.035	0.111	0.964
ST_LOANS_SALES	57,014	0.074	0.263	0.000	0.000	0.000	0.033	2.126
LT_LOANS_SALES	57,014	0.049	0.270	0.000	0.000	0.000	0.000	2.362
LEVERAGE	57,014	0.488	0.302	0.000	0.232	0.473	0.734	0.989
CURRENT	57,014	3.555	5.388	0.000	0.960	1.538	3.109	22.548
CASH_SALES	57,014	0.352	1.081	0.000	0.017	0.066	0.220	8.546
INTEREST_COV	57,014	4.753	12.281	-14.587	0.007	1.865	21.187	22.445
CAPEX_SALES	57,014	0.074	0.408	-1.653	0.000	0.012	0.058	2.777
SALES_GROWTH	57,014	0.186	0.958	-1.000	-0.119	0.028	0.200	7.063
ROE	57,014	0.155	0.521	-0.907	0.002	0.085	0.316	1.201
ROIC	57,014	0.247	1.302	-4.487	-0.000	0.076	0.256	8.685

In empirical literature on corporate finance, the variables measuring the presence of financial constraints are compared by *Farre-Mensa and Ljungqvist (2014)* and *Elsas and Klepsch (2016)*. Of the variables used by the authors, we measure the presence of constraints on access to short-term bank credit against the three most common variables: the Kaplan–Zingales, the Whited–Wu and the Cleary indices for unlisted firms. We define the presence of financial constraints based on the statistical procedures summarised in *Farre-Mensa and Ljungqvist (2014)*. The Kaplan–Zingales and Whited–Wu indices and the Cleary index multiplied by minus one are arranged in ascending order, and then observations above the upper 30th percentile are defined as unconstrained, and those below the lower 30th percentile are defined as constrained. Items in the middle 40 per cent belong to neither class and are considered as representing a grey zone. We applied this procedure and the sorting to the entire sample and to the entire firm/year observation. This means

that the classification of certain firms may change over the year, but change is not characteristic of the sample. The three classifications are compared in *Table 5*. Except for large corporations, the Kaplan–Zingales approach can be used to classify firms more or less independently of firm size. With large firms, the share of firms at the limit of their capacity constraints is about 50 per cent, which indicates the poor reliability of the KZ procedure (in their original paper, *Kaplan and Zingales (1997)* used a relatively homogeneous sample to estimate the index for categories based on self-assessment). The Cleary procedure has a moderate tendency, while the Whited–Wu procedure has a strong tendency to classify smaller firms, particularly microenterprises, as financially constrained. It is to be noted that these procedures are also not considered to be perfect alternatives in the international literature, which is why we use several classifications for our analyses.

Table 5
Classification of financial constraint measures

Size	Year	Number of firms	Number of financially constrained firms in each classification									
			KZ	CLEARY	WW	KZ & CLEARY	KZ & WW	CLEARY & WW	All three	KZ / All	CLEARY / All	WW / All
Large	2011	68	39	5	0	3	0	0	0	57%	7%	0%
Large	2012	68	35	5	1	2	1	1	1	51%	7%	1%
Large	2013	67	30	8	1	3	1	1	1	45%	12%	1%
Large	2014	67	34	6	0	2	0	0	0	51%	9%	0%
Large	2015	34	14	6	1	1	0	1	0	41%	18%	3%
Medium	2011	1,269	434	143	5	63	5	0	0	34%	11%	0%
Medium	2012	1,310	431	181	11	80	9	7	5	33%	14%	1%
Medium	2013	1,298	416	195	15	79	11	12	9	32%	15%	1%
Medium	2014	1,254	365	192	18	64	13	12	9	29%	15%	1%
Medium	2015	734	184	128	9	43	6	9	6	25%	17%	1%
Small	2011	3,966	1,177	698	110	317	72	50	32	30%	18%	3%
Small	2012	4,201	1,233	867	121	353	86	74	49	29%	21%	3%
Small	2013	4,168	1,150	864	134	369	92	103	68	28%	21%	3%
Small	2014	3,967	1,041	818	131	324	86	90	57	26%	21%	3%
Small	2015	2,333	498	494	72	160	47	57	35	21%	21%	3%
Micro	2011	8,282	2,497	2,791	4,110	1,085	1,538	1,827	656	30%	34%	50%
Micro	2012	8,581	2,735	3,273	4,392	1,200	1,684	2,164	799	32%	38%	51%
Micro	2013	7,288	2,179	2,706	3,448	957	1,264	1,659	576	30%	37%	47%
Micro	2014	6,266	1,758	2,311	2,873	773	976	1,390	446	28%	37%	46%
Micro	2015	3,618	854	1,413	1,652	388	477	888	240	24%	39%	46%

In the first analysis, we examine how accounts payable are related to short-term and long-term bank loans and generally to accounts receivable under constrained access to bank finance. For that purpose, the following regression equations are estimated:

$$AP/SALES = ST_LOANS/SALES + LT_LOANS/SALES + AR/SALES + \text{controls} + \text{Firm FE} + \text{Year FE} \quad (1)$$

$$AP/SALES = ST_LOANS/SALES + LT_LOANS/SALES + AR/SALES + \text{FINC_Dummy} + \text{FINC_Dummy} \times ST_LOANS/SALES + \text{controls} + \text{Firm FE} + \text{Year FE} \quad (1A)$$

$$AP/SALES = ST_LOANS/SALES + LT_LOANS/SALES + AR/SALES + \text{FINC_Dummy} + \text{FINC_Dummy} \times LT_LOANS/SALES + \text{controls} + \text{Firm FE} + \text{Year FE} \quad (1B)$$

$$AP/SALES = ST_LOANS/SALES + LT_LOANS/SALES + AR/SALES + \text{FINC_Dummy} + \text{FINC_Dummy} \times AR/SALES + \text{controls} + \text{Firm FE} + \text{Year FE} \quad (1C)$$

where *FINC_Dummy* (financially constrained) is the indicator variable for financially constrained firms beyond their capacity for bank debt. Firm FE indicates firm fixed effects, and Year FE the time fixed effects. The control variables include all variables that, in addition to the effects in individual firms, adequately capture individual operational and growth characteristics: firm size ($\log(\text{ASSET})$), operating margin (*EBIT_SALES*), profitability (*ROE*), interest coverage ratio (*INTEREST_COV*), leverage ratio (*LEVERAGE*) lagged, liquidity ratio (*CURRENT*) lagged, cash to sales ratio (*CASH_SALES*) lagged, capital expenditures to sales (*CAPEX_SALES*) lagged, dividend ratio (*DIVIDEND_RATIO*), sales growth (*SALES_GROWTH*).

The first idea for constructing Equation (1A) came from *Cunningham (2005)*, but we use different set of base and control variables. Equation (1B) explores the relationship of accounts payable and long-term credit in order to determine whether firms increase their long-term credit by increasing trade credit. Equation (1C) is not derived from the theoretical model, but is based rather on practical considerations. In this case, we seek to determine whether in cases with higher AR/Sales ratios firms pass financing on to their suppliers.

In the second analysis, we also examine access to trade credit. In the theoretical model, *Burkart and Ellingsen (2004)* consider the wealth of the firm to determine the finance option in terms of which the firm will become constrained. In the case of high wealth, firms are less reliant on external funds, and neither constraint applies. In the case of medium wealth, they exploit their capacities for bank debt, but not

that for trade debt. In the case of low wealth, they are strongly reliant on external funds, and as such face constraint in both forms. As firms' wealth is difficult to observe, average profitability (EBIT/ASSET, i.e. Earnings before interest and taxes to Total assets) is used as a proxy for wealth, the rationale being that a highly profitable firm has the ability to utilise internal funds. We divide the database into three sub-samples according to low (<0 per cent), medium (0–14 per cent) and high (>14 per cent) business profitability based on the firms' average EBIT/ASSET ratios. On the sub-samples, of the previous model specifications we run (1A) for the three constraint indicators, of which we use only the indicator variable for firms constrained in their access to bank finance, and consider the control group as including both the grey zone and firms unconstrained in their access to bank credit.

In our third analysis, we seek to determine the extent to which firm size influences the relationship between trade credit and short-term bank credit. In the case of US firms, *Nilsen (2002)* found firm size to be the factor that adequately captured access, although he did not carry out his analyses on micro-level panel data. We re-estimate our previous Equations (1) and (1A) for four sub-samples, i.e. the groups of large, medium, small and microenterprises, using the criteria for sales (EUR 50 million, EUR 10 million, and EUR 2 million, respectively; assuming an exchange rate of EUR/HUF 300) from the size definitions commonly applied to Hungarian firms. In a supplier-buyer relationship, larger size may give a firm more bargaining power over its supplier, as a result of which larger firms rarely face strict supplier constraints. Primarily, we seek to determine the extent to which this relationship is captured by firm size, but exploring the differences between firms of various sizes could in itself also produce noteworthy findings. For that purpose, the previously constructed Equation (1A) is again estimated with the three constraint dummies.

The fourth analysis quantifies the sensitivity of trade credit to business cycles. For each sub-sample (segmented by size), the following regression equation is used to obtain a direct estimate of days payable outstanding:

$$\begin{aligned}
 AP/COST_SALES \times 365 = & AR/SALES \times 365 + ST_LOANS/SALES + LT_LOANS/SALES + \\
 & + FINC_Dummy + FINU_Dummy + \sum_{t=1}^4 YEAR_Dummy_t + \\
 & + \sum_{t=1}^4 YEAR_Dummy_t \times FINC_Dummy + \quad (2) \\
 & + \sum_{t=1}^4 YEAR_Dummy_t \times FINU_Dummy + \\
 & + controls + Firm FE
 \end{aligned}$$

where the dummies FINC and FINU (financially unconstrained) do not cover the entire sample, given that 40 per cent of it falls into the grey zone. From the estimate, we seek to establish time fixed effects and the interactions with those effects. In other words, we seek to determine the developments in days payable outstanding, solely attributable to cyclical effects and other things being equal, for the segments of firms with varying access to finance, assuming their capacities for bank debt to be exploited. For this purpose, we multiply the dependent variable by 365 days in order to construe the result in terms of days. As regards the COST_SALES variable, we also had to take into account the fact that not all firms follow the same logic in calculating the effect of the cost of sales on profits. Some of the firms submitted P/L accounts on a total cost basis, and others on a turnover basis. The base which, if not settled immediately, could become accounts payable, was the sum of materials and consumables and contracted services in the first case, and the sum of the cost of goods sold and the value of services sold in the second case. Both types of reporting included the item of material costs, which differed from the sum calculated using the two methods in fewer than 1 per cent of all cases, and even then the difference was small.

Regressions were estimated with the *fixed effects within panel* method, where unique firm fixed effects were filtered using the *within* estimator, and time (year) effects using dummy variables. The panels are unbalanced because it was possible for the firms observed to be included in the sample for less than five years. Both the Breusch–Pagan and the Honda LM test rejected the null hypothesis that the time effect was not significant in the regressions. Unique (firm) effects were also mostly significant. The Hausmann tests run for the estimates supported the use of fixed effects versus random effects. The error terms of the estimated models implied heteroscedasticity, but in the error terms autocorrelation was not significant for four observation years. Consequently, we used White’s heteroscedasticity-consistent standard errors for our statistical tests.

5. Results

The estimation results for the first question of our analysis (Equations (1), (1A), (1B) and (1C)) are summarised in *Table 6*, where the coefficients of the basic equation are shown in the first column, followed by the three equations with the constraint indicators derived using classifications based on the Kaplan–Zingales, Whited–Wu and Cleary indices. The dependent variable is the ratio of accounts payable to sales, which is reported as multiplied by 100 to ensure the readability of the coefficients.

Table 6
Relationship of accounts payable or bank credit to accounts receivable in financially constrained firms

(based on Kaplan–Zingales, Whited–Wu and Cleary classifications)

	(1)		(1A)		(1B)		(1C)			
	Base	KZ	WW	CLEARY	KZ	WW	CLEARY	KZ	WW	CLEARY
AP_SALES × 100	12.69*** (2.24)	5.03 (2.74)	8.85*** (2.28)	5.39** (1.80)	12.52*** (2.24)	12.60*** (2.25)	12.67*** (2.24)	11.82*** (2.20)	12.82*** (2.24)	12.59*** (2.24)
LOANS_LT_per_SALES	0.64 (2.31)	0.48 (2.28)	0.50 (2.30)	0.04 (2.29)	0.37 (3.17)	-0.70 (2.27)	0.35 (2.98)	0.56 (2.33)	0.69 (2.29)	0.60 (2.28)
AR_per_SALES × 100	0.38*** (0.03)	0.37*** (0.03)	0.37*** (0.03)	0.37*** (0.03)	0.37*** (0.03)	0.38*** (0.03)	0.38*** (0.03)	0.25*** (0.03)	0.35*** (0.03)	0.32*** (0.03)
D.KZ.constr		2.26** (0.71)			3.39*** (0.69)			-1.27 (0.88)		
LOANS_ST_per_SALES:D.KZ.constr		11.48** (3.74)								
LOANS_LT_per_SALES:D.KZ.constr					-4.61 (3.87)					
AR_SALES × 100:D.KZ.constr										
D.WW.constr			-0.62 (0.86)			0.18 (0.88)			-0.76 (0.97)	
LOANS_ST_per_SALES:D.WW.constr			17.20** (5.26)							
LOANS_LT_per_SALES:D.WW.constr						7.88 (5.33)				
AR_SALES × 100:D.WW.constr									0.06 (0.04)	

Table 6
Relationship of accounts payable or bank credit to accounts receivable in financially constrained firms
(based on Kaplan–Zingales, White–Wu and Cleary classifications)

	(1)	(1A)		(1B)		(1C)	
	Base	KZ	WW	CLEARY	KZ	WW	CLEARY
AP_SALES × 100							
D.CLEARY.constr				-1.38** (0.44)			-1.61** (0.56)
LOANS_ST_per_SALES:D.CLEARY.constr				13.68*** (2.91)			
LOANS_LT_per_SALES:D.CLEARY.constr					0.51 (3.24)		
AR_SALES × 100:D.CLEARY.constr							0.08** (0.03)
lag(LOGASSET)	-6.94*** (0.73)	-6.85*** (0.73)	-6.84*** (0.73)	-6.82*** (0.73)	-6.89*** (0.74)	-6.93*** (0.73)	-6.94*** (0.73)
EBIT_per_SALES	-8.93*** (0.92)	-8.81*** (0.92)	-8.73*** (0.92)	-8.75*** (0.92)	-8.88*** (0.92)	-8.86*** (0.92)	-8.86*** (0.93)
ROE	0.18 (0.50)	0.64 (0.52)	0.18 (0.50)	0.26 (0.51)	0.65 (0.52)	0.18 (0.50)	0.09 (0.50)
COVERAGE_RATIO	0.00** (0.00)	0.00** (0.00)	0.00** (0.00)	0.00** (0.00)	0.00*** (0.00)	0.00** (0.00)	0.00** (0.00)
lag(LEVERAGE)	10.13*** (1.71)	9.15*** (1.72)	9.87*** (1.71)	9.97*** (1.70)	9.31*** (1.72)	10.07*** (1.71)	10.11*** (1.71)
lag(CURRENT)	0.12* (0.06)	0.11 (0.06)	0.11* (0.06)	0.12* (0.06)	0.11 (0.06)	0.12* (0.06)	0.11* (0.06)

Table 6
Relationship of accounts payable or bank credit to accounts receivable in financially constrained firms
(based on Kaplan-Zingales, White-Wu and Cleary classifications)

	(1)		(1A)		(1B)		(1C)					
	Base		KZ	WW	CLEARLY	KZ	WW	CLEARLY	KZ	WW	CLEARLY	
AP_SALES x 100												
lag(CASH_per_SALES)	1.88** (0.59)		1.89** (0.59)	1.87** (0.59)	1.90** (0.59)	1.88** (0.59)	1.85** (0.59)	1.88** (0.59)	1.92** (0.59)	1.86** (0.59)	1.93** (0.59)	
lag(CAPEX_per_SALES)	-0.14 (0.80)		-0.11 (0.81)	-0.09 (0.80)	-0.18 (0.81)	-0.11 (0.80)	-0.18 (0.81)	-0.14 (0.81)	-0.08 (0.81)	-0.11 (0.80)	-0.17 (0.81)	
SALES_GROWTH	-2.67*** (0.39)		-2.73*** (0.39)	-2.65*** (0.40)	-2.75*** (0.40)	-2.64*** (0.40)	-2.69*** (0.38)	-2.72*** (0.40)	-2.72*** (0.38)	-2.60*** (0.40)	-2.67*** (0.40)	
DIVIDEND_RATIO	-0.00*** (0.00)		-0.00*** (0.00)									
YEAR2013	1.23*** (0.25)		1.25*** (0.25)	1.18*** (0.25)	1.11*** (0.25)	1.22*** (0.25)	1.26*** (0.25)	1.23*** (0.25)	1.18*** (0.24)	1.24*** (0.25)	1.26*** (0.25)	
YEAR2014	2.01*** (0.31)		2.05*** (0.31)	1.94*** (0.31)	1.90*** (0.31)	1.99*** (0.31)	2.09*** (0.31)	2.01*** (0.31)	1.90*** (0.30)	2.00*** (0.31)	2.02*** (0.31)	
YEAR2015	2.01*** (0.40)		2.11*** (0.39)	1.93*** (0.40)	1.90*** (0.39)	2.01*** (0.40)	2.14*** (0.39)	2.02*** (0.40)	1.95*** (0.39)	2.00*** (0.40)	1.99*** (0.40)	
N	42,460		42,460	42,460	42,460	42,460	42,460	42,460	42,460	42,460	42,460	
n	14,554		14,554	14,554	14,554	14,554	14,554	14,554	14,554	14,554	14,554	
T	2012–2015		2012–2015	2012–2015	2012–2015	2012–2015	2012–2015	2012–2015	2012–2015	2012–2015	2012–2015	
R ²	0.1968		0.2002	0.2006	0.2017	0.1972	0.1982	0.1969	0.2119	0.1976	0.1976	

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; $p < 0.1$; while character ‘.’ represents interaction between variables.

Based on an overview of the table, we find that the level of accounts payable was increased by the presence of bank credit in all estimates, which is implied by the significantly positive ST_LOAN_SALES coefficients. Due to the specific nature of the *within* estimator, R-square values are low, but overall the variables significantly explain the ratio of accounts payable with a 1 per cent threshold for each of the regression equations.

In the estimate for the basic equation, with firms where the ratio of short-term loans to sales was 10 percentage points (0.1) higher, the ratio of accounts payable was also 1.269 percentage points higher on average. This essentially suggests that if Burkart and Ellingsen's theoretical model is assumed to be correct, and this kind of complementarity is observed in low-wealth firms that are constrained in their access to both trade credit and bank credit, then according to the authors' theory most Hungarian firms are financially constrained. This could also mean that the credit constraint indicators introduced capture the extent of the capacity constraint rather than the fact of its presence.

A brief overview of the coefficients obtained for the control variables is also appropriate. We found that larger firms and those working with higher profit margins have a lower average AP-to-sales ratio. The ROE indicator, as a measure of profitability, failed to significantly explain the size of accounts payable, while the coefficient of the dividend ratio showed zero effect. For firms with a higher level of indebtedness, we found higher accounts payable in the following year. The AP-to-sales ratio was observed to be higher for firms with higher cash ratios, and lower for firms with higher growth rates. All of this is not contrary to general business intuition.

Financial constraint indicators did not produce consistent results in every case. Other than the above argument, this could also result from the differences in constructing the indicators, as well as from the fact that the indices used for classification were constructed on US data, while differences in the firms' operations may also undermine the reliability of application. Consequently, we only consider estimation results as robust where all three indicators produced similar results. Note that in interpreting our results, we refer to firms as constrained where the value of bank credit is one, and unconstrained where the value is zero (grey zone and firms marked as unconstrained above the 30th percentile collectively).

In analysing Equations (1A), we essentially found short-term loans and accounts payable to be complementary. In themselves, constraint dummy indicators produced mostly insignificant or inconsistent results, but interactions were significant and combined effects (with an average volume of short-term loans) also imply complementarity.

In the case of Equations (1B), no correlation was found between long-term credit and accounts payable: estimating for the entire sample, there is no evidence of any relationship between taking out additional long-term loans and the use of accounts payable.

Estimates for Equations (1C) show that in financially unconstrained firms, an increase in accounts receivable in a given year will be accompanied by a higher ratio of trade credit. On the entire sample, for financially constrained firms combined effects are both insignificant and contradictory. This implies that the ability of firms to pass financing on to their suppliers may depend on sector, size and other factors.

Table 7
Relationship between trade credit and short-term bank credit by wealth

	Low "wealth" (EBIT/ASSET)				Medium "wealth" (EBIT/ASSET)				High "wealth" (EBIT/ASSET)			
	(1) Base	KZ	(1A) WW	CLEARY	(1) Base	KZ	(1A) WW	CLEARY	(1) Base	KZ	(1A) WW	CLEARY
AP_SALES × 100												
LOANS_ST_per_SALES	18.15*** (3.81)	6.48 (4.34)	11.53*** (4.06)	7.14 (3.92)	5.79* (2.26)	0.98 (3.42)	4.66* (2.14)	3.23* (1.63)	13.22 (8.84)	26.30 (15.42)	20.41 (11.67)	6.14 (4.65)
LOANS_LT_per_SALES	-1.03 (4.67)	-1.46 (4.59)	-1.61 (4.59)	-0.97 (4.64)	2.55 (2.70)	2.58 (2.68)	2.61 (2.71)	2.07 (2.65)	-4.76 (3.77)	-4.80 (3.95)	-4.62 (3.77)	-5.01 (3.83)
AR_SALES × 100	0.41*** (0.04)	0.41*** (0.04)	0.41*** (0.04)	0.41*** (0.04)	0.35*** (0.04)	0.35*** (0.04)	0.35*** (0.04)	0.35*** (0.04)	0.22*** (0.06)	0.21*** (0.06)	0.22*** (0.06)	0.22*** (0.06)
D.KZ.constr		-0.57 (1.60)			4.07*** (0.69)					5.15 (2.68)		
LOANS_ST_per_SALES:D.KZ. constr		16.20** (5.77)			7.76 (4.74)					-23.18 (13.71)		
D.WW.constr			-2.12 (1.76)				1.18 (1.13)				0.31 (1.40)	
LOANS_ST_per_SALES:D.WW. constr			23.80** (7.61)				7.49 (6.41)				-21.50 (11.15)	
D.CLEARY.constr				0.22 (1.08)				-1.54** (0.52)				-2.60*** (0.59)

Table 7 Relationship between trade credit and short-term bank credit by wealth (EBIT/ASSET)											
	Low "wealth" (EBIT/ASSET)			Medium "wealth" (EBIT/ASSET)			High "wealth" (EBIT/ASSET)				
	(1) Base	(1A) WW	CLEAR CLEAR	(1) Base	(1A) WW	CLEAR CLEAR	(1) Base	(1A) WW	CLEAR CLEAR		
AP_SALES x 100			15.44***			7.00				14.63	
LOANS_ST_per_SALES:D. CLEARY.constr			(4.54)			(3.80)				(11.15)	
lag(LOGASSET)	-10.64*** (1.42)	-10.60*** (1.41)	-10.49*** (1.40)	-3.01** (0.94)	-2.77** (0.93)	-2.90** (0.93)	-1.98* (0.77)	-1.97* (0.77)	-1.94* (0.76)	-1.95** (0.75)	
EBIT_per_SALES	-8.45*** (1.02)	-8.36*** (1.02)	-8.22*** (1.02)	-9.21*** (2.46)	-9.08*** (2.46)	-9.09*** (2.48)	-7.86* (4.32)	-7.99* (4.32)	-7.94* (4.28)	-7.88* (4.33)	
ROE	-0.63 (0.85)	-0.47 (0.88)	-0.67 (0.84)	-0.10 (0.67)	0.66 (0.69)	0.01 (0.66)	1.61 (1.06)	1.59 (1.11)	2.38* (1.17)	1.10 (1.02)	
COVERAGE_RATIO	0.01*** (0.00)	0.01** (0.00)	0.01** (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	
lag(LEVERAGE)	14.01*** (3.47)	13.34*** (3.44)	13.32*** (3.48)	4.60* (2.15)	2.89 (2.17)	4.47* (2.16)	2.31 (2.27)	2.24 (2.29)	1.31 (2.35)	2.52 (2.29)	
lag(CURRENT)	0.16 (0.13)	0.16 (0.13)	0.15 (0.13)	0.07 (0.08)	0.05 (0.08)	0.07 (0.08)	0.11 (0.06)	0.11 (0.06)	0.10 (0.06)	0.11 (0.06)	
lag(CASH_per_SALES)	2.87** (0.90)	2.87** (0.91)	2.86** (0.90)	0.95 (0.81)	0.96 (0.80)	0.94 (0.81)	0.08 (0.74)	0.10 (0.74)	0.18 (0.73)	0.03 (0.75)	

Table 7
Relationship between trade credit and short-term bank credit by wealth
(EBIT/ASSET)

	Low "wealth" (EBIT/ASSET)				Medium "wealth" (EBIT/ASSET)				High "wealth" (EBIT/ASSET)			
	(1)	KZ	WW	CLEARLY	(1)	KZ	WW	CLEARLY	(1)	KZ	WW	CLEARLY
AP_SALES × 100												
lag(CAPEX_per_SALES)	-0.95 (1.47)	-0.74 (1.48)	-0.78 (1.46)	-0.94 (1.48)	0.45 (0.83)	0.41 (0.83)	0.47 (0.83)	0.43 (0.83)	0.49 (1.30)	0.95 (1.35)	0.71 (1.34)	0.18 (1.35)
SALES_GROWTH	-3.68*** (0.66)	-3.75*** (0.66)	-3.73*** (0.68)	-3.39*** (0.70)	-2.77*** (0.62)	-2.84*** (0.62)	-2.61*** (0.64)	-3.00*** (0.66)	0.03 (0.50)	-0.04 (0.48)	0.04 (0.51)	-0.31 (0.51)
DIVIDEND_RATIO	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	0.02 (0.03)	0.03 (0.04)	0.02 (0.03)	0.01 (0.04)
YEAR2013	2.26** (0.74)	2.33** (0.73)	2.18** (0.74)	2.04** (0.74)	0.37 (0.25)	0.40 (0.25)	0.35 (0.25)	0.33 (0.25)	0.56 (0.37)	0.63 (0.35)	0.57 (0.36)	0.62 (0.37)
YEAR2014	4.29*** (0.88)	4.29*** (0.88)	4.21*** (0.88)	4.16*** (0.88)	0.35 (0.29)	0.46 (0.29)	0.33 (0.29)	0.31 (0.29)	0.58 (0.55)	0.70 (0.54)	0.60 (0.54)	0.57 (0.55)
YEAR2015	4.06*** (1.19)	4.10*** (1.18)	3.82** (1.18)	3.84** (1.18)	0.21 (0.40)	0.39 (0.39)	0.22 (0.40)	0.19 (0.40)	0.60 (0.59)	0.71 (0.58)	0.56 (0.57)	0.69 (0.59)
N	12,383	12,383	12,383	12,383	21,746	21,746	21,746	21,746	8,331	8,331	8,331	8,331
n	4,544	4,544	4,544	4,544	7,021	7,021	7,021	7,021	2,989	2,989	2,989	2,989
T	2012– 2015	2012– 2015	2012– 2015	2012– 2015	2012– 2015	2012– 2015	2012– 2015	2012– 2015	2012– 2015	2012– 2015	2012– 2015	2012– 2015
R ²	0.2389	0.2418	0.2450	0.2436	0.1627	0.1696	0.1639	0.1650	0.0981	0.1098	0.1049	0.1058

Note: ***p < 0.001; **p < 0.01; *p < 0.05; ·p < 0.1; while character ‘·’ represents interaction between variables.

The results of the second analysis are summarised in *Table 7*. We estimated Equations (1) and (1A) on three sub-samples. These included firms (a) with negative average EBIT/ASSET ratios (representing the lower 26.6 per cent of the sample); (b) ratios between 0 and 14 per cent; and (c) ratios above 14 per cent (representing the upper 18.7 per cent of the sample). The selection of the thresholds was determined by the fact that firms with average negative performance are adequately described by Burkart and Ellingsen's low wealth, while the average 14 per cent threshold was supported by the tendency that in business practice, a higher level of profitability already enables firms to raise internal funds, and the fact that the number of items in the sub-sample is not very low either.

The propositions of Burkart and Ellingsen's model are only partially confirmed by our empirical results. In the first sub-sample, we are able to confirm the predicted complementarity. In this subsample, according to the theory, constraints apply in terms of both trade credit and bank credit, which is why a breakdown of the analysis by KZ, WW and Cleary-based indicators would not make any sense. Nevertheless, these indicators, constructed on the basis of other theories, will classify firms as both constrained and unconstrained even in this range. In such cases, the indicators are construed as suggesting that it is difficult for all participants to access bank credit due to their low profitability, and the capacity constraint actually even more applies to firms that are also marked by the indices. This will make short-term loans and accounts payable complements on the sub-sample: a higher ratio of loans will be accompanied by higher accounts payable; however, the result is only robust for severely constrained firms where the combined effects also include interactions. The second sub-sample shows barely significant complementarity in Equation (1). In such a case, for constrained firms the theoretical model would predict substitution, but we found none. We consider it plausible that under Burkart and Ellingsen's theory this falls in the category of low wealth, which explains the complementarity found. In the third sub-sample, that of highly profitable firms, we found no correlation at a significance level of 0.05, and neither is correlation predicted according to the theory. If we accept that we were overly lenient in selecting the wealth threshold, then zero effect could mean that the sub-sample includes firms characterised by substitution as well as firms characterised by complementarity. By way of comparison to research carried out on Canadian data for similar purposes, *Cunningham (2005)* did not succeed in identifying the third category, but did identify complementarity in the first category and then substitution in the second.

Table 8 Relationship between trade credit and short-term bank credit by firm size																
	Large corporations				Medium-sized enterprises				Small enterprises				Microenterprises			
	(1)	KZ	WW	CLEARLY	(1)	KZ	WW	CLEARLY	(1)	KZ	WW	CLEARLY	(1)	KZ	WW	CLEARLY
AP_SALES x 100																
LOANS_ST_per_SALES	-36.73** (12.74)	-41.95** (15.63)	-51.78** (17.35)	-64.18*** (13.11)	38.53*** (10.89)	43.48* (21.36)	38.58*** (10.67)	11.48* (5.40)	15.05*** (3.92)	11.19* (4.80)	11.20** (3.61)	4.49. (2.58)	7.58*** (2.57)	0.18 (2.87)	2.92 (2.61)	3.05 (2.19)
LOANS_IT_per_SALES	-20.85 (90.09)	-7.36 (91.76)	-33.24 (77.11)	-50.17 (76.53)	42.91* (20.11)	43.80* (21.17)	42.95* (20.07)	33.22. (16.96)	5.35 (4.17)	5.75 (4.28)	5.38 (4.22)	3.12 (4.26)	-1.06 (2.42)	-1.54 (2.38)	-1.22 (2.39)	-1.31 (2.43)
AR_SALES x 100	1.10*** (0.32)	1.08*** (0.31)	1.18*** (0.30)	1.20*** (0.21)	0.46*** (0.11)	0.46*** (0.11)	0.45*** (0.11)	0.45*** (0.11)	0.54*** (0.06)	0.54*** (0.06)	0.53*** (0.06)	0.52*** (0.06)	0.31*** (0.03)	0.31*** (0.03)	0.31*** (0.03)	0.31*** (0.03)
D.KZ.constr		-1.92 (1.44)				0.54 (1.63)				-0.29 (0.94)				4.59*** (1.12)		
LOANS_ST_per_SALES.D.KZ. constr		9.38 (9.52)				-6.58 (27.27)				5.50 (5.98)				11.42** (4.14)		
D.WW.constr			-24.71* (11.32)				-1.21 (10.13)					-1.44 (2.87)			-0.97 (0.88)	
LOANS_ST_per_SALES.D.WW. constr							-64.22 (56.24)					35.76* (14.11)			16.82** (5.31)	
D.CLEARLY.constr				0.31 (2.44)				-2.95*** (0.76)								-0.92 (0.61)
LOANS_ST_per_SALES.D.CLEARLY. constr				155.71*** (39.94)				41.21*** (8.26)								8.33* (3.36)
log(LOGASSET)	-2.46 (2.34)	-2.39 (2.31)	-2.57 (2.18)	-1.94 (2.09)	-4.40** (1.48)	-4.35** (1.39)	-4.34** (1.49)	-4.61** (1.46)	-8.42*** (1.58)	-8.45*** (1.59)	-8.19*** (1.56)	-8.12*** (1.55)	-6.25*** (0.89)	-6.08*** (0.89)	-6.15*** (0.88)	-6.18*** (0.89)
EBIT_per_SALES	2.80 (6.94)	3.97 (6.85)	-2.42 (5.97)	38.99*** (10.99)	-9.15* (3.89)	-9.30* (4.30)	-9.23* (3.97)	-7.43* (3.68)	-17.53*** (2.36)	-17.44*** (2.36)	-16.26*** (2.45)	-17.02*** (2.32)	-7.11*** (1.00)	-7.00*** (1.00)	-6.95*** (0.99)	-7.03*** (1.00)
ROE	-4.76 (2.65)	-4.95 (2.65)	-4.91 (2.51)	-5.67* (2.26)	1.49 (0.94)	1.54 (0.99)	1.47 (1.06)	1.11 (0.96)	-0.43 (0.83)	-0.43 (0.86)	-0.45 (0.82)	-0.01 (0.83)	0.30 (0.67)	1.13 (0.69)	0.26 (0.66)	0.31 (0.67)

Table 8 Relationship between trade credit and short-term bank credit by firm size		Large corporations				Medium-sized enterprises				Small enterprises				Microenterprises						
		(1)		(1A)		(1)		(1A)		(1)		(1A)		(1)		(1A)				
		Base	KZ	WW	CLEAR	Base	KZ	WW	CLEAR	Base	KZ	WW	CLEAR	Base	KZ	WW	CLEAR			
AP_SALES × 100																				
COVERAGE_RATIO	0.01 (0.01)	0.01 (0.01)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00* (0.00)											
lag(LEVERAGE)	21.68* (9.36)	21.76* (9.60)	17.54* (8.65)	22.32** (8.47)	-1.00 (6.23)	-0.93 (5.70)	-0.90 (5.68)	0.09 (5.85)	19.27*** (3.28)	19.16*** (3.34)	18.12*** (3.27)	18.48*** (3.24)	7.61*** (2.06)	6.27*** (2.08)	7.41*** (2.06)	7.55*** (2.06)	7.55*** (2.06)			
lag(CURRENT)	0.23 (0.81)	0.30 (0.84)	0.08 (0.81)	0.50 (0.74)	-0.00 (0.12)	0.01 (0.10)	-0.06 (0.11)	0.00 (0.12)	0.03 (0.17)	0.03 (0.17)	0.03 (0.17)	0.02 (0.17)	0.11 (0.06)	0.10 (0.06)	0.11 (0.06)	0.11 (0.06)	0.11 (0.06)			
lag(CASH_per_SALES)	228.23*** (57.67)	229.71*** (58.31)	219.84*** (51.99)	150.36*** (32.03)	-0.25 (3.61)	-0.30 (3.63)	-0.20 (3.58)	0.00 (3.55)	1.95 (2.16)	1.95 (2.16)	2.15 (2.12)	2.16 (2.14)	2.03** (0.62)	2.04** (0.62)	2.01** (0.62)	2.03** (0.62)	2.03** (0.62)			
lag(CAPEX_per_SALES)	142.86* (60.80)	140.65* (61.03)	137.98** (50.62)	67.48* (27.13)	7.19 (5.83)	6.93 (5.15)	7.36 (5.87)	7.89 (5.83)	-0.83 (1.19)	-0.83 (1.19)	-0.83 (1.18)	-0.95 (1.19)	-0.16 (0.90)	-0.15 (0.90)	-0.11 (0.90)	-0.19 (0.91)	-0.19 (0.91)			
SALES_GROWTH	-2.87 (2.70)	-2.83 (2.76)	-3.92 (2.75)	-3.61 (2.52)	-0.84 (0.83)	-0.81 (0.87)	-0.88 (0.79)	-1.27 (0.85)	-1.18 (0.90)	-1.19 (0.90)	-1.38 (0.87)	-1.45 (0.94)	-3.15*** (0.44)	-3.23*** (0.44)	-3.15*** (0.47)	-3.20*** (0.46)	-3.20*** (0.46)			
DIVIDEND_RATIO	0.33 (0.23)	0.26 (0.21)	0.16 (0.22)	0.06 (0.25)	0.23 (0.20)	0.23 (0.20)	0.23 (0.19)	0.25 (0.20)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)			
YEAR2013	1.04 (1.13)	0.97 (1.14)	0.86 (1.06)	0.27 (0.97)	0.92* (0.38)	0.95* (0.40)	0.94* (0.37)	0.86* (0.38)	1.44*** (0.33)	1.43*** (0.33)	1.35*** (0.33)	1.20*** (0.32)	1.04** (0.39)	1.10** (0.39)	1.00* (0.39)	0.98* (0.39)	0.98* (0.39)			
YEAR2014	0.95 (1.14)	0.98 (1.14)	0.77 (1.12)	0.93 (1.12)	1.11 (0.62)	1.14 (0.67)	1.09* (0.54)	0.97 (0.62)	2.25*** (0.38)	2.23*** (0.38)	2.12*** (0.39)	2.00*** (0.37)	1.85*** (0.50)	1.98*** (0.49)	1.76*** (0.50)	1.80*** (0.50)	1.80*** (0.50)			
YEAR2015	2.00 (1.77)	2.05 (1.77)	1.78 (1.69)	1.06 (1.47)	0.62 (0.60)	0.92 (0.68)	0.84 (0.57)	0.78 (0.59)	2.77*** (0.51)	2.76*** (0.50)	2.56*** (0.49)	2.49*** (0.49)	1.65** (0.63)	1.86** (0.62)	1.55* (0.63)	1.59* (0.63)	1.59* (0.63)			
N	235	235	235	235	4,432	4,432	4,432	4,432	13,750	13,750	13,750	13,750	24,043	24,043	24,043	24,043	24,043			
n	68	68	68	68	1,332	1,332	1,332	1,332	4,354	4,354	4,354	4,354	8,800	8,800	8,800	8,800	8,800			
T	2012–2015				2012–2015				2012–2015				2012–2015				2012–2015			
R ²	0.8115	0.8125	0.8222	0.8641	0.4647	0.4653	0.4668	0.4911	0.4442	0.4447	0.4549	0.4563	0.1339	0.1386	0.1380	0.1358	0.1358			

Note: ***p < 0.001; **p < 0.01; *p < 0.05; ·p < 0.1, while character ‘·’ represents interaction between variables.

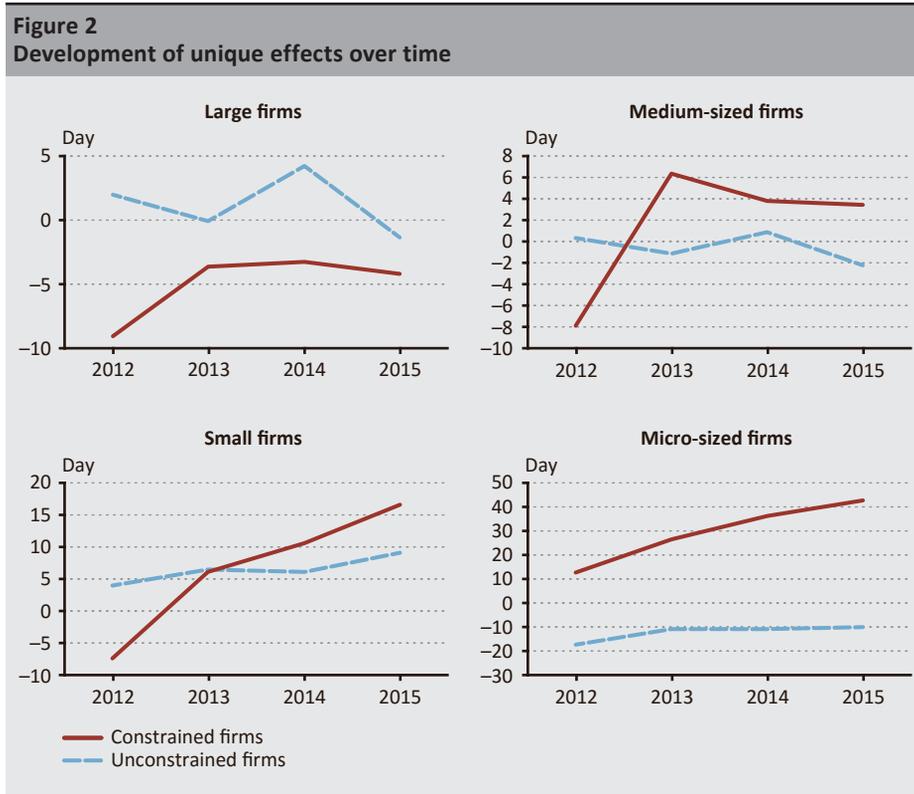
In terms of substitution, noteworthy features are identified by the analysis carried out on sub-samples by firm size, the results of which are reported in *Table 8*. In the sub-sample of large corporations, the estimated coefficient of short-term loans is always negative. This implies substitution in the case of large corporations (regardless of whether they have become financially constrained according to the three indices). The model did not produce an estimate for the Whited–Wu interaction (presumably because only zero or one company is marked by the index as constrained), and the other two indices were found to be inconsistent. In this scenario, we should argue that the relationship between large corporations and banks is different from other cases, i.e. that often the size of large corporations makes it easier for them to use trade credit rather than bank credit, which represents the case of “medium wealth” in Burkart and Ellingsen’s theory. In the category of medium and small enterprises, we found significant complementarity between bank credit and trade credit for financially unconstrained firms. The KZ and WW indices also failed on this sub-sample; only the interaction of the CLEARY dummy was significant. With microenterprises (which are presumably constrained in terms of both forms of funding), we found no correlation in firms where the bank constraint did not actually apply, and found complementarity in firms where it did. This scenario corresponds to the case of “low wealth” in Burkart and Ellingsen’s model.

Table 9			
Summary of the results of sub-sample analyses			
	Wealth (EBIT/ASSET)		
	Low (<0%) <i>(trade credit constraint applies)</i>	Medium (0–14%) <i>(no trade credit constraint applies)</i>	High (>14%) <i>(no trade credit constraint applies)</i>
In general <i>(where segmentation is solely based on EBIT/ASSET)</i>	strong complementarity*	complementarity	no correlation*
Constrained <i>(measured separately)</i>	strong complementarity*	complementarity* <i>(not robust)</i>	<i>(the category is not applicable to the model)</i>
Unconstrained <i>(measured separately)</i>	<i>(the category is not applicable to the model)</i>	<i>(the category is not applicable to the model)</i>	no correlation*
	Firm size		
	Micro	Small/Medium	Large
Constrained	complementarity	complementarity (not robust)	substitution
Unconstrained	no correlation	complementarity (not robust)	substitution

*Note: In the upper part of the table, items marked with * indicate results that are consistent with the theory.*

Table 9 summarises the relationship between bank credit and trade credit on a case-by-case basis.

In the context of the empirical analysis, a few methodological limitations also emerged, in relation to which improvements may be appropriate in future research. It would be beneficial to construct the Kaplan–Zingales, Whited–Wu and Cleary indices on Hungarian data, and classify Hungarian firms on that basis. This would presumably produce more consistent estimates compared to our current results. Notwithstanding that, literature relying on US data is also divided over classification procedures. It is even uncertain whether average profitability can adequately capture the wealth variable used in the theoretical model. Although in prior investigations we arrived at similar results in breakdowns by other variables, we do not believe that an indicator consisting of a single component could be a good separator. The outcome was somewhat improved when firm size was also taken into account. For that reason, it could be appropriate to use the Burkart and Ellingsen’s model to construct a trade credit constraint index on Hungarian data, which could be the Hungarian proxy for the model’s “wealth” variable.



The trends of trade credit are addressed in our fourth analysis, which relates to Equation (2). Here we investigate the average development of days payable outstanding over time for financially constrained and financially unconstrained firms, adjusted for the unique (sector-specific, operational, etc.) characteristics of firms. By analysing the time effect, we wish to capture typical individual (non-aggregated) responses to cyclical developments, i.e. a behavioural pattern. The results of the estimation are reported in *Table 10*, and the parts related to cyclicity are illustrated in *Figure 2*.

Table 10				
Estimation of the development of days payable outstanding over time				
AP_COST_SALES × 365	(1A)			
	Large	Medium	Small	Micro
AR_SALES × 365	0.40** (0.13)	0.44*** (0.12)	0.56*** (0.07)	0.34*** (0.03)
LOANS_ST_per_SALES	-19.17 (30.11)	170.07*** (47.69)	71.65*** (19.32)	33.26*** (10.05)
LOANS_LT_per_SALES	87.99 (136.89)	191.95* (86.87)	54.42 (31.44)	-3.46 (10.89)
YEAR2013	1.69 (3.93)	3.12 (2.65)	4.99*** (1.46)	5.06 (2.88)
YEAR2014	-2.79 (3.74)	4.19 (3.54)	6.83*** (1.78)	3.11 (3.53)
YEAR2015	1.92 (4.63)	-0.25 (4.35)	9.25*** (2.46)	3.67 (3.77)
D.KZ.constr	-9.07* (4.53)	-7.92 (6.35)	-7.43 (4.84)	12.21* (5.83)
D.KZ.unconstr	2.01 (4.74)	0.28 (3.41)	3.98 (2.44)	-17.42*** (3.73)
YEAR2013:D.KZ.constr	3.74 (5.66)	11.07 (6.56)	8.51 (4.64)	8.62 (5.42)
YEAR2014:D.KZ.constr	8.54 (5.64)	7.42 (4.84)	11.18* (4.65)	20.35** (6.83)
YEAR2015:D.KZ.constr	2.91 (7.66)	11.57 (7.02)	14.83* (6.94)	26.36** (9.17)
YEAR2013:D.KZ.unconstr	-3.73 (5.45)	-4.60 (3.55)	-2.58 (2.06)	1.67 (4.03)
YEAR2014:D.KZ.unconstr	4.96 (6.27)	-3.62 (3.46)	-4.77* (2.30)	3.62 (4.50)
YEAR2015:D.KZ.unconstr	-5.29 (5.90)	-2.31 (5.32)	-4.20 (2.83)	3.40 (4.82)
lag(LOGASSET)	-6.79 (4.12)	-20.08** (6.99)	-31.20*** (7.21)	-19.82*** (3.80)
EBIT_per_SALES	14.91 (10.99)	-23.91 (16.35)	-27.63* (10.97)	-3.42 (3.27)

Table 10
Estimation of the development of days payable outstanding over time

AP_COST_SALES × 365	(1A)			
	Large	Medium	Small	Micro
ROE	-6.79 (4.16)	5.70 (4.03)	-4.80 (4.18)	7.14* (2.84)
COVERAGE_RATIO	0.01 (0.01)	0.01 (0.01)	0.01 (0.00)	0.01* (0.00)
lag(LEVERAGE)	30.72 (18.66)	3.05 (25.04)	59.10*** (15.64)	14.91 (8.25)
lag(CURRENT)	0.03 (1.88)	0.37 (0.57)	0.19 (0.79)	0.60* (0.27)
lag(CASH_per_SALES)	292.14*** (85.81)	1.81 (13.33)	18.86 (9.76)	6.27** (2.34)
lag(CAPEX_per_SALES)	222.38* (86.39)	3.22 (40.23)	-11.66 (7.17)	-4.37 (3.46)
SALES_GROWTH	-10.26 (6.06)	-2.98 (3.58)	-3.48 (3.47)	-8.09*** (1.62)
DIVIDEND_RATIO	0.54 (0.41)	0.82 (0.73)	0.02 (0.01)	-0.00*** (0.00)
N	235	4,432	13,739	24,000
n	68	1,332	4,352	8,790
T	2012–2015	2012–2015	2012–2015	2012–2015
R ²	0.7289	0.3613	0.2706	0.0748

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; · $p < 0.1$; while character ‘·’ represents interaction between variables.

Although the estimation coefficients were not found to be statistically significant for large corporations and medium-sized enterprises, and thus the results are primarily applicable to small and microenterprises, all categories are represented. The diagrams are based on the average days payable outstanding in 2012 for unclassified firms. The diagrams show the differences relative to the base for each type.

Based on the estimated coefficients, days payable outstanding are found to have increased in financially constrained microenterprises. Between 2012 and 2013, a relatively moderate increase is also indicated for firms which are unconstrained according to their Kaplan–Zingales classification. The increase is also observed for constrained small enterprises, which implies pro-cyclical behaviour. By contrast, rather than an increase at a similar rate, stagnation is found for unconstrained firms.

6. Summary

This paper investigates the role of trade credit in Hungarian non-financial, non-utility firms that submit annual accounts prepared in Hungarian forints. In 2010–2015, firms initially faced difficulties in accessing finance and then benefited from the easing of terms. In this study, we sought to determine how firm's use of trade credit was related to their use of bank credit. For our empirical analysis, we relied on *Burkart and Ellingsen's (2004)* theory. Although it was only possible to partially confirm the theory, some findings of practical relevance were made beyond the scope of the theory on the basis of our data. On the entire sample, we found short-term loans and accounts payable to be complementary, the effect of which was stronger in firms that better exploited their capacity for bank debt. For firms with low own funds (equivalent to low wealth in Burkart and Ellingsen's model and a low EBIT/Sales ratio in the statistical analysis), we found strong complementarity, while no substitution could be demonstrated for highly profitable firms. We repeated the analysis in a breakdown by firm size, which indicated substitution in the category of large corporations. Although firm size cannot be considered as a proxy for wealth, we thought it important to document the differences in the other size categories as well. In the paper, primarily for illustration purposes, we also represented the trend of trade credit over time in a breakdown by size, and found that in the period of improving finance, the role of trade credit became more prominent in financially constrained firms, and remained level in unconstrained firms.

References

- Bálint, M. – Fellner, Z. (2016): *Hitelezési folyamatok – 2016. augusztus (Trends in Lending – August 2016)*. <https://www.mnb.hu/kiadvanyok/jelentesek/hitelezesi-folyamatok/hitelezesi-folyamatok-2016-augusztus>. Downloaded: 8 August 2017.
- Bálint, M. – Fellner, Z. (2017): *Hitelezési folyamatok – 2017. május (Trends in Lending – May 2017)*. <https://www.mnb.hu/kiadvanyok/jelentesek/hitelezesi-folyamatok/hitelezesi-folyamatok-2017-majus>. Downloaded: 8 August 2017.
- Biais, B. – Gollier, C. (1997): Trade credit and credit rationing. *The Review of Financial Studies*, 10(4): 903–937. <https://doi.org/10.1093/rfs/10.4.903>
- Bokor, Cs. – Fellner, Z. – Plajner, Á. (2014): *The use and expected impact of loans taken out in the Funding for Growth Scheme – results of a questionnaire-based survey*. In: MNB (2014): *The Funding for Growth Scheme – The first 18 months*, MNB. <https://www.mnb.hu/letoltes/nhp-tanulmanykotet-eng-b5-digit.pdf>. Downloaded: 8 August 2017.

- Bougheas, S. – Mateut, S. – Mizen, P. (2007): *The inventory channel of trade credit: Theory and Evidence*. Working Paper. Department of Economics, University of Sheffield. ISSN 1749–8368.
- Brennan, M. – Maksimovic, V. – Zechner, J. (1988): *Vendor financing*. The Journal of Finance, 43(5): 1127–1141. <https://doi.org/10.1111/j.1540-6261.1988.tb03960.x>
- Burkart, M. – Ellingsen, T. (2004): *In-kind finance: a theory of trade credit*. American Economic Review, 94(3): 569–590. <https://doi.org/10.1257/0002828041464579>
- Chittenden, F. – Bragg, R. (1997): *Trade credit, cash-flow and SMEs in the UK, Germany and France*. International Small Business Journal, 16(1): 22–35. <https://doi.org/10.1177/0266242697161002>
- Cleary, S. (1999): *The relationship between firm investment and financial status*. The Journal of Finance, 54(2): 673–692. <https://doi.org/10.1111/0022-1082.00121>
- Csubák, T.K. – Fejes, J. (2014): *A magyar kkv-k 21. századi banki finanszírozásának áttekintése és kiútkeresés a hitelválság csapdájából (Overview of bank financing for Hungarian SMEs in the 21st century, and the search for and escape from the debt trap)*. Financial and Economic Review, 13(2): 174–194.
- Cunningham, R.M. (2005): *Trade credit and credit rationing in Canadian firms*. Bank of Canada Working Paper, No. 2004–49. <https://doi.org/10.2139/ssrn.643023>
- Elsas, R. – Klepsch, C. (2016): *How and when do firms adjust their investments toward targets?* <https://dx.doi.org/10.2139/ssrn.2775091>
- Endrész, M. – Harasztosi, P. – Lieli, R. (2015): *The Impact of the Magyar Nemzeti Bank's Funding for Growth Scheme on Firm Level Investment*. MNB Working Papers 2. <https://www.mnb.hu/letoltes/ehl-wp-2015-final-1.pdf>. Downloaded: 8 August 2017.
- Emery, G.W. (1984): *A pure financial explanation for trade credit*. Journal of Financial and Quantitative Analysis, 19(3): 271–285. <https://doi.org/10.2307/2331090>
- Fazzari, S.M. – Petersen, B.C. (1993): *Working capital and fixed investment: new evidence on financing constraints*. The RAND Journal of Economics, 24(3): 328–342. <http://www.jstor.org/stable/2555961>
- Farre-Mensa, J. – Ljungqvist, A. (2016): *Do measures of financial constraints measure financial constraints?* The Review of Financial Studies, 29(2): 271–308. <https://doi.org/10.1093/rfs/hhv052>
- Ferrando, A. – Mulier, K. (2013): *Do firms use the trade credit channel to manage growth?* Journal of Banking & Finance, 37(8): 3035–3046. <https://doi.org/10.1016/j.jbankfin.2013.02.013>

- Ferris, J.S. (1981): *A transactions theory of trade credit use*. The Quarterly Journal of Economics, 96(2): 243–270. <https://doi.org/10.2307/1882390>
- Freixas, X. (1993): *Short-term credit versus account receivable financing*. Economics Working Paper 27. Department of Economics and Business, Universitat Pompeu Fabra. <https://ideas.repec.org/p/upf/upfgen/27.html>
- García-Teruel, P.J. – Martínez-Solano, P. (2010): *Determinants of trade credit: A comparative study of European SMEs*. International Small Business Journal, 28(3): 215–233. <https://doi.org/10.1177/0266242609360603>
- Juhász, P. – Reszegi, L. (2017): *Gátak a magyar vállalati növekedésben (Barriers to Hungarian Corporate Growth)*. Vezetéstudomány, 48(6–7): 27–38.
- Kaplan, S.N. – Zingales, L. (1997): *Do investment-cash flow sensitivities provide useful measures of financing constraints?* The Quarterly Journal of Economics, 112(1): 169–215. <https://doi.org/10.1162/003355397555163>
- Kling, G. – Paul, S.Y. – Gonis, E. (2014): *Cash holding, trade credit and access to short-term bank finance*. International Review of Financial Analysis, Vol. 32(March): 123–131. <https://doi.org/10.1016/j.irfa.2014.01.013>
- MNB (2015): *Time series of credit institutions limited by shares (ended 31.12.2014)*. <https://www.mnb.hu/felugyelet/idosorok/i-penz-es-hitelpiaci-szervezetek>. Downloaded: 7 August 2017.
- MNB (2017): *Financial Stability Report (May 2017)*. <https://www.mnb.hu/en/publications/reports/financial-stability-report/financial-stability-report-may-2017>. Downloaded: 8 August 2017.
- Módos, D. – Bokor, Cs. – Hidasi, B. (2014): *Main features of FGS loans and companies borrowing under the scheme*. In: MNB (2014): *The Funding for Growth Scheme – The first 18 months*, MNB. <https://www.mnb.hu/letoltes/nhp-tanulmánykotet-eng-b5-digit.pdf>. Downloaded: 8 August 2017.
- Nábelek, F. (2016): *Késedelmes fizetés a magyar vállalkozások körében (Overdue payment among Hungarian firms)*. Institute for Economic and Enterprise Research. http://gvi.hu/kutatas/467/kesedelmes_fizetes_a_magyar_vallalkozasok_koreben
- Nilsen, J.H. (2002). *Trade credit and the bank lending channel*. Journal of Money, Credit, and Banking, 34(1): 226–253. <https://doi.org/10.1353/mcb.2002.0032>
- Nyíró, Zs. – Hajdu, M. (2015): *A GVI 2015. januári KKV Körkép felvételének eredményei (Evidence from the GVI SME Survey of January 2015)*. Institute for Economic and Enterprise Research. http://gvi.hu/kutatas/398/kkv_korkep_2015_januar__javulo_uzleti_varakozasok_a_kkv_szektorban

- Petersen, M.A. – Rajan, R.G. (1997): *Trade credit: theories and evidence*. Review of Financial Studies, 10(3): 661–661. <https://doi.org/10.1093/rfs/10.3.661>
- SAFE (2011): *2011 SMEs' access to finance*. Survey. Analytical Report. European Commission, Enterprise and Industry. <http://ec.europa.eu/DocsRoom/documents/7922/attachments/1/translations/en/renditions/native>. Downloaded: 8 August 2017.
- SAFE (2013): *2013 SMEs' access to finance*. Survey. Analytical Report. European Commission, Enterprise and Industry. <http://ec.europa.eu/docsroom/documents/7864/attachments/1/translations/en/renditions/native>. Downloaded: 8 August 2017.
- SAFE (2014): *Survey on the access to finance of small and medium-sized enterprises in the euro area, October 2013 to March 2014*. European Central Bank. <https://www.ecb.europa.eu/pub/pdf/other/accesstofinancesmallmediumsizedenterprises201404en.pdf>. Downloaded: 8 August 2017.
- SAFE (2015): *Survey on the access to finance of enterprises (SAFE)*. Analytical Report 2015. European Commission. <http://ec.europa.eu/DocsRoom/documents/14321/attachments/1/translations/en/renditions/native>. Downloaded: 8 August 2017.
- Schwartz, R.A. (1974): *An economic model of trade credit*. Journal of Financial and Quantitative Analysis, 9(4): 643–657. <https://doi.org/10.2307/2329765>
- Smith, J.K. (1987): *Trade credit and informational asymmetry*. The Journal of Finance, 42(4): 863–872. <https://doi.org/10.1111/j.1540-6261.1987.tb03916.x>
- Szűcs, N. (2008): *Lánctartozás a kereskedelmi hitelezés irodalmának tükrében (Circular debt as reported in the literature on trade credit)*. Financial and Economic Review, 7(4): 399–417.
- Whited, T.M. – Wu, G. (2006): *Financial constraints risk*. Review of Financial Studies, 19(2): 531–559. <https://doi.org/10.1093/rfs/hhj012>
- Ying, Q. – Guo, W. – Yang, T. (2014): *The substitution effect of trade credit financing in an emerging economy: empirical evidence from China*. In: Xu, J. – Cruz-Machado, V. – Lev, B. – Nickel, S. (eds.): Proceedings of the Eighth International Conference on Management Science and Engineering Management. Advances in Intelligent Systems and Computing, vol. 280, Springer, Berlin, Heidelberg, pp. 551–561. http://doi.org/10.1007/978-3-642-55182-6_48

Explanations of Asset Price Bubbles

Gergely Lakos – Tibor Szendrei

The study provides an overview of the theories explaining asset price bubbles found in the literature. Six different explanations are identified that can cause asset price bubbles on their own as well as in some combination: informational frictions, short selling constraints, limited liability of financial intermediaries, herd behaviour among market participants, perverse incentives in the provision of information, and market participants' bounded rationality. Mispricing arising from asset price bubbles can considerably hamper the efficient allocation of resources, which may warrant government intervention. Nevertheless, most of the underlying factors that induce asset price bubbles can only be limited or offset rather than fully eliminated. Financial stability is mainly threatened by debt-financed asset price bubbles. With respect to asset price bubbles, the mandate and available instruments of macroprudential policy are limited to the mitigation of cyclical fluctuations in the related lending.

Journal of Economic Literature (JEL) codes: G10, G12, R31

Keywords: asset price bubble, excessive lending, bounded rationality, financial stability, macroprudential policy

1. Introduction

Bubbles develop in the price of an asset when the asset is substantially overvalued for a relatively long time. Such bubbles can emerge in the price of various assets.¹ Significant mispricing distorts the allocation of resources, goods and services, resulting in considerable economic losses. In this context, the impact of asset price bubbles which threaten financial stability is the main concern for macroprudential authorities. This is because the build-up of real and financial asset price bubbles may lead to systemic risk in the financial system. When these risks materialise, there is a high probability that the disruptions in the financial intermediary system will considerably damage the overall economy.

*Lakos Gergely is a lead economist at the Magyar Nemzeti Bank. E-mail: lakosg@mnbb.hu
Tibor Szendrei is an economist at the Magyar Nemzeti Bank. E-mail: szendreit@mnbb.hu*

The Hungarian manuscript was received on 24 March 2017.

DOI: <http://doi.org/10.25201/FER.16.4.122150>

¹ A comprehensive account of major historical examples can be found, for example, in *Kindleberger and Aliber (2011)*.

The main systemic risks consist of excessive lending risks linked to asset price bubbles. In a developed asset price bubble, economic actors mainly purchase the overvalued asset on credit. During price corrections, the solvency of debtors diminishes, since selling the assets only yields a fraction of the expected income. In addition, the assets concerned are usually the collateral for the loans. Therefore, decreasing asset prices also lower the value of potential collateral pledged when borrowing. Both effects considerably dampen credit demand. Furthermore, asset price corrections also affect lenders negatively, as they incur heavy losses due to part of their loan portfolio becoming non-performing. One method of stabilising their weakened capital position is to curb lending. All in all, when debt-financed asset price bubbles burst, both credit demand and credit supply may contract considerably. This may even impair the realisation of companies' efficient investment plans as well as households' ability to temporally smooth consumption.

Macroprudential policy can only be effective if we have a proper understanding of the reasons behind the emergence of asset price bubbles. The primary aim of this study is to present the factors identified in the literature. In the case of perfectly rational actors interacting on a perfect market, overvaluation of assets cannot be sustained, since in such a scenario holding the asset yields much lower profits than the cost of purchasing or what can be received by selling it. Therefore, the explanations of asset price bubbles hinge on the presence of investors exhibiting bounded rationality or some imperfection in the market allocation mechanism. In line with the literature, we distinguish five main groups of the latter: informational frictions, short selling constraints, limited liability of financial intermediaries, herd behaviour among market participants, and perverse incentives in the provision of information.

Chapter 2 of the study presents a more accurate definition of asset price bubbles, a key element of which is the definition of the fundamental value. Due to the difficulty in observing the fundamental value, the identification of asset price bubbles in practice poses challenges even in hindsight. This observability predicament is partly the reason why asset price bubbles can emerge. *Chapter 3* presents the basic features of the different phases in the asset price cycle. An asset price bubble emerges in the so-called boom phase, and develops during the so-called euphoria period. *Chapter 4* gives an overview of the theoretical explanations of asset price bubbles, grouped by explanatory factors. Through the short description of a few relevant theoretical models, the paper illustrates how the different reasons can create asset price bubbles. Since asset price bubbles mainly threaten financial stability when coupled with excessive lending, *Chapter 5* briefly discusses why and how the two phenomena can strengthen each other. Finally, the paper summarises the major lessons for state intervention – with a focus on macroprudential policy – that can be drawn from the explanations behind debt-financed asset price bubbles.

2. Definition of asset price bubbles

Financial or real assets are overvalued when their market price exceeds the expected present value of the future cash flows and other benefits derived from holding the asset for an extended period of time. The latter is also referred to as the fundamental value and is defined in mathematical form as follows:²

$$F_t = \sum_{\tau=t+1}^{\infty} \frac{E_t(d_\tau)}{1+r_{t,\tau}} \quad (1)$$

The fundamental value of the asset in the t^{th} period is F_t . d_τ denotes the monetary value of the profits realised in the τ^{th} period, which is uncertain in the t^{th} period. $r_{t,\tau}$ denotes the expected return on investments in general from the t^{th} period to the τ^{th} period, which also includes the risk premium compensating for the uncertainty of investments. E_t is the expected value operator derived from the information available in the t^{th} period.

Substantial and persistent overvaluation of an asset is commonly referred to as an asset price bubble.³ The extent of the price bubble is the difference between the market price and the current fundamental value. Since the fundamental value is the value of uncertain future returns discounted to the present, the fundamental value cannot be observed directly. Identifying exactly what information pertaining to future returns is considered available in the present is already difficult. This is because the various economic actors have different amounts of information with varying degrees of accuracy. This is also influenced by the economic actors' effort to acquire relevant information. It is similarly difficult to observe the extent of uncertainty surrounding potential future returns and to gauge the realistic risk premium expected in exchange for this uncertainty. Due to the above, even with the benefit of hindsight, in practice it is hard to establish whether a bubble emerged in an asset's price or to determine the size of a bubble.

Nevertheless, in several cases, overvaluation or asset price bubbles can be identified with great certainty. One such example is from the study by *Huberman and Regev (2001)*, where the authors examine the price developments of the

² This formula is equivalent to the general pricing formula of the asset pricing theory:

$$F_t = E_t \left(\sum_{\tau=t+1}^{\infty} m_{t,\tau} d_\tau \right)$$

where $m_{t,\tau}$ is the so-called stochastic discount factor (see Chapter 1 of *Cochrane (2005)*, especially pages 24–25).

³ The definition of asset price bubbles varies slightly. The definition used in this study is the dominant variety (see *Brunnermeier 2008*, *Brunnermeier – Oehmke 2013*, *Hirshleifer 2001*, *Scherbina 2013* and *Xiong 2013*), although this is not accurate either, since it only roughly distinguishes asset price bubbles from overvaluation. As we detail later, the inaccuracy is due to the fact that these are two very similar phenomena, which primarily differ in their magnitude.

shares of EntreMed,⁴ a biotech firm, on the market. On 3 May 1998, *The New York Times* published an article on the development of cancer-treatment medicine, mentioning the name EntreMed several times. In the following weeks, the share price of the company soared from USD 12 to USD 52, in parallel with a steep rise in the volume of trading in the shares. However, EntreMed's research results had already been published in *Nature* magazine in November 1997, which had also boosted EntreMed's share price, albeit by far less than in May the following year. It is reasonable to claim that the fundamental value of EntreMed shares rose only in November, and therefore a major part of the price increase in May contributed to the emergence of an asset price bubble.

3. Phases of the asset price cycle

The starting point in the explanation of asset price bubbles was established by *Minsky (1992)*, and was later detailed by *Kindleberger and Aliber (2011)*. The study divides the asset price cycle into six⁵ phases: *displacement, boom, euphoria, profit taking, Minsky Moment, and panic (Figure 1)*. The theoretical models of asset price bubbles seek to explain this comprehensive asset price cycle. Despite the considerable progress in recent decades, no uniform theory has been devised that is able to accurately and simultaneously capture all the phases of the asset price cycle described by Minsky. In this chapter, Minsky's intuitive characterisation will be reviewed, while the next chapter focuses on describing market imperfections that can cause asset price bubbles to build up and persist.

Displacement is the phase in the life of the asset price cycle when the fundamental value suddenly rises on account of some factor (*Figure 2*). In the case of stocks, this can occur, for example, when a listed firm develops or uses a more efficient technology, successfully ends a legal dispute, or acquires another company that it can make more efficient through reorganisation. In the case of properties, the fundamental value may rise, for example, when larger employers appear in a given geographical region, which raises the future rents for residential and commercial properties. The asset's market transitions from displacement to the boom phase endogenously.

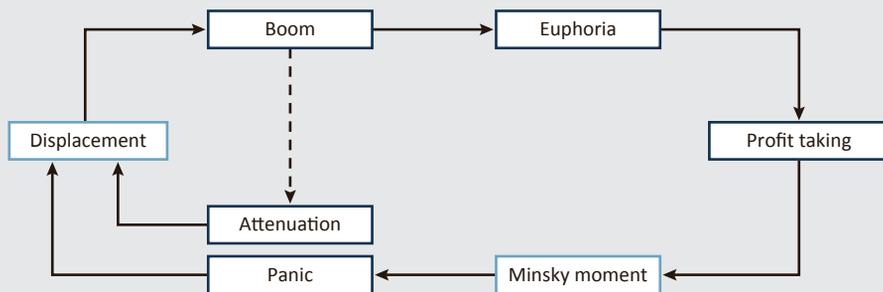
In the *boom* phase, the demand for the now undervalued asset rises, lifting the market price and moderating undervaluation. In this period, the economic environment is usually stable; the volatility of asset prices is low, and is coupled with increasing lending and investments. During this stage, the expansion in lending mostly finances investments that seek to utilise the increased efficiency, thus

⁴ Today, the company is called CASI Pharmaceuticals.

⁵ Some divide Minsky's financial instability hypothesis into five phases (see: *Brunnermeier and Oehmke 2012*), as the Minsky Moment merely refers to a point in time.

boosting the asset's fundamental value. In the case of real assets, this manifests in the production of more real assets, for example the construction of properties. In the case of financial assets, the expansion of activities that ensure future cash flows may be targeted. In the context of stocks, this can take the form of increasing the market share of a listed company using new technology. Often the change that raises the fundamental value is a novelty, the future impact of which is difficult to evaluate, leading to the conundrum of estimating the asset's new fundamental value. Therefore, the asset may be inaccurately priced in the boom phase, making the undervalued asset overvalued.

Figure 1
Schematic illustration of the asset price cycle



Note: According to the model, the light blue phases of the asset price cycle occur due to exogenous effects.

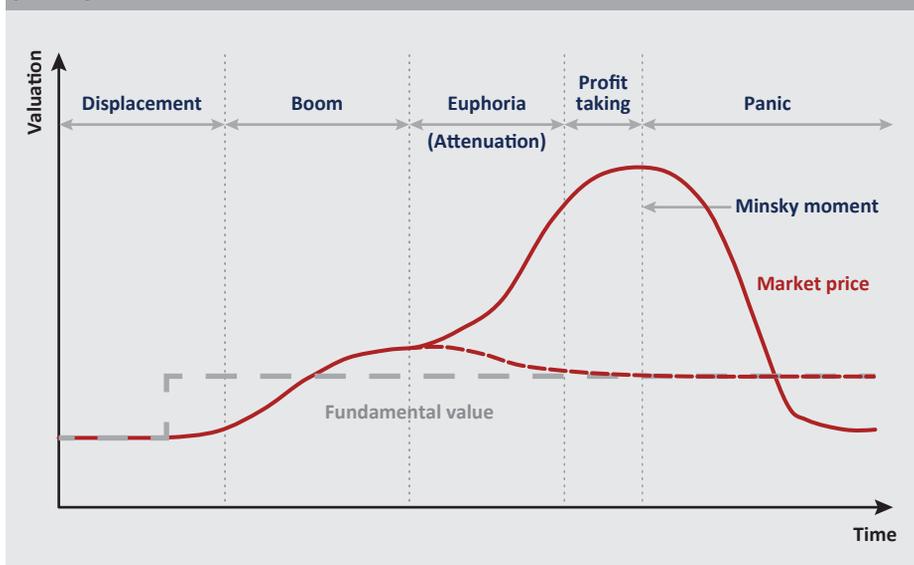
In the *euphoria* phase, the overvaluation of the asset increases and persists for a longer period, i.e. the asset price bubble emerges during this stage. The asset price increases steeply, coupled with large price volatility and a high number of transactions. Based on the accelerating price dynamics or due to the fundamental value that is becoming increasingly easy to identify, more and more market participants may realise that an asset price bubble has built up. However, they may have various reasons for holding the asset for some time, thereby sustaining the asset price bubble.

During the euphoria phase, new types of investors appear on the market of the assets. To use *Keynes's* (1936) terms, professional investors are accompanied by swarms of speculators. The first group was already present on the market of the asset during the boom, and invested in the asset to secure the higher future payments from holding the asset. However, speculators principally purchase the asset in the hope of the expected profit generated from selling the asset later, at a higher price. Demand for the asset is further boosted by the fact that lending may now be used to finance the purchase of the asset, in addition to financing the underlying investments. In the stock example, loans are taken out not only by the companies with increased efficiency to expand their activities, but also by a myriad

of leveraged investors who purchase shares in the company on credit. The supply of the asset can only catch up to the rapidly expanding demand with a considerable lag, which exacerbates asset price bubbles. One typical example is the world of real estate, where renovations and construction take more time.

It should be noted that the boom phase is not necessarily followed by euphoria, and even if it is, it is not always easy to identify when this happens, even in hindsight. If it becomes clear in time that the asset is slightly overvalued, an investment frenzy is unlikely to occur which would be necessary for the boom period to transition into euphoria. However, if information on the new fundamental value disseminates slower, the probability of the emergence of an asset price bubble is greater; therefore, it is also harder to establish the exact onset of the euphoria phase. Certain characteristics of the asset price bubbles are easier to observe, e.g. the rapid rise in market prices, the related excessive lending or the ensuing distortion in resource allocation. On account of this, the empirical studies on asset price bubbles often identify the bubbles with the help of such variables.

Figure 2
Development of the market price and the fundamental value in a schematic asset price cycle



The final stage of euphoria marks the start of *profit taking*. In this phase, so-called sophisticated market participants, who recognised the asset price bubble earlier, sell the asset, while others do not start selling in large numbers. New actors, mainly speculators, continue to enter the market, who proceed to buy the assets from the sophisticated market participants. Thus, the sophisticated players get a high price for their assets, achieving considerable gains. This is when trading in the asset becomes

analogous to a Ponzi scheme, since the price is maintained by less sophisticated participants who will probably sell their newly acquired assets at a loss. In practice, isolating the phases of euphoria and profit taking is even harder than delineating the boom and euphoria phases. This is because the withdrawal of capital cannot be accurately observed, and because distinguishing sophisticated market participants from unsophisticated players is even more difficult than identifying fundamental traders and speculators.

The *Minsky Moment* is an event that impacts market participants' expectations about the future development of the asset price simultaneously, en masse and considerably. As market participants revise their price expectations downwards, *panic* ensues on the asset's market: market participants seek to sell the asset in large numbers, and the price nosedives, often at a pace exceeding that of the earlier price increase. The intention of selling as soon as possible is rational at the individual level, since the expected price is continuously decreasing, but collectively it results in a suboptimal situation. The resulting price correction is often too large due to the panic, making the asset undervalued. Many are unable to repay the loans amassed in the euphoria phase at the pace stipulated in the loan contract, leading to mass debt problems. The bigger the asset price bubble created by excessive lending in the euphoria phase, the greater the drop in prices, the undervaluation, and the extent of the debt problems.

4. Reasons behind asset price bubbles

The theoretical literature has identified six factors that can cause and maintain asset price bubbles in certain combinations: informational frictions, short selling constraints, limited liability of financial intermediaries, herd behaviour among market participants, perverse incentives in the provision of information, and bounded rationality of market participants (*Table 1*). This chapter presents these in more detail, through the brief introduction of some relevant models. The chapter does not aim to provide a comprehensive review of the literature; it merely undertakes to show examples of how asset price bubbles are formed, using a selection of studies.⁶

4.1. Informational frictions

One necessary condition for the price correction of overvalued assets is that enough information about the fundamental values is available on the markets of the assets in a timely manner. The factors hampering the spread of information are called

⁶ Brunnermeier and Oehmke (2013), Hirshleifer (2001), Scherbina (2013) and Xiong (2013) give a comprehensive overview about the explanations behind asset price bubbles from different aspects.

informational frictions, due to which market participants' beliefs can be persistently disparate.⁷

In the model of *Abreu and Brunnermeier (2003)*, the informational friction is embodied in the phenomenon called *sequential awareness*. In this framework, the traders do not simultaneously realise when the asset price exceeds the fundamental value, and they do not know when other traders become aware of this. Therefore, the development of the asset price bubble never becomes common knowledge. The other important assumption of the model is that no trader is “big” enough to make a considerable impact on prices with their intention to sell. Thus, for the market price to drop a critical mass of the traders have to recognise that there is an asset price bubble simultaneously and act accordingly by selling at the same time. Since traders forego the extra profits derived from selling later at a higher price by selling too early, the traders who are aware of the asset price bubble are unwilling to part with the asset until this critical mass comes together. However, the requirement of selling together introduces a coordination problem. In such a scenario, all readily observable and perhaps even irrelevant news can cause a selling wave, which entails a rapid correction of the prices and uncertainty regarding the time of the correction. This conclusion is supported by the empirical studies of *Cutler et al. (1989)* and *Fair (2002)*, where the authors find that most large price movements on the stock exchange did not happen as a result of news pertaining to fundamental values.

Explanatory factor	Content
Informational frictions	Beliefs regarding the fundamental value converge to the real figure too slowly.
Short selling constraints	The option of short selling is limited.
Limited liability	Financial intermediaries are required to absorb the losses only to the extent of their capital.
Herd behaviour	Market participants excessively imitate each other's market behaviour.
Perverse incentives in the provision of information	There is a conflict of interest for certain organisations tasked with assessing products and market participants.
Bounded rationality	Market participants are not fully rational.

⁷ The extent of persistently different beliefs is difficult to measure directly, however, in line with, for example, the questionnaire-based survey of *Mankiw et al. (2004)* on the different inflation expectations, we can argue that this phenomenon can truly be significant. Persistently different beliefs can result not only from informational frictions but also from the bounded rationality of market participants. The role of bounded rationality is discussed in Chapter 4.6. One comprehensive overview of the literature on the role of heterogeneous beliefs in the explanations of asset price bubbles can be found in *Xiong (2013)*.

Another type of informational friction is caused by *heterogeneous priors*. If the initial beliefs regarding the asset's fundamental value differ among traders, they also assess obtainable new information differently. Their beliefs about the fundamental value may differ even if they share all their information with each other. Therefore, even if news on the fundamental value spreads fast on the market, the exaggerated optimism of certain traders only wears off slowly.

With heterogeneous priors, the market price may even exceed the assessment of the most optimistic traders regarding the fundamental value. This is illustrated by the following example which is based on the model by *Harrison and Kreps (1978)*. Let us assume that two types of traders are active on the market for an asset. At the onset of the trading process, all of them believe that the same fundamental value is likely. Next, news about the fundamental value arrives in several waves, until it becomes fully observable. Traders in one group trust their initial beliefs and thus they adjust their beliefs less in response to the new information. In this example, the initial market price of the asset will be higher than the common assessment of the traders. This is because both types of traders can expect lucrative future selling opportunities. If in the subsequent periods good news on the fundamental value arrives, more adaptive traders are willing to purchase the asset even at a higher price than the initial assessment because of the good news. If bad news on the fundamental value arrives, less adaptive traders will still be willing to purchase the asset, even at slightly lower prices than the initial assessment. Accordingly, both types of traders can sell the asset to the other type at a relatively high price precisely at the time when they wish to sell. This mutually beneficial resell option lifts the market price of the asset above the initial assessments.

4.2. Short selling constraints

The prospect of pouncing on the arbitrage inherent in a mispriced asset is also hampered by limiting the possibility of short selling. Short selling refers to the immediate selling of a borrowed asset and its return after its future purchase. This makes the selling of an asset available to those who currently do not own enough of it. This trading technique is important because accurate information on the fundamental value does not necessarily reach the market participants who currently hold the given asset in large volumes.

There is often no option for short selling because the market for the given asset is not standardised enough. One typical example is the property market. In other cases, administrative barriers hinder its use. Restrictive measures are usually implemented to avoid a panic due to simultaneous, large-scale short selling, which would make prices fall excessively and rapidly. *Bris et al. (2007)* analysed the regulations on and the practice of short selling in securities markets of 46 countries between 1990 and 2001, and found that short selling was prohibited in 10 countries during the period reviewed. At the end of the period, it was permitted in 35 countries, but it was

common practice in only 25. The lower number of countries where short selling is common practice can be attributed to the fact that in certain countries short selling was hampered by a considerable tax burden.

In *Miller's (1977)* model, the asset price reflects the beliefs of optimistic traders due to short selling constraints, which makes the asset overvalued. The buying intentions of optimistic traders are reflected in the market price without restrictions, in contrast to the selling intentions of pessimistic traders. *Chen et al. (2002)* expanded the model and formulated theoretical predictions regarding the impact of short selling on stock prices, which were tested empirically. The study approximated the effectiveness of short selling constraints through the number of shareholders. In the case of fewer owners, it is more likely that market actors who believe the fundamental value to be lower will hold shares in the company in small proportion; therefore, the limits on short selling can take hold more effectively. The authors find that the stocks held by fewer owners yielded lower returns later on, which is consistent with their theoretical conclusion that the limits on short selling make these stocks more overvalued.

4.3. Limited liability

Trading in overvalued assets often involves financial intermediaries managing others' savings. Savers usually trust them with their money because financial intermediaries are better at picking good investments. Financial intermediaries can use this information advantage partly to their own advantage and to the detriment of their clients. This practice can be further strengthened by the fact that financial intermediaries share the returns generated by the investments with their clients, but since their liability is limited, they do not have to incur losses in excess of their own capital.

In the model of *Allen and Gale (2000)*, financial intermediaries use debt financing for their activities, i.e. in exchange for these funds they promise interest that is largely independent from their profitability. In such a scenario, it is worthwhile for financial intermediaries to invest in risky assets with high expected returns, since they can retain a larger portion of the potential profit, while they are not obligated to repay their debt in full in the case of losses due to their limited liability. Since lenders are at an information disadvantage compared to the financial intermediaries, they cannot appropriately limit this risk shifting. Excessive risk appetite can generate excessive demand for certain assets, which can lead to the evolution of asset price bubbles.

In the model of *Allen and Gorton (1993)*, there are two types of financial intermediaries both with limited liability but with different levels of sophistication. In the model, the asset price bubble is generated by the less sophisticated financial intermediaries. The difference between the two types of financial intermediaries

is similar to the distinction made by *Keynes (1936)* mentioned earlier: more sophisticated intermediaries are able to determine the fundamental value of the assets, while less sophisticated ones cannot. Savers cannot directly observe the quality of financial intermediaries. They pay the financial intermediaries by ceding a fixed share of the money entrusted to them, and in the case of losses, they do not pay the intermediaries anything but incur all the losses. If the less sophisticated financial intermediaries are not active enough on the market, they will reveal to their clients that they are unable to pick the undervalued assets, which could mean the end of their contract. They seek to avoid this, which encourages active behaviour on the market. They do not have to fear buying overvalued assets, since in the event of a potential price correction, their limited liability shields them from large losses anyway.

4.4. Herd behaviour

Herd behaviour is when economic actors imitate others' market behaviour too much, in the belief that they will be better off as a result. This process entails wasting the information dispersed among the individual economic actors. In this case, even if the majority believe that an asset price bubble has emerged, they are prone to buying the given asset in line with the trend on the market, thereby contributing to the persistence and potentially further inflating the asset price bubble. Several factors may motivate herd behaviour. *Bikhchandani and Sharma (2000)* distinguish three types of herd behaviour on financial markets (*Table 2*): information-based herding, reputation-based herding and compensation-based herding.

Explanatory factor	Content
Information-based herding	Investors imitate the investment decisions of others because they believe that the others can better identify good investments.
Reputation-based herding	Investors imitate the investment decisions of others because this makes them competent in the eyes of their (potential) employers.
Compensation-based herding	Investors imitate the investment decisions of others because their compensation depends partly on their performance relative to others.

In the case of *information-based herding*, all market participants have information regarding the potential investment decisions that only they can observe (so-called private information). Market participants do not have the time, expertise and other resources available to examine all of the investment opportunities in depth themselves. Therefore, nobody is perfectly informed. Furthermore, the private information obtained by the various actors can be dissimilar. In such a situation, market participants aim to gain an insight into others' private information through their observable behaviour on the market so that they can make the best investment decision.

In the basic models describing information-based herding (*Banerjee 1992; Bikhchandani et al. 1992; Welch 1992*), assets are traded sequentially, i.e. investors acting later can observe the earlier investment decisions of the other market participants before making their own investment decision. All investors can purchase the same asset at the same prices. Players making the decisions earlier, who engage in similar investments, may easily induce herd behaviour. This is because all investors acting later can deduce that there are many earlier investors whose private information is similar to each other, which can convince them to copy earlier decisions, even if their own private information is contrary to this. In such a scenario, market participants exert negative externalities on those making decisions later. Therefore, the emerging market practice will appropriately reflect the information of only those market participants who make decisions earlier and will not include the private information of those who make decisions later. Investments can be concentrated in certain assets even if the private information of the majority does not substantiate this investment strategy.

Due to the constant asset price assumed in basic models, such models are not suitable for explaining asset price bubbles. However, *Avery and Zemsky (1998)* have relaxed this assumption. In their model, they examine two types of investors, who can trade in only one asset. Of the two types of investors one has accurate private information regarding the fundamental value of the asset, while the other has access only to inaccurate private information. All investors know their own type, but they do not know exactly what the proportion of well-informed investors to poorly informed investors is on the market. The investment decisions are made sequentially, thus information-based herding can emerge in line with the logic of the basic model. In this model, poorly informed investors imitate the observed earlier investment decisions because they assume that the decisions were made by mostly well-informed investors, even if that was not the case. In such a scenario, the asset price on the market becomes detached from the fundamental value, which may lead to an asset price bubble.

The basic models of *reputation-based herding* rest on the assumption that financial experts differ in terms of expertise, experience, diligence, and other qualitative aspects, which cannot be directly observed by their employers. In such situations, the employees seek to perform their duties in a way that their observable performance convinces their current and potential employers that they are better than the others, since this creates better career opportunities for them. However, in certain cases professional reputation can only be improved by imitating other financial experts' investment decisions even if the employees believe that these decisions are not optimal for their employers (*Scharfstein – Stein 1990; Trueman 1994; Zweibel 1995; Prendergast – Stole 1996; Graham 1999*).

In the model of *Scharfstein and Stein (1990)*, the employed investors all have some private information about the good investment opportunities, independently of each other. The authors distinguish two types of investors: better investors have more accurate private information than bad ones. Each investor has a different employer, but neither investors nor the employers know which investors belong to which category. However, employers can observe the decisions made by all the employees (not only their own), and the profits generated. In such a situation, employers should remunerate their employed investors based not only on the final performance (profits), but also based on how much their decisions differ from other investors' decisions. This is because employers know that the private information of better investors differs less than that of poor investors. If all investors followed their private information, worse investors would be more likely to make decisions deviating from market trends, which would be indicative of their poor quality. Investors strive to avoid this, even if based on their private information they risk making an investment with low returns. Therefore, in equilibrium, investors making sequential decisions can imitate earlier investment decisions, which may generate excessive demand for assets with a lower fundamental value, thereby potentially contributing to the build-up of asset price bubbles.

Zwiebel's (1995) model differs from the previous one in three major assumptions. First, employed investors can be of more than two qualities, and what is even more important, they are aware of their own type. Second, employers cannot observe the quality of the investment decisions made by their employees, only the profits generated from the investments made by their own investors and those employed by others. Third, in this model being a better quality investor does not entail that their private information about investment opportunities is more accurate, instead it merely increases the profitability of all chosen investments deterministically and to the same extent. The large group of average investors making their decisions simultaneously pick the usual investment opportunities, which are in line with earlier market practices, and which are not necessarily the most efficient, because if many investors do so, when employers compare the profits, they can identify

the relatively good investors quite accurately. Only few investors (the best and the worst) choose investment opportunities diverging from the market practices, therefore their relative performance is more uncertain *ex ante*. The best trust their excellence, while the worst need to take this risk, since they can only avoid their employers finding out about their poor quality in the event they make a fortuitous investment.

The latter model is a transitional type on the road towards the third group of models, comprising of *compensation-based herding* models (see, for example, *Maug – Naik 1996; Admati – Pfleiderer 1997*). In these models, herd behaviour ensues because the remuneration of the employed investors also depends on the observable performance of other, similar investors, i.e. employers use a relative performance assessment system.

In the model of *Maug and Naik (1996)*, employers tie the remuneration of the employed investors to a benchmark on the market, i.e. the more the investors outperform the benchmark, the more their compensation grows. In such a case, imitating the practice widespread on the market can be profitable for risk-averse investors because it reduces the possibility of their performance falling short of the benchmark on the market. This can occur even if based on the private information of a market investor, another investment strategy is expected to yield more profits. Just as in the model of the authors, employers usually use relative performance assessment because they can urge their employees to make greater efforts (tackle moral hazard), and because it helps them select the best candidates (tackle adverse selection), whom they can task with more complex duties. This is because relative performance assessment can be used to pit the individual investors against each other, and better investors exerting more effort outperform others. In other words, employers should use relative performance assessment even if it causes herd behaviour generating poorer investments as an adverse side effect.

4.5. Perverse incentives in the provision of information

One group of economic actors are tasked with providing information on certain products and market participants. This is performed, for example, by credit rating agencies, audit companies and financial advisers. If these organisations do not work adequately, the information necessary for the correction of asset price bubbles can spread only slowly. Certain *perverse incentives* can partly divert the behaviour of economic actors from this basic information provision goal (*Scherbina 2013*).

According to the study cited, giving a better assessment than warranted about certain firms is worthwhile for financial analysts for various reasons. First, in order to preserve the option of obtaining appropriate information in the future, financial analysts are averse to infuriating the companies under review with negative assessments. Second, the financial analysts of investment banks profit

from generating transactions for the trading department with a positive stock assessment. Furthermore, in the context of short selling constraints, it is easier to generate trading if the financial analysts make recommendations for more purchases. *Scherbina's (2007)* empirical study has confirmed that the negative information withheld by financial analysts was incorporated into stock prices with a considerable lag.

The clients of audit organisations and credit rating agencies are usually the companies under review themselves, and they are not interested in making these organisations disclose all the negative information revealed about them. In such a situation, the extent to which audit companies and credit rating agencies perform their job in line with their core function depends on how much they are willing to risk their reputation in the hopes of higher short-term profits. This is because withholding some of the negative information increases their chances of receiving further work from their current clients. However, some of the negative information may become public over time, which undermines the perceived competence and reliability of auditors and credit rating agencies.

4.6. Bounded rationality

Despite the fact that asset price bubbles are widely considered to be irrational phenomena, the market imperfections described so far can explain asset price bubbles with the assumption of perfectly rational market participants. Nevertheless, psychological reasons that limit perfect rationality do indeed play a central role in the build-up of asset price bubbles. In the models of behavioural economics, a more realistic decision-maker is assumed instead of a perfectly rational one, i.e. the market participants' ability to process the available information is influenced by the imperfections of memory, attention, calculation skills, and self-control, among other factors. According to another important example, the information reception can also be selective in order to preserve one's own convictions and self-image. The different forms of bounded rationality appearing in the models of behavioural economics are based on empirical psychological results. Decision-makers exhibiting bounded rationality usually create an asset price bubble through their bounded rationality leading to overly optimistic beliefs, which in turn can generate significant purchasing pressure for an overvalued asset. The explanations presented in the study are summarised in *Table 3*.⁸

⁸ For a more detailed overview of the models explaining asset price bubbles with the assumption of bounded rationality, see *Hirshleifer (2001)*.

Table 3	
Forms of bounded rationality in the presented models	
Explanatory factor	Content
Overconfidence	Investors consider their initial private information more accurate than it actually is.
Biased self-attribution	Investors attribute past successes to themselves and failures to outside factors.
Conservatism bias	Investors change their beliefs less based on new information as compared to their perfectly rational peers.
Representativeness heuristic	Investors overestimate the probability of the future occurrence of past events.
Newswatching	Investors disregard the portion of others' private information that can be deduced through observing market prices.
Momentum traders	Investors only take into account the information discernible from market prices.
Noise traders	Investors' beliefs differ to an uncertain extent from the perfectly rational expected value of the future market price.

Daniel et al. (1998) explain asset price bubbles with the combination of *overconfidence* and *biased self-attribution*. The essence of overconfidence in the model is that investors regard their private information concerning the fundamental value of the asset to be more accurate than it actually is; biased self-attribution means that investors attribute past successes to themselves and failures to outside factors. In the model, subsequent news confirming the initial private information exacerbates the actors' overconfidence.

Investors perform their own analyses at the beginning of the period under review, thereby gaining private information on the fundamental value of the asset. Later information observable by all investors arrives until the end of the period, making the fundamental value accurately observable. The investors who, based on their private information, assume at the beginning of the period that a value higher than the fundamental value is likely, overreact to the initial information due to their overconfidence, becoming overly optimistic. The public news arriving later can hardly correct this excessive optimism, since overconfident investors trust public news less. In fact, due to biased self-attribution, their excessive optimism is typically heightened even further, since these investors assign greater weight to the new information confirming their initial optimism than contrary information. Over time, the increasing amount of public news convinces everyone about the real fundamental value, but the optimism among the investors exhibiting bounded rationality is greater and more persistent than what would emerge among perfectly rational investors.

Barberis et al. (1998) explain asset price bubbles with the help of two other assumptions limiting rationality, *conservatism bias* and *representativeness heuristic*. Decision-makers characterised by conservatism bias change their beliefs less based on new information than their perfectly rational peers. The decision-makers following the representativeness heuristic overestimate the probability of events that are noticeably similar to already observed earlier events.

The authors analyse securities, the return of which exhibit random walk over time. Accordingly, the best rational estimate regarding the return for the next period equals the current return. Market investors do not know that the return exhibit random walk and establish their beliefs concerning future returns based on past returns. This process is influenced by the two biases mentioned above. In this model, the excessive optimism regarding future returns and thus the current fundamental value of the securities emerges as follows. If by chance increasing returns are realised in a few consecutive periods, the representativeness heuristic takes hold. The pattern of earlier, similarly increasing returns is easy to remember, and it may indicate to investors that returns are on an upward trend once again. Inevitably, the realised returns will diverge from the assumed trend and conservatism bias can manifest in the investors i.e. they do not update their expectations about the security adequately. Eventually, on account of the returns repeatedly falling short of the expected rising trend, investors update their expectations and assume that the returns fluctuate around a lower average value. This updated expectation is a “conservative” because it assumes that the next period’s returns will be closer to the expected average rather than the same value that is currently realised. Ultimately, the excessive optimism is curbed, and the assets’ price drops to around the fundamental value.

In the model by *Hong and Stein (1999)*, there are two types of traders exhibiting bounded rationality. So-called newswatchers establish their beliefs on the fundamental value of the asset solely based on the private information derived from the analysis of the asset under review, and they disregard the information from market prices that is otherwise observable. So-called momentum traders do the opposite: they disregard the information pertaining to the fundamental value, and instead estimate the fundamental value of the asset based on the development of market prices observed earlier. In this model, the new positive information pertaining to the fundamental value spreads slowly among investors, since first it only reaches newswatchers who concentrate only on their own information, and not on the potentially new information of other investors. In view of the relatively slow rise in prices, momentum traders can rightly conclude that the asset is undervalued, and thus worth buying. However, since based solely on the price movements they are unable to accurately judge when exactly the undervaluation of the asset ceases,

the price increase does not stop at the real fundamental value, and overvaluation ensues.

DeLong et al. (1990b) include a perfectly rational trader in a model framework similar to the previous one. The main finding of the analysis is that these traders do not use the opportunity for arbitrage inherent in the asset price bubble created by their peers exhibiting bounded rationality. They are even better off if they exploit their peers' bounded rationality and first make them inflate an even larger asset price bubble. In the model, this can be achieved by increasing demand for the asset more than warranted on account of the initial positive news. This generates a huge initial price increase, which convinces more momentum traders in the next period that the asset should be bought, which drives the price of the asset even higher. The large number of momentum traders entering the market will buy the asset from the rational traders even at this high price, resulting in greater gains than would be possible by not generating extra demand at the beginning of the period. One important lesson from the model is that the asset price bubble engendered by bounded rationality is not necessarily shrunk by the introduction of rational traders, but in fact, it can even be bloated by it.

Perfectly rational traders and those exhibiting bounded rationality are present at the same time in the model of *DeLong et al. (1990a)* as well. Traders exhibiting bounded rationality are the so-called noise traders who estimate the future expected value of the asset under review inaccurately as compared to rational expectations. The extent of their error is random in all periods, fluctuating around an average; however, this error is persistent in time. If rational traders are risk-averse and impatient enough, they cannot eliminate the overvaluation created by the currently overly optimistic noise traders. Rational traders know that in such a situation the asset should be sold, but they are unwilling to engage in short selling, since that is profitable in the short run only if the market price starts dropping promptly (this is because the borrowed and sold asset has to be bought and given back later). Noise traders remain overly optimistic for a while, and over the short term they can become even more optimistic, which can further increase asset prices. Therefore risk-averse rational traders who heavily discount future profits are prone to putting off large-scale selling until the excessive optimism of noise traders diminishes. In this model as well, the presence of rational traders does not ensure that the overvaluation created by traders' exhibiting bounded rationality disappears through a large-scale selling wave immediately when the mispricing is observed.

5. The role of excessive lending

Excessive lending is the situation where there is a high risk that debtors will not repay a major portion of their loans. Similar to the cyclical development of asset prices, the periods characterised by excessively intensive and subdued lending alternate. According to international experiences, the corresponding phases in asset price cycles and credit cycles are prone to aligning with each other, and when that happens, the cyclical positions are greater than on average (*Claessens et al. 2011*). Furthermore, asset price bubbles that are linked to excessive lending are more likely to end in a financial crisis (*Borio – Lowe 2002*) and entail greater losses for the real economy (*Brunnermeier – Schnabel 2015; Jordà et al. 2015*) than those that are not coupled with excessive lending. Furthermore, the dangerous mix of excessive lending and asset price bubbles can emerge relatively easily if one of them is already present (in the case of properties see: *Anundsen – Jansen 2013: Tables 1 and 2; Mian – Sufi 2011*). Therefore, excessive lending can considerably exacerbate the adverse effects of asset price bubbles. The focus of this chapter is the reasons behind the emergence of excessive lending related to asset price bubbles, and its effect on the bubbles.

Lending linked to asset price bubbles can easily become excessive for many reasons. One of the main factors is collateral pledged during borrowing. In the case of mortgages, properties are widely used as collateral, just like securities in interbank lending. Loans extended for an overvalued asset as collateral can easily become excessive, since the correction of the asset price over time devalues the collateral, which thus cannot provide adequate compensation to the lender in the case of the debtor's non-performance. Moreover, excessive lending and asset price bubbles can easily reinforce each other's impact. This requires the debtor to invest the loan in the same asset type as the one pledged as collateral. One typical example for this is the mortgage extended for purchasing residential and commercial properties.

The other important reason behind the linkage between excessive lending and asset price bubbles is that the appearance of the bubbles can prompt excessive risk-taking among not only borrowers but also lenders. Excessive lending, i.e. loans provided by lenders to subprime debtors in exchange for inadequate collateral, is only one form of excessive risk-taking. Banks may also be willing to reduce their funding costs just to extend their lending activity. The typical examples for this include disproportionate leverage and overly tight maturity and FX structure. These factors combined can entail significant systemic risk, since in such a scenario an unforeseeable large liquidity need of one of the banks' clients, or a large loss sustained by one of the banks can cause disruption in financial intermediation.

Excessive risk-taking by banks can be explained by several market imperfections that are partly identical to the reasons behind asset price bubbles.⁹ Investors who wish to invest in an overvalued asset more than their own capital generate large credit demand vis-à-vis financial intermediaries. When assessing loans, even financial intermediaries themselves often try to gauge whether it is worth investing in the given asset, since in certain cases it is known that the loan will finance the purchase of such an asset. In this scenario, lenders are in a similar situation as the economic actors who wish to invest in the asset directly. That is why it comes as no surprise that the *limited liability* of financial intermediaries can contribute to the emergence of not only asset price bubbles, but also the related excessive lending. Similarly, *herd behaviour* in lending practices can emerge among banks. Additionally, a banks' management's or owners' *bounded rationality* can also explain the emergence of excessive lending.

Among the explanations of excessive risk-taking of financial intermediaries there is also an important specific element: larger financial intermediaries can receive various forms of *state assistance* in the event of systemic stress. In the case of a systemic liquidity shortage, commercial banks can use the central bank as the lender of last resort. Deposit insurance payments protect them from the self-fulfilling bank runs of small depositors. Finally, insolvent financial intermediaries may also receive state capital injections. The instruments of state assistance seek to mitigate the real economy losses of an already developed financial stress. At the same time, they encourage excessive risk-taking on the part of the financial intermediary *ex ante*, since a large portion of the potential losses are incurred by the state instead of the owners, while the potential profits go to the latter.¹⁰ State assistance is more likely if multiple financial institutions are in dire state, since such an event threatens the functioning of the financial system as a whole. Therefore, state assistance encourages not only excessive but also correlated risk-taking, thereby strengthening herd behaviour among the institutions.

It is important to note that, according to the empirical literature, the mutually reinforcing processes of asset price bubbles and excessive lending are especially dangerous in the case of residential properties.¹¹ *Claessens et al. (2012)* expanded the study by *Claessens et al. (2009)* and examined the business and financial cycles

⁹ For more details, see, for example, Chapter 4 of *Freixas et al. (2015)*.

¹⁰ The resolution of banks on the brink of insolvency is also a form of state assistance, which, in contrast to the other interventions mentioned, reduces the excessive risk-taking of the banking system *ex ante*. This is because efficient resolution can prevent insolvency, which is often only alleviated with a capital injection by the state. The state-owned organisation in charge of the resolution temporarily assumes the ownership and management rights, and it uses these rights to separate the good assets of the distressed financial institution, and to sell them to other, solvent market participants. Timely reorganisation can steadily maintain the critical functions of the financial institution concerned, for example access to bank deposits and corporate credit lines.

¹¹ One recent, comprehensive description of the empirical and theoretical results pertaining to the macroeconomic role of residential properties can be found in *Piazzesi and Schneider (2016)*.

of 21 developed OECD countries and 24 emerging countries between 1960 and 2010, and found, among other things, that a rapid drop in housing prices entailed twice as large a decline in GDP on average than other crises. This loss was even greater when the drop in housing prices was coupled with a major contraction of the volume of credit, although the sample contained too few such elements to make the result statistically significant.

Jordà et al. (2015) examined stock prices, housing prices, banking system credit and other macroeconomic data from 17 developed countries between 1870 and 2013. The periods of credit boom coupled with housing price bubbles were more likely to be followed by a recession with financial stress than the periods of mere credit booms. On average, during recessions that were not preceded by asset price bubbles, real GDP per capita declined by approximately 2 per cent (until the trough) in roughly a year, and it took another year for it to return above its initial level. On average, for recessions following housing price bubbles, the trough was reached after around one year as well, and they entailed a similar extent of real economy losses, but a longer recovery given that the housing price bubble was not coupled with rapid credit growth. However, if the recession followed a housing price bubble coupled with a credit boom, real GDP per capita fell even in the second year of the crisis, by about 4 per cent on average before reaching the trough, and it was not able to return to its original level even after three years. Estimates for stock price bubbles yielded similar results between the bubbleless events and the housing price bubble scenarios. All in all, it can be said that the longest and most severe economic crises were preceded by a housing price bubbles coupled with large-scale lending.

Three special features of residential properties make the emergence of a housing price bubbles lead to considerable real economy losses (*Crowe et al. 2013*). First, housing prices are often prone to becoming decoupled from fundamental values. There are several reasons for this. Homes are not standardised products, they have no standardised market, and many actors are insufficiently informed. The latter is due to the fact that purchasing a home is simultaneously a crucial consumption decision (housing) and a complex investment decision (real asset). In such an environment, the role of informational frictions can be substantial, which hampers the rapid spread of the relevant information on the fundamental value. The high transaction costs incurred during the purchase of a home (e.g. moving, duties, retaining a real estate broker and a lawyer, rebuilding in line with personal preferences) and the uncertain supply reactions due to the time-consuming nature of construction, make identification of the current fundamental value even more complex. In addition, information-based herding can easily emerge among the many insufficiently informed market participants. Average home buyers are not familiar with the methodologies to accurately assess investments; therefore, many rely on

the naïve extrapolation of past market price developments or similar rules of thumb, i.e. they may be characterised by bounded rationality as well.

Second, normally home purchases are mainly financed from loans with the collateral being the home itself. The price is usually much higher than the average steady income, and therefore external funds are required as well. Furthermore, the home itself is a durable, immovable real asset; hence lenders readily accept it as collateral. Third, homes play a central macroeconomic role. Even without a housing price bubble and excessive lending, homes comprise a huge portion of households' wealth, and mortgages secured with homes comprise a major part of banks' assets. In addition, home construction contributes substantially to GDP and employment.

6. Conclusions

6.1. Asset price bubbles are endogenous phenomena affecting many assets

The explanations of asset price bubbles note that they do not necessarily emerge and burst as a result of some unexpected shock coming from outside the asset's market. All of the explanations claim that asset price bubbles are the aggregate result of the individually optimal decisions of the economic actors active on the market for the asset. Accordingly, state intervention has to change the decisions of market participants so that the aggregate of the optimal individual decisions creates as small an asset price bubble as possible.

The theoretical literature has identified several reasons behind asset price bubbles. Therefore, state intervention must potentially employ several types of targeted regulatory instruments. The explanations are less specific about what type of real or financial instruments prices are prone to becoming a bubble. Thus, the need for state interventions aimed at the elimination or mitigation of anomalies can arise on several asset markets. According to international experience, asset price bubbles coupled with excessive lending usually threaten greater losses in the real economy, especially in the case of residential properties.

6.2. Most factors that induce bubbles are difficult to eliminate

A large part of market participants' bounded rationality cannot be changed. Similarly, the basic reason behind the existence of financial intermediary institutions is that they have more expertise in investments than savers, and if their limited liability was eliminated, few actors would take the elevated business risk, and there would be a significant shortage of a vital service. It is also unlikely that we will have much more accurate and readily available information about future uncertain payments. Therefore, the current fundamental value of assets remains difficult to identify, and the information pertaining to them will not spread faster. Thus, we should always expect informational frictions and information-based herding.

The incentives due to the information advantage of financial experts are hard to overcome, and therefore reputation-based and compensation-based herding cannot be completely eliminated either. Nonetheless, all the above-mentioned reasons can be constrained or offset to some extent, and state interventions should seek to do so.

Of the reasons behind asset price bubbles the easiest to tackle are short selling constraints and the perverse incentives in the provision of information. The former is often enforced by state regulation, which can be changed, and in the case of stocks several analyses call into question the social benefits of short selling constraints (*Marsh – Payne 2012; Beber – Pagano 2013; Boehmer – Wu 2013*). Some perverse incentives in the provision of information can be corrected effectively through more efficient regulation, and state organisations can complement the activities of market participants.

6.3. Partly paternalistic state interventions are necessary

State intervention has to be partly paternalistic due to the bounded rationality of market participants. As portrayed in the examples of Chapter 4.6, asset price bubbles emerge in line with the desire of most market participants, not as an unintended and adverse side effect of a functioning market. The question in regard to state intervention is of normative basis, since everything happens in line with the expectations of most market participants. Furthermore, the reason behind state intervention can vary depending on the specific form of bounded rationality.

According to one possible approach, the state can reasonably guide decision-makers towards decisions they would probably make themselves if their rationality was not bounded, and if they are willing to shed these constraints (*Thaler – Sunstein 2003*). One example for this is the excessive optimism arising from limited memory and calculation skills.

6.4. Lessons for macroprudential policy

One of the most important state interventions related to asset price bubbles is macroprudential policy which is in charge of preventing and mitigating financial crises. Since asset price bubbles threaten the stability of the financial system through the related excessive lending, it is vital that macroprudential policy identifies when the boom phase of the asset price cycle turns into euphoria, where the asset is increasingly purchased from credit. Identifying the transition in time can help stop excessive lending before it fully unfolds. The spread of herd behaviour, excessive and correlated risk-taking of financial intermediaries with limited liability and the mass entry to the market of less informed and overly optimistic market participants exhibiting bounded rationality all take time. However, once these processes get underway, they can continue to strengthen in a self-reinforcing fashion, which is increasingly difficult to stop and neutralise.

The explanations of asset price bubbles reviewed also point out that in many respects the overvaluation emerging in the boom phase differs from the asset price bubble developing in the euphoria phase only in its extent rather than qualitatively. They share several reasons, and often the only difference is the time available for the effect of these reasons to take hold. Accordingly, the transition between the two periods is usually smooth and difficult to observe.

Two special factors can be identified that increase the asset's overvaluation primarily in the euphoria phase. The first is the excessive optimism of the market participants exhibiting bounded rationality, which can persist for a long time and thus foster euphoria. The second is excessive lending itself, which is also encouraged by the greater number of actors exhibiting bounded rationality entering the market, and which provides additional funds for boosting demand for the overvalued asset. Therefore, macroprudential policy striving to put an end to excessive lending can only seek to identify one special reason – apart from excessive lending itself – behind the euphoria that spread on the market: excessive optimism (or more broadly bounded rationality). Therefore, the euphoria can be mostly identified through the consequences of the asset price bubble rather than the special reasons, for example a rapid rise in market prices, more volatile prices, the increasing number of transactions or the distortion in resource allocation.

Macroprudential policy can efficiently constrain asset price bubbles in certain cases by spreading information pertaining to the fundamental value. Several of the explanations behind asset price bubbles are based on the difficulty of making the information dispersed among market participants widely available. This includes informational frictions, herd behaviour, the perverse incentives of information provision and certain cases of bounded rationality. When monitoring the asset markets concerned, the current overvaluation should be sought to be estimated as accurately as possible. One important aspect here is that the organisation shaping macroprudential policy should not necessarily know more about the fundamental value than what market participants know collectively. It is sufficient if it has incentives to collect this information, even partially, and to make it widely available. This is what the market often performs inadequately.

In other cases, only passing on information is not enough to halt asset price bubbles. This is the case with short selling constraints, the limited liability of financial intermediaries, and asset price bubbles that emerge due to certain types of bounded rationality. In such cases, the primary task of macroprudential policy is to rein in the potential related excessive lending and to mitigate its adverse consequences on the financial system (see, for example, *ESRB 2014a: Chapter 3*). International experience shows that the so-called debt cap rules are the most effective in curbing excessive lending (see, for example, *McDonald 2015; Cerutti et al. 2017*, and the references in them). Debt cap rules typically limit the credit amount that can be borrowed by households relative to the value of the collateral or

income, and they also limit the repayment instalments pledged as a percentage of income. Surplus capital requirements are also used, and they principally strengthen banks' resilience to external shocks. Another important supplementary feature is that they help cover banks' losses on not only household loans but also corporate and project loans.

It may be important from the perspective of the accurate calibration of interventions that in certain cases identifying excessive lending is easier than observing asset price bubbles. For example, according to the *ESRB (2014b)*, various credit-to-GDP figures can be a good measure of excessive lending. The study claims that current values that significantly exceed their long-term trend indicate an amount of credit that makes a financial stress likely in the following couple of years.

References

- Abreu, D. – Brunnermeier, M.K. (2003): *Bubbles and crashes*. *Econometrica*, Vol. 71, January: 173–204.
- Admati, A.R. – Pfleiderer, P. (1997): *Does it all add up? Benchmarks and the compensation of active portfolio managers*. *The Journal of Business*, Vol. 70, July: 323–350. <https://doi.org/10.1086/209721>
- Allen, F. – Gale, D. (2000): *Bubbles and crises*. *The Economic Journal*, Vol. 110, January: 236–255. <https://doi.org/10.1111/1468-0297.00499>
- Allen, F. – Gorton, G. (1993): *Churning bubbles*. *The Review of Economic Studies*, Vol. 60, October: 813–836. <https://doi.org/10.2307/2298101>
- Anundsen, A.K. – Jansen, E.S. (2013): *Self-reinforcing effects between housing prices and credit*. *Journal of Housing Economics*, Vol. 22, September: 192–212. <https://doi.org/10.1016/j.jhe.2013.07.001>
- Avery, C. – Zemsky, P. (1998): *Multidimensional uncertainty and herd behavior in financial markets*. *American Economic Review*, Vol. 88, September: 724–748.
- Banerjee, A.V. (1992): *A simple model of herd behavior*. *The Quarterly Journal of Economics*, Vol. 107, August: 797–817. <https://doi.org/10.2307/2118364>
- Barberis, N. – Shleifer, A. – Vishny, R. (1998): *A model of investor sentiment*. *Journal of Financial Economics*, Vol. 49, September: 307–343. [https://doi.org/10.1016/S0304-405X\(98\)00027-0](https://doi.org/10.1016/S0304-405X(98)00027-0)
- Beber, A. – Pagano, M. (2013): *Short-Selling Bans Around the World: Evidence from the 2007–09 Crisis*. *The Journal of Finance*, Vol. 68, January: 343–381. <https://doi.org/10.1111/j.1540-6261.2012.01802.x>

- Bikhchandani, S. – Hirshleifer, D. – Welch, I. (1992): *A theory of fads, fashion, custom, and cultural change as informational cascades*. Journal of Political Economy, Vol. 100, October: 992–1026. <https://doi.org/10.1086/261849>
- Bikhchandani, S. – Sharma, S. (2000): *Herd behavior in financial markets*. IMF Economic Review, Vol. 47., July: 279–310.
- Boehmer, E. – Wu, J.J. (2013): *Short Selling and the Price Discovery Process*. Review of Financial Studies, Vol. 26, September: 287–322. <https://doi.org/10.1093/rfs/hhs097>
- Borio, C. – Lowe, P. (2002): *Asset prices, financial and monetary stability: exploring the nexus*. BIS Working Papers, 114. <https://doi.org/10.2139/ssrn.846305>
- Bris, A. – Goetzmann, W.N. – Zhu, N. (2007): *Efficiency and the bear: Short sales and markets around the world*. The Journal of Finance, Vol. 62, May: 1029–1079. <https://doi.org/10.1111/j.1540-6261.2007.01230.x>
- Brunnermeier, M.K. (2008): *Bubbles*. In Durlauf, S.N. – Blume, L.E. (Eds.): The New Palgrave Dictionary of Economics, Second Edition, Palgrave Macmillan.
- Brunnermeier, M.K. – Oehmke, M. (2013): *Bubbles, Financial Crises, and Systemic Risk*. In: Constantinides, G.M. – Harris, M. – Stulz, R.M. (Eds.): Handbook of Economics and Finance 2B, Elsevier, pp. 1221–1288. <https://doi.org/10.1016/B978-0-44-459406-8.00018-4>
- Brunnermeier, M.K. – Schnabel, I. (2015): *Bubbles and central banks: Historical perspectives*. In: Bordo, M.D. – Eitheim, O. – Flandreau, M. – Qvigstad, J.F. (Eds.): Central Banks at a Crossroads: What can we learn from History? Cambridge University Press, pp. 493–562.
- Cerutti, E. – Claessens, S. – Laeven, L. (2017): *The use and effectiveness of macroprudential policies: New evidence*. Journal of Financial Stability, Vol. 23, February: 203–224. <https://doi.org/10.1016/j.jfs.2015.10.004>
- Chen, J. – Hong, H. – Stein, J.C. (2002): *Breadth of ownership and stock returns*. Journal of Financial Economics, Vol. 66, November-December: 171–205. [https://doi.org/10.1016/S0304-405X\(02\)00223-4](https://doi.org/10.1016/S0304-405X(02)00223-4)
- Claessens, S. – Kose, M.A. – Terrones, M.E. (2009): *What happens during recessions, crunches and busts?* Economic Policy, Vol 24, October: 653–700. <https://doi.org/10.1111/j.1468-0327.2009.00231.x>
- Claessens, S. – Kose, M.A. – Terrones, M.E. (2011): *Financial cycles: What? How? When?* In: Clarida, R. – Giavazzi, F. (Eds.): NBER International Seminar on Macroeconomics 2010, The University of Chicago Press, pp. 303–343. <https://doi.org/10.1086/658308>

- Claessens, S. – Kose, M.A. – Terrones, M.E. (2012): *How do business and financial cycles interact?* Journal of International Economics, Vol 87, May: 178–190. <https://doi.org/10.1016/j.jinteco.2011.11.008>
- Cochrane, J.H. (2005): *Asset pricing, revised edition*. Princeton University Press.
- Cutler, D.M. – Poterba, J.M. – Summers, L.H. (1989): *What moves stock prices?* The Journal of Portfolio Management, Vol. 15., Spring: 4–12. <https://doi.org/10.3905/jpm.1989.409212>
- Crowe, C. – Dell’Ariccia, G. – Igan, D. – Rabanal, P. (2013): *How to deal with real estate booms: Lessons from country experiences*. Journal of Financial Stability, Vol. 9, September: 300–319. <https://doi.org/10.1016/j.jfs.2013.05.003>
- Daniel, K. – Hirshleifer, D. – Subrahmanyam, A. (1998): *Investor psychology and security market under- and overreactions*. The Journal of Finance, Vol. 53, December: 1839–1885. <https://doi.org/10.1111/0022-1082.00077>
- De Long, J.B. – Shleifer, A. – Summers, L.H. – Waldmann, R.J. (1990a): *Noise trader risk in financial markets*. Journal of Political Economy, 98(4): 703–738. <https://doi.org/10.1086/261703>
- De Long, J.B. – Shleifer, A. – Summers, L.H. – Waldmann, R.J. (1990b): *Positive feedback investment strategies and destabilizing rational speculation*. The Journal of Finance, Vol. 45, August: 379–395. <https://doi.org/10.2307/2328662>
- ESRB (2014a): *The ESRB Handbook on Operationalising Macro-prudential Policy in the Banking Sector*. https://www.esrb.europa.eu/pub/pdf/other/140303_esrb_handbook_mp.en.pdf. Accessed: 10 October 2017
- ESRB (2014b): *Operationalising the countercyclical capital buffer: indicator selection, threshold identification and calibration options*. ESRB Occasional Paper, 5.
- Freixas, X. – Leaven, L. – Peydró, J.L. (2015): *Systemic Risk, Crises, and Macroprudential Regulation*. The MIT Press.
- Fair, R.C. (2002): *Events that shook the market*. The Journal of Business, Vol. 75, October: 713–731. <https://doi.org/10.1086/341640>
- Graham, J.R. (1999): *Herding among investment newsletters: Theory and evidence*. The Journal of Finance, Vol. 54, February: 237–268. <https://doi.org/10.1111/0022-1082.00103>
- Harrison, J.M. – Kreps, D.M. (1978): *Speculative investor behavior in a stock market with heterogeneous expectations*. The Quarterly Journal of Economics, Vol. 92, May: 323–336. <https://doi.org/10.2307/1884166>

- Hirshleifer, D. (2001): *Investor psychology and asset pricing*. The Journal of Finance, Vol. 56, August: 1533–1597. <https://doi.org/10.1111/0022-1082.00379>
- Hong, H. – Stein, J.C. (1999): *A unified theory of underreaction, momentum trading, and overreaction in asset markets*. The Journal of Finance, Vol. 54, December: 2143–2184. <https://doi.org/10.1111/0022-1082.00184>
- Huberman, G. – Regev, T. (2001): *Contagious speculation and a cure for cancer: A nonevent that made stock prices soar*. The Journal of Finance, Vol. 56, February: 387–396. <https://doi.org/10.1111/0022-1082.00330>
- Jordà, Ò. – Schularick, M. – Taylor, A.M. (2015): *Leveraged bubbles*. Journal of Monetary Economics, Vol. 76, December, pp. S1–S20. <https://doi.org/10.1016/j.jmoneco.2015.08.005>
- Keynes, J.M. (1936): *The General Theory of Employment, Interest and Money*. Palgrave Macmillan.
- Kindleberger, C.P. – Aliber, R.Z. (2011): *Manias, Panics and Crashes: A History of Financial Crises*. Palgrave Macmillan.
- Mankiw, N.G. – Reis, R. – Wolfers, J. (2004): *Disagreement About Inflation Expectations*. NBER Macroeconomics Annual 2003, Vol. 18: 209–248. <https://doi.org/10.1086/ma.18.3585256>
- Marsh, I.W. – Payne, R. (2012): *Banning short sales and market quality: The UK's experience*. Journal of Banking & Finance, Vol. 36, July: 1975–1986. <https://doi.org/10.1016/j.jbankfin.2012.03.005>
- Maug, E. – Naik, N. (1996): *Herding and delegated portfolio management*. IFA Working Paper, 223.
- McDonald, C. (2015): *When is macroprudential policy effective?* BIS Working Papers, 496.
- Mian, A. – Sufi, A. (2011): *House prices, home equity-based borrowing, and the US household leverage crisis*. The American Economic Review, Vol. 101, August: 2132–2156. <https://doi.org/10.1257/aer.101.5.2132>
- Miller, E.M. (1977): *Risk, uncertainty, and divergence of opinion*. The Journal of Finance, Vol. 32, September: 1151–1168. <https://doi.org/10.1111/j.1540-6261.1977.tb03317.x>
- Minsky, H.P. (1992): *The financial instability hypothesis*. The Jerome Levy Economics Institute Working Paper, 74.
- Piazzesi, M. – Schneider, M. (2016): *Housing and Macroeconomics*. In: J.B. Taylor – H. Uhlig (Eds.): Handbook of Macroeconomics, Volume 2, North Holland, pp. 1547–1640.

- Prendergast, C. – Stole, L. (1996): *Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning*. *Journal of Political Economy*, Vol. 104, December: 1105–1134. <https://doi.org/10.1086/262055>
- Scharfstein, D.S. – Stein, J. C. (1990): *Herd behavior and investment*. *The American Economic Review*, Vol. 80, June: 465–479.
- Scherbina, A. (2007): *Suppressed negative information and future underperformance*. *Review of Finance*, Vol. 12, October: 533–565. <https://doi.org/10.2139/ssrn.981063>
- Scherbina, A. (2013): *Asset Price Bubbles; A Selective Survey*. IMF Working Paper, 13/45. <https://doi.org/10.5089/9781475515299.001>
- Thaler, R.H. – Sunstein, C.R. (2003): *Libertarian Paternalism*. *The American Economic Review*, Vol. 93, May: 175–179. <https://doi.org/10.1257/000282803321947001>
- Trueman, B. (1994): *Analyst forecasts and herding behavior*. *Review of Financial Studies*, Vol. 7, May: 97–124. <https://doi.org/10.1093/rfs/7.1.97>
- Welch, I. (1992): *Sequential sales, learning, and cascades*. *The Journal of Finance*, Vol. 47, June: 695–732. <https://doi.org/10.1111/j.1540-6261.1992.tb04406.x>
- Xiong, W. (2013): *Bubbles, crises, and heterogeneous beliefs*. In: Fouque, J.P. – Langsam, J.A. (Eds.): *Handbook on Systemic Risk*. Cambridge University Press, pp. 664–713.
- Zwiebel, J. (1995): *Corporate conservatism and relative compensation*. *Journal of Political Economy*, Vol. 103, February: 1–25. <https://doi.org/10.1086/261973>

Operation of the International Monetary and Financial System – Structural Tensions of a “Non-system”

Gusztáv Báger

Over the past few decades, the international monetary and financial system has been a subject of constant economic debate. More recently, topics such as the activity of global financial institutions and corporations, developments in money and capital flows, international regulation and the role of central banks were thrust into the limelight by the outbreak of the financial crisis in 2007 and its culmination into a global crisis in 2008. The depth and impact of the crisis highlighted the deficiencies of the system, prompting participants to seek solutions that may facilitate sustainable, inflation-free and balanced economic growth by eliminating or at least mitigating the negative externalities of the system.

With that in mind, one of the objectives of this study is to provide a brief presentation of the conceptual framework and main structural components of the existing system. This is also necessitated by the considerable changes that have taken place relative to the Bretton Woods system. The second objective of the study is to present a comprehensive overview of the complicated operation of the system by outlining the role and the interaction between the key elements of the system without intending to offer an in-depth examination of the details.

We found that, owing to its significant deficiencies, the international monetary and financial system fails to meet the criteria of a well-functioning system. It is not suitable to keep exchange rate and capital flow developments in balance on a global scale, to ensure an optimal level of international liquidity, and to harmonise the relationship between international (economic and monetary policy) cooperation, spillovers and diverging volatility levels.

Journal of Economic Literature (JEL) codes: G01, G010, G21, G23, F33

Keywords: monetary and financial system, global financial institutions and corporations, IMF, central banks

Gusztáv Báger is a member of the Magyar Nemzeti Bank's Monetary Council. E-mail: bagerg@gmail.com

The Hungarian manuscript was received on 4 July 2017.

DOI: <http://doi.org/10.25201/FER.16.4.151186>

1. Main structural components of the international monetary and financial system

The international monetary and financial system (IMFS) includes arrangements, mechanisms and institutions designed to organise and regulate the financial exchanges and transactions between countries in relation to goods and services. *Table 1* illustrates the components that make up the system and their changes.

Following the gold standard system and the gold-exchange standard operating between the two world wars, under the direction of the IMF the Bretton Woods system, which was put into a legal framework by the IMF's Articles of Agreement, was implemented and successfully operated between 1946 and 1971.

One of the most typical features of the period since the inception of the system to date is a quest to find a way to build a sustainable global monetary system built on trust rather than bullion. In the quarter century following World War II, the US dollar component of the gold-dollar currency structure was, in effect, trust-based, fiduciary money.¹ In this case, the source of confidence was the United States' obligation to exchange dollars for gold at a fixed exchange rate² and the fact that the rest of the national currencies were pegged to the dollar.

	Bretton Woods system	Current system
Monetary anchor	Exterior: Ultimately gold	Internal goals (e.g. price stability)
Exchange rates	Fixed, but variable exchange rate	Mixed type (floating exchange rate in focus)
Key currencies	Actually, US dollars	Dollar dominance (less excluded)
Capital mobility	Limited	Mixed type (without limitation in focus)
International pass-through*	Regulated (conventions)	Cooperative collaborations, ad hoc
Organisation, coordination*	IMF, central banks	G8, G20, IMF, central banks

*Note: * these system components are the author's additions.*
Source: BIS, 2015

In the Bretton Woods system, (partly) linking the US dollar to gold and the exchange rate negotiation mechanism operated by the IMF played a dominant role. Under the mechanism, the parity of national currencies was established in terms of the US dollar and maintained for some time (the dollar was the only currency convertible into gold). National governments were required to negotiate with the IMF with

¹ The Latin term *fides* means confidence or trust.

² 1 ounce of gold = USD 35

respect to any change in the exchange rate of their national currencies.³ When a country was running a balance of payments deficit, it had access to credit from the contributions of IMF member states and from the reserves of the IMF. As such, the mechanism was based on the principle of mutual assistance.

The progress, however, was only half a step forward, as gold still remained in the system. On a positive note, the United States implicitly agreed to maintain its stock of gold through its balance of payments deficit (around USD 1.5 billion per year). The outflow of purchasing power from the United States to foreign countries significantly contributed to the recovery of global economy in general and Western Europe in particular. Meanwhile, substantial US dollar holdings had accumulated in the reserves of Western European banks, and the banks slowly started to exchange dollars for gold. Consequently, by the end of the 1960s, the gold reserves of the United States fell to USD 10 billion from around USD 25 billion in the years following the war. In the early 1950s, experts had assumed that the United States would not allow its stock of gold to fall below this – safe – level. When the US gold stock approached this level, the United States unilaterally terminated the dollar's convertibility to gold on 15 August 1971 (for more detail, see *Báger 2011*).

The structure and characteristic features of the new IMFS emerging from 1971–1973 were markedly different from those of the Bretton Woods system. As opposed to the exchange of dollars for gold as an external monetary anchor, the role of the *internal monetary anchor* increased sharply in ensuring price stability and also in terms of a number of other important economic indicators, such as the employment level in the USA. Consequently, domestic factors and objectives became the primary drivers of monetary policies. (Note that at the time of the – original – gold standard regime, exchange rate stability was practically the only economic policy objective of governments/central banks and it remained a crucial aspect in the Bretton Woods system as well).

Another important difference was the transition to a *floating (market) exchange rate regime*, at least in key developed countries. With the existence of various other exchange rate systems in several other countries, the new regime is considered to be a hybrid system.

While the possibility had been raised several times in professional circles (and indeed attempts have been made to that effect), no steps were taken to restore the fixed exchange rate regime. This can be partly attributed to the rapid development of information technology, which afforded quick and comprehensive access to

³ If the exchange rate change was less than 10 per cent, it sufficed to notify the Board of Governors of the IMF. Any change in excess of 10 per cent had to be approved by IMF members representing at least 75 per cent of the IMF's quota.

financial markets for large-scale international capital flows that by then rejected any new restrictions. To explain the reasons, we refer to *Isard (2005)*:

- The United States, Japan and the euro area are subject to different political and economic shocks and exhibit different stabilisation responses to these shocks.
- In terms of economic performance, the three countries (groups) are in a different position, and they cannot risk giving priority to exchange rate stability over domestic macroeconomic stability.
- Maintaining greater stability of key currency exchange rates would require such an immense and consistent effort that would not have clear beneficial effects either for the countries concerned or for the rest of the world.

It should be noted that the reasons listed above can only be interpreted in relation to the Bretton Woods system; they do not rule out the possibility that the stabilisation of exchange rates would be beneficial for smaller groups of countries, integrations or currency areas.⁴

In the Bretton Woods system, the US dollar was clearly the world's leading currency, and it *retained its dominant role* even under the new IMFS. Meanwhile, however, other currencies also gained significance (initially the euro and more recently, the Chinese renminbi). As of 1 October 2016, the latter became the fifth member of the SDR basket with a weight of 10.92 per cent.

It is yet another important difference that, as opposed to the previous system, the restriction on capital flows was lifted in most developed and emerging countries in the new IMFS. The gradual unfolding of financial globalisation over the past four decades was a key driver in this process. In the context of economic globalisation, financial globalisation may be defined as the complex integration process of the financial system of a country and the country's relationship with international financial markets and institutions. In the 1980s and 1990s, this integration process was primarily dedicated to development finance. In addition to such drivers as commercial openness, the development of domestic financial systems, the level of economic development, regional integration and the establishment of financial centres, the liberalisation of domestic financial systems also facilitated the process. Besides governments and financial systems, the International Monetary Fund and the World Bank played a pivotal role in tearing down the barriers to capital flows (for more detail, see *Báger 2010*).

In relation to the shift to floating exchange rates and the abandonment of controls on international capital flows we should also mention the *impossible trinity*

⁴ See, for example, *Mundell (1960; 1961a)*

(“trilemma”) concept, which offers an important interpretation option with regard to the change in the IMFS. According to the theory developed in the early 1960s (Fleming 1962; Mundell 1961b; 1962), it is impossible for any country to maintain a fixed exchange rate regime, an autonomous monetary policy aimed at domestic stabilisation objectives, and the strong presence of mobile foreign capital at the same time. Of the three factors, the creators of the Bretton Woods system must have recognised – and obviously gave priority to – the significance of the autonomy of national monetary policy in the 1940s already, and allowed individual countries to place a restriction on international capital flows. This solution contributed significantly to ensuring the efficient operation of the Bretton Woods system up until the mid-1970s.

In the Bretton Woods system, monetary systems and conditions exerted their mutual effect on one another in a regulated, institutional form, primarily through the transactions of international trade and balance of payments items. This feature ensured and at the same time eased *international macroeconomic coordination*. In the new IMFS, this objective is served by the various forms of cooperation stemming from the increased role of the internal monetary anchor on the one hand and the highly diverse international spillover effects on the other hand. According to Caruana (2015), *spillover effects* take hold through the following four channels:

- monetary policy stance, which may support loose or tight monetary conditions;
- the international use of currencies (USD, euro, etc.), which may influence – through the monetary policy of the given country – the user country’s monetary policy stance;
- the integration of financial markets, which allows global common factors to move bond and equity prices;
- the availability of external funding, which can be a complement to internal loans in times of boom and economic crisis.

Through the transactions of these channels, monetary and financial regimes can reinforce each other, but they can also amplify domestic imbalances to the point of instability.

After the abandonment of the Bretton Woods system, developed countries had a clash of opinions regarding the formulation of the *institutional conditions* of the new IMFS. After the Group of Ten (the ten most advanced economies) had come to the foreground, the focus shifted once again to the role of the IMF and the Group of Twenty, as a ministerial-level Advisory Board. This Board had an adequate institutional background and political support on the part of both developed and emerging economies. However, it did not have any acceptable plan about how

to restore the stability of the international payment system; consequently, the principle of exchange rate stability – despite the proposal of France and the USA in support of it – was ultimately rejected. In addition, the IMF’s controlling role of exchange rate regime proved to be far weaker than anticipated, and survived only within the framework of bilateral and multilateral supervision, as reflected in the second amendment to the IMF’s Articles of Agreement in 1978.

In light of the growing instability of the exchange rate regime, in the first half of the 1980s France, the USA and a few other countries repeatedly called for “a new Bretton Woods”. The Group of Five (G5, composed of France, Germany, Japan, the United Kingdom and the United States), however, stood up against the proposal, taking steps in 1985–1987 such as the Plaza Agreement signed on 22 September 1985 in New York in an effort to halt the sustained multiannual appreciation of the dollar and to stabilise the rates around a new equilibrium (*Boughton 2009*).

From the beginning of the 1980s, various *collegial leadership* arrangements such as the G5, the G7 (G5 plus Canada and Italy) and then the G20 assumed an increasingly important role in *international macroeconomic coordination*, representing a higher-level alternative practice than independent national policies. Experts describe the role of the USA in these bodies as first among equals.

2. Main features of the operation of the international monetary and financial system

Due to space limitations, the primary focus of this study is on *countries wielding global influence* in general and on the activity of the USA and its role in the operation of the IMFS in particular. Narrowing down our topic allows us to provide an overview of the trends in the operation of the IMFS, the main progressive changes related to the factors constituting the system and the shortfalls of the system’s operation.

2.1. Changes in exchange rates

Floating rates are of great significance for inflation targeting monetary policies, as the exchange rate influences changes in prices. Changes in the real exchange rate also affect the current account balance. However, it poses a problem that the volatility of floating rates is greater than that of a fixed exchange rate; therefore, the intent to mitigate this volatility is understandable.

Practical experiences show that the operation of the floating exchange rate regime falls short of “textbook” expectations in several regards.

- A large part of the world employs different exchange rate regimes (fixed exchange rate regime, currency board, etc.) with significant asymmetry between some of

these, which may give rise to inconsistent political decisions from a national and global perspective.

- The *volatility of exchange rates* is a typical feature, and the magnitude of this volatility is larger than warranted by either the macroeconomic fundamentals or the exchange rates of the key currencies.
- The – occasionally manipulated – changes in exchange rates contributed to the build-up of significant and persistent imbalances (the substantial balances of the current items of international balances of payments) and to the *accumulation of excessively large official foreign exchange reserves*.⁵ Numerous studies have pointed out that the deficit on the USA's current account could only be reduced with a sharp decline in the US dollar's real exchange rate.⁶
- Owing to the dollar's *dominant role in the IMFS* (Saccomann, 2012) (for example, it accounted for 87 per cent of the turnover of forex markets in April 2013 and made up 62.9 per cent of foreign exchange reserves in 2014 Q4), the United States enjoyed a privileged position in financing its external deficit at low cost.

According to its real effective exchange rate indicator, before the global financial crisis the US dollar depreciated by 25 per cent between 2002 and 2007, while the balance of payments of the USA remained above 5 per cent of GDP. As illustrated by *Figure 1*, following its outstanding appreciation in the first few years of the 1980s, the depreciation of the US dollar in the period of 1985–1991 was also initially accompanied by a high balance of payments deficit; the two indicators converged to a near-balance position in 1991. This convergence, however, did not occur in the period of 2002–2007: *the difference between the indicators continued to grow until 2006, whereas the balance of payments deficit stood at 5.5 per cent of GDP even in 2007.*

According to an analysis by the IMF (IMF 2008), the diverging trends in the two indicators can be attributed to factors such as:

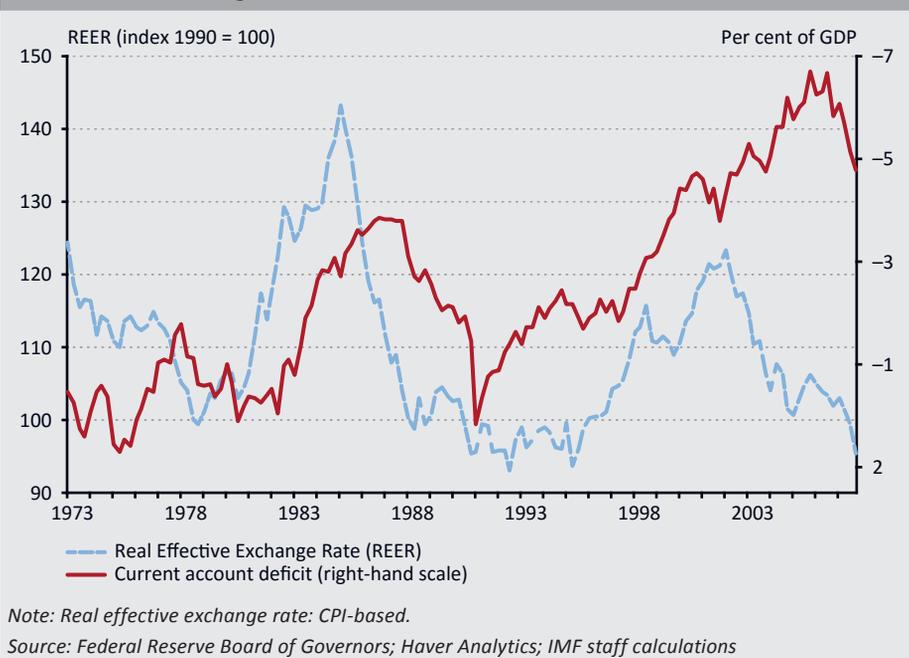
- the erosion of the USA's competitiveness in the light of the real effective exchange rate, which resulted from increased trade with “low-price” emerging and developing economies in the early 1990s (Thomas – Marquez – Fahle 2008);

⁵ Official foreign exchange reserves represent the reserve assets held by central banks, which may comprise the five freely usable currencies acknowledged by the IMF (US dollar, euro, British pound sterling, Japanese yen and, from 2016, the Chinese renminbi). The bulk of the reserves consists of allocated reserves, the currency composition of which is released by the official foreign exchange reserve databank of the IMF (COFER).

⁶ According to econometric estimations, a 10–20 per cent real effective depreciation of the US dollar is needed in order to reduce the current account deficit of the United States by 1 per cent of GDP (Krugman 2006; Mussa 2004). See also Edwards (2005).

- favourable business cycle developments up until 2006, rapid economic growth accompanied by rising imports;
- a surge in oil prices, which raised the balance of payments deficit of oil importers; and
- a financial market situation where the balance of payments deficit of the USA could be financed from FDI inflows.

Figure 1
Real effective exchange rate and current account balance



According to the assessment of the IMF's Consultative Group on Exchange Rate Issues (CGER), with the low exchange rate level reached in 2007 following the dollar's depreciation, the dollar's exchange rate became broadly consistent with its medium-term equilibrium. At the same time, however, the national currencies of numerous countries with positive balance of payments developed a close connection to the dollar, which hampered the necessary adjustment process. Indeed, rather than reducing existing imbalances or preventing the build-up of new ones, this merely resulted in the redistribution of global imbalances.

2.2. The emergence of global imbalances

Another momentous phenomenon of the operation of the IMFS was the emergence of unsustainable global imbalances. The current account deficit – or surplus – persisted in numerous countries and eventually posed a global challenge.

It was heightened demand for the US dollar as a reserve currency that led to the accumulation of global imbalances after the breakdown of the Bretton Woods system, especially in the years following the 1997–1998 South-East Asian crisis. Apart from China’s successful growth, this was fostered by the decision of several Asian countries to accumulate dollar reserves at an increasing rate in order to prevent such crises originating from exchange rate changes as the one seen in 1997–1998. In addition, the substantial savings of Chinese households and the low savings rate of the United States also contributed to the emergence of global imbalances, the increase in China’s dollar reserves and the current account deficit of the United States.⁷

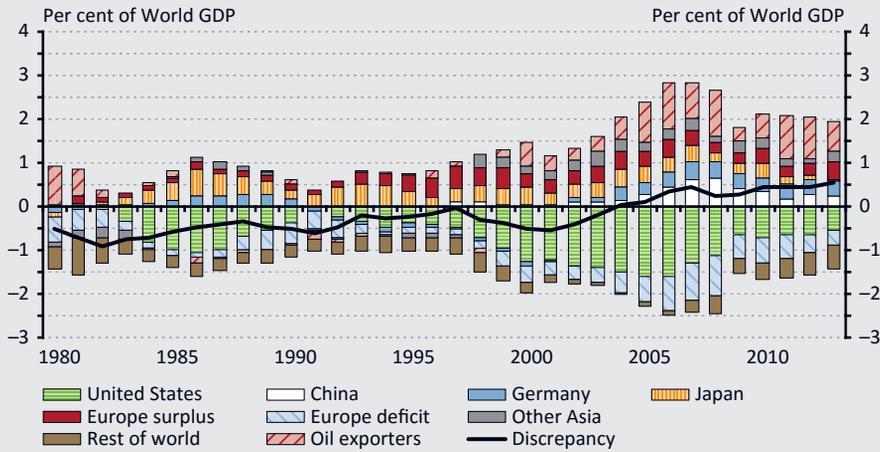
These processes raised the deficit on the US current account balance in such a way that the increase in the deficit did not exert a significant adjustment pressure, which “normally” would have called for tight monetary policy. This, however, was not the case. We should note that even countries running balance of payments surpluses failed to experience inflationary pressures that would have prompted adjustment in similar cases. In such a situation the USA’s low savings rate not only allowed but indeed, supported the accumulation of dollar reserves as did the fact that a shortage of dollars, as a reserve currency, did not cause any economic or political problems for the United States.

As a result of these underlying processes, from 1998 global imbalances started to build up on a large scale, exceeding the levels seen in previous years and peaking around 2006 (*Figure 2*).

In 2006, the aggregate balance of current accounts (deficits and surpluses together) equalled 5.6 per cent of world GDP; it declined by nearly one third at the height of the global crisis in 2009 and then, after a moderate increase in 2010, dropped to 3.6 per cent in 2013. Deficits accounting for larger distortions were recorded in 2006–2008 in the USA and some parts of Europe, while China, Germany, Japan and oil exporters reported considerable surpluses.

⁷ *Kürthy (2013)* provides a multi-faceted, detailed analysis of the triggers of global imbalances in his study entitled “*Global imbalances in a stock-flow consistent model*”. It is a significant accomplishment of the study that the author sought a systemic explanation to the imbalances by: analysing the conflicts arising from the co-existence of economic and social coordination mechanisms; rethinking the role of the international financial system and investigating why mainstream economics fails to give a consistent answer to the sustainability issues of global imbalances.

Figure 2
Global current account (“flow”) imbalances



Note: Oil exporters = Algeria, Angola, Azerbaijan, Bahrain, Bolivia, Brunei, Chad, Republic of Congo, Ecuador, Equatorial Guinea, Gabon, Iran, Iraq, Kazakhstan, Kuwait, Libya, Nigeria, Norway, Oman, Qatar, Russia, Saudi Arabia, South Sudan, Timor-Leste, Trinidad and Tobago, Turkmenistan, United Arab Emirates, Venezuela, Yemen, Hong Kong Special Administrative Region, India, Indonesia, South Korea, Malaysia, Philippines, Singapore, Taiwan, Thailand. European economies (excluding Germany and Norway) are sorted into surplus or deficit each year by the signs (positive or negative, respectively) of their current account balances.

Source: IMF Staff calculations (IMF World Economic Outlook, October 2014)

The constellation of deficits and surpluses also changed significantly (Table 2).

Representing a substantial weight among the deficit economies, the US deficit accounted for -5.8 per cent of GDP in 2006. Of the ten largest deficit economies, the US figure was exceeded by Greece (-11.3 per cent), Portugal (-10.7 per cent), Spain (-9.0 per cent) and Turkey (-6.0 per cent). In 2013, US deficit shrank to -2.4 per cent of GDP, while Turkey's deficit rose to -7.9 per cent. In the same year, the deficit of the United Kingdom rose to 4.5 per cent of GDP, and Brazil (-3.6 per cent), Indonesia (-3.3 per cent), Canada (-3.2 per cent) and Australia (-3.2 per cent) joined the ranks of the ten largest deficit economies.

Table 2
Largest deficit and surplus economies, 2006 and 2013

	2006			2013		
	USD billions	Per cent of GDP	Per cent of World GDP	USD billions	Per cent of GDP	Per cent of World GDP
1. Largest deficit economies						
United States	-807	-5.8	-1.6	United States	-400	-2.4
Spain	-111	-9	-0.22	United Kingdom	-114	-4.5
United Kingdom	-71	-2.8	-0.14	Brazil	-81	3.6
Australia	-45	-5.8	-0.09	Turkey	-65	-7.9
Turkey	-32	-6	-0.06	Canada	-59	-3.2
Greece	-30	-11.3	-0.06	Australia	-49	-3.2
Italy	-28	-1.5	-0.06	France	-37	-1.3
Portugal	-22	-10.7	-0.04	India	-32	-1.7
South Africa	-14	-5.3	-0.03	Indonesia	-28	-3.3
Poland	-13	-3.8	-0.03	Mexico	-26	-2.1
Total	-1,172		-2.3	Total	-891	-1.2
2. Largest surplus economies						
China	232	8.3	0.46	Germany	274	7.5
Germany	182	6.3	0.36	China	183	1.9
Japan	175	4	0.35	Saudi Arabia	133	17.7
Saudi Arabia	99	26.3	0.2	Switzerland	104	16
Russia	92	9.3	0.18	Netherlands	83	10.4
Netherlands	63	9.3	0.13	South Korea	80	6.1
Switzerland	58	14.2	0.11	Kuwait	72	38.9
Norway	56	16.4	0.11	United Arab Emirates	65	16.1
Kuwait	45	44.6	0.09	Qatar	63	30.9
Singapore	37	25	0.07	Taiwan	58	11.8
Total	1,039		2.1	Total	1,113	1.5

Source: IMF, World Economic Outlook database

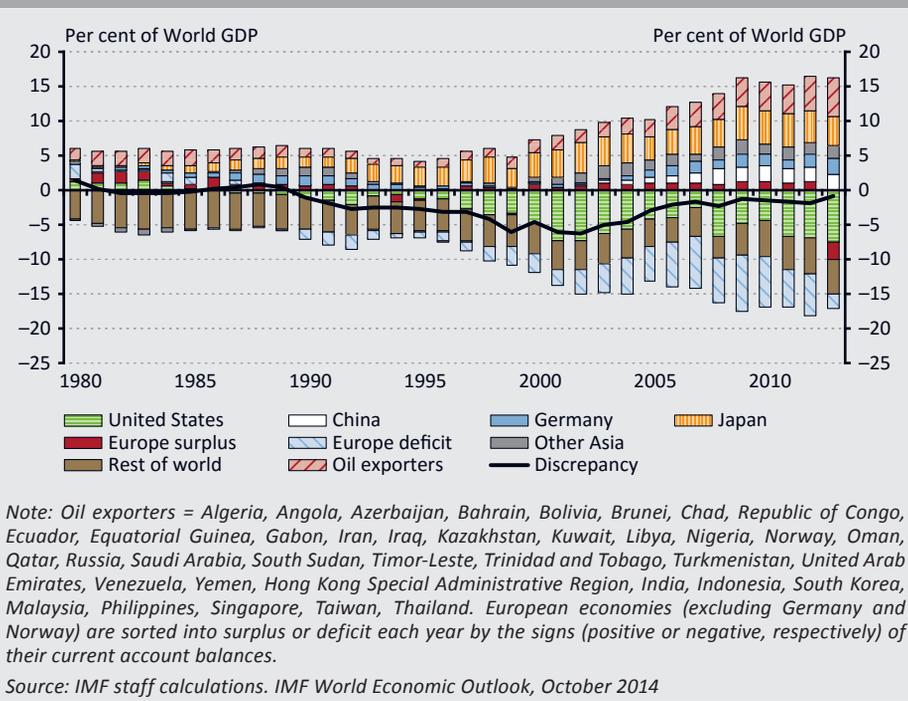
In the lead of countries running surpluses, China and Germany switched places: Germany, which took second place in 2006 with a balance of payments surplus of USD 182 billion, ran a USD 274 billion surplus in 2013, while the surplus of China amounted to USD 183 billion. In 2013, Saudi Arabia and Kuwait increased their respective surpluses of USD 99 billion and USD 45 billion in 2006 to USD 133 billion and USD 72 billion, respectively. It is noteworthy that in 2013 South Korea, Kuwait, the United Arab Emirates and Qatar joined the top 10 countries with balance of payments surpluses, while Russia dropped out.

Based on the “flow” analysis of global imbalances presented so far, we may draw the reassuring conclusion that by 2013, the total current account balance moderated to an acceptable level as a percentage of world GDP. The USA’s balance of payment deficit and the surpluses of China and Japan almost halved, although the large surpluses of some European countries and those of oil exporters persisted. In this regard, it is also a notable change that the systemic risks threatening the global economy abated significantly. However, two significant risks remained. One of them – the “price” of the moderation of external imbalances – is the exacerbation of domestic imbalances (increasing unemployment and rising output gap costs). The other risk is associated with the fact that while balance of payments (“flow”) positions improved, the investment/savings positions of creditors and debtors (“stock” imbalances) diverged further.

This is why it is necessary to also analyse the evolution of imbalances on a “stock” basis, i.e. based on net external assets calculated from the balance of payments statistics. Such an analysis is especially warranted in cases when external money market conditions render economies vulnerable; for example, when external sources of finance dry up unexpectedly or a credit crunch takes hold (*Catao – Milesi-Ferretti 2013*).

In this analytical dimension, global imbalances continued to grow even in the period of 2006–2013, in stark contrast with the result of the “flow” analysis (*Figure 3*).

Figure 3
Global net foreign assets (“stock”) imbalances



This can be attributed to the following three groups of factors (IMF 2014b)

- Even after the significant narrowing, global flow imbalances remained positive; therefore, stock imbalances continued to widen.
- Valuation effects can change asset positions. It should be mentioned that valuation effects are influenced by the initial international investment position of the countries (creditor or debtor) and the composition of their gross assets and liabilities. The debtor position of the USA was unique in this regard: after having experienced a downward revision of its growth prospects, the value of US assets continued to rise even as its international investment positions weakened.
- Growth effects also supported higher imbalances as a share of GDP: for creditor economies, GDP growing ahead of net foreign assets lowered net foreign asset ratios, whereas in debtor economies these effects contributed to lower net foreign liability ratios.

The level of *foreign exchange reserves* – another important indicator of global imbalances signalled by the position of current account balances – *increased and its distribution was disproportionate* (Table 3).

Table 3				
Global distribution of international reserves				
<i>(USD trillions)</i>				
	1998	2010	2013	Change between 1998 and 2013
World	1.6	9.3	12.1	10.5
Developed economies	1.0	3.1	3.4	2.4
Emerging and developing countries	0.6	6.2	8.7	8.1

Source: IMF (2014a)

As indicated by the table, the foreign exchange reserves of the world surged to USD 12.1 trillion in one and a half decades (representing a USD 10.5 trillion increase), accounting for 16.2 per cent of world GDP in 2013. In particular, compared to the USD 2.4 trillion growth of developed countries, the growth recorded in emerging and developing countries approached USD 8.1 trillion. On the one hand, these substantial reserves boosted the countries' resilience to potential financial shocks; on the other hand, they allowed some emerging and developing countries (such as China and Saudi Arabia) to record – in addition to the exports of goods and services – significant capital outflows to advanced economies, primarily the United States. The magnitude of the US dollar reserves also poses a risk to the USA by allowing Asian creditor countries to become important actors, which may increase volatility.

Consequently, with regard to global imbalances we can conclude that in the review period, this operational deficiency of the IMFS played a prominent role in economic growth falling short of its potential in several countries, as the countries concerned were unable to reduce their balance of payments deficits and surpluses to a noticeable extent. The consequences of this failure are also evident in the euro area, where Mediterranean countries accumulated a sustained deficit, while surpluses were built up in Northern economies.

Another consequence of the imbalances was the emergence of a *global savings glut*.⁸ After the initial positive expectations, the surge in cash hoarding led to asset bubbles and a contraction in global investment opportunities. It should be borne in mind in this regard that balance of payments surpluses and the accumulation of foreign exchange reserves were not the only source of the global savings glut in several emerging economies in Asia: oil revenues deriving from the sharp rise in

⁸ Global savings glut. The term was first used by former Fed Chairman Ben Bernanke (*Bernanke 2005*).

international oil prices – and the simultaneous unfolding of the two processes – represented another important contributor.

2.3. Liquidity

Frequent swings in the ebb and flow of global liquidity represented another weakness in the operation of the IMFS. For example, the USA's accommodative monetary policy was an important factor in the liquidity tide in the early 2000s, which was also buoyed by low interest rates and risk spreads. In this environment, investors' appetite for risk increased; investments in the USA were deemed by the rest of the world both liquid and secure.

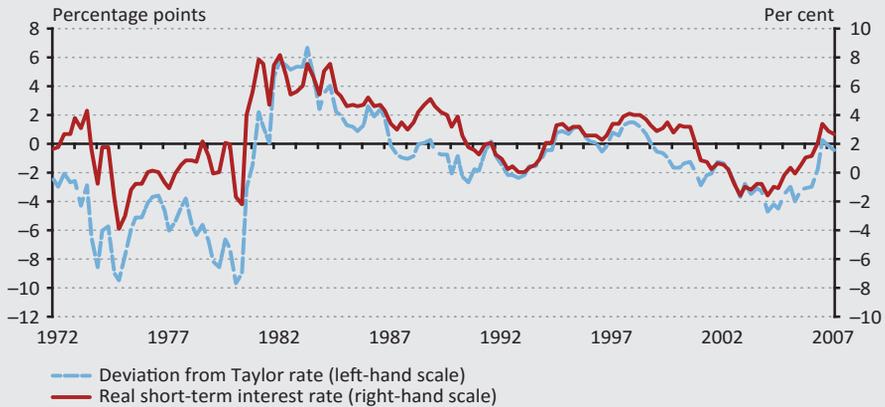
Before illustrating liquidity developments, we should briefly present the concept itself. Initially, the term "international liquidity" denoted the changeability of foreign exchange reserves. Later on, with financial globalisation and the interconnectedness of money markets the concept and the sources of liquidity expanded (Coene 2012).

The complex, multifaceted concept of global liquidity is defined by the Bank for International Settlements (BIS) and by the IMF through two components: official liquidity and market liquidity. *Official liquidity* is a funding opportunity that is unconditionally available to settle needs/claims through monetary authorities (central banks). This opportunity is made available through the activity of central banks with the assistance of such instruments as the foreign exchange reserves, lending, swap lines, IMF programmes and SDR allocations. These instruments are aimed at mobilising the available official liquidity the accumulation of which, ultimately, is one of the tasks of central banks. *Market (private) liquidity* is created by banks and other financial institutions through their operations. The common element of these two liquidity components is "ease of financing" (BIS 2011).

Owing to the extreme complexity of the concept, global liquidity developments can only be approximated with proxy indicators. According to the IMF (IMF 2007), the official liquidity position can be captured by the *short-term real policy rates*, the evolution of which is largely determined by the monetary policy stance (degree of accommodation) of globally important central banks. Another proxy measure is the *Taylor rate* or the *Taylor rule*, which determines the short-term policy rate based on the deviation of actual inflation from the inflation target, the neutral (long-term) real interest rate and the cyclical position of the economy. *Figures 4* and *5* illustrate the key policy rates for the USA and for the euro area and their deviation from the Taylor rate.

Since the real interest rates calculated from the policy rates were negative, in the period between 1998 and 2005 the monetary policy stance of the USA can be described as strongly accommodative before – towards the end of the review period

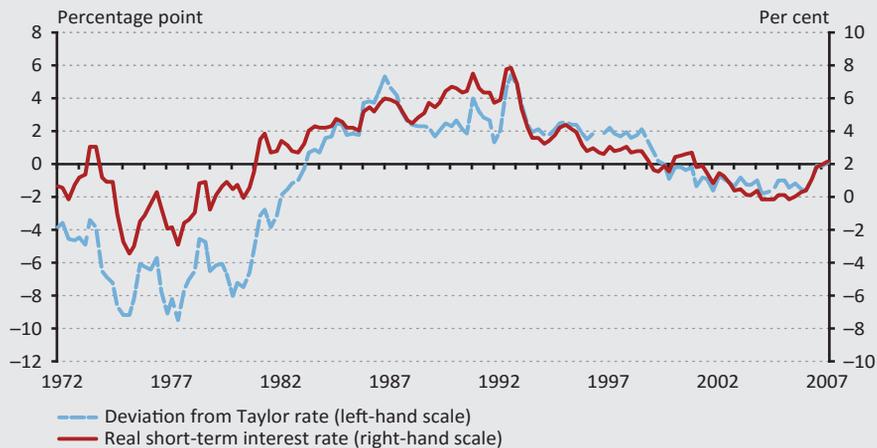
Figure 4
Interest rate-based measure of liquidity for the USA



Source: IMF staff calculations. World Economic Outlook, October 2007

– the applied interest rate approached the level of the equilibrium or neutral rate of interest. In 2003–2005, monetary policy was even more accommodative than in the early 1990s. Despite the long-term similarity of the monetary adjustment path, the European Central Bank (ECB) was characterised by less pronounced monetary accommodation in the period 2003–2005 than in the 1970s, when the accommodative monetary policy stance peaked at a historical level even in the USA.

Figure 5
Interest rate-based measure of liquidity for the euro area

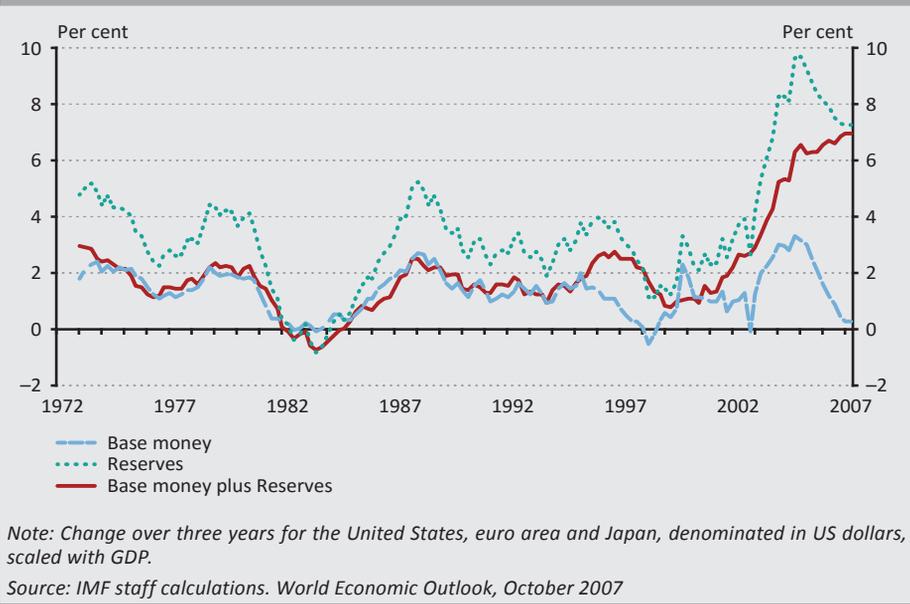


Note: Before 1998, purchasing-power-parity-weighted average of euro-area Member States.

Source: IMF staff calculations. World Economic Outlook, October 2007

According to the third, *quantitative measure* of global liquidity (Figure 6), monetary policy was more accommodative in the first half of the decade between 2000 and 2010 than either in the 1970s or in the middle of the 1970s, with the adjustment peaking in 2005. The increase in international foreign exchange reserves, as seen in Section 2.2, largely contributed to the liquidity expansion.

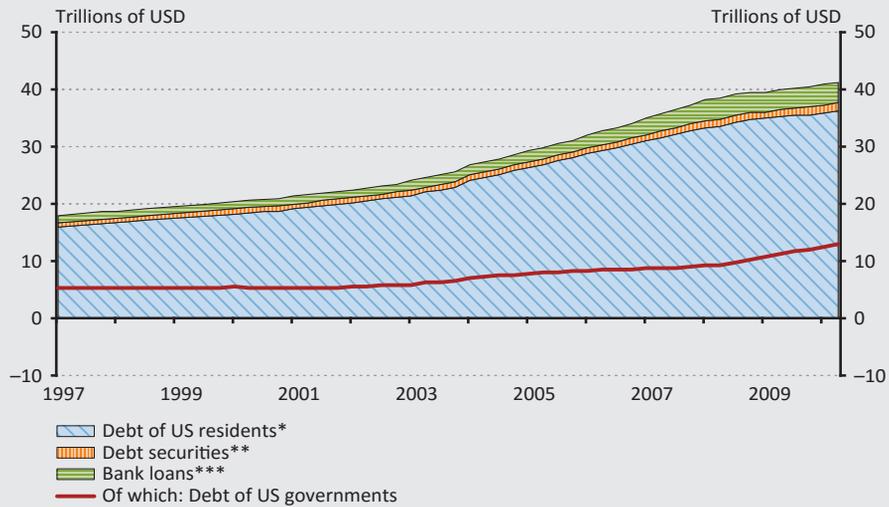
Figure 6
Quantitative measures of liquidity



Another group of global liquidity indicators measures the size of cross-border *global (bank) loans and bond issues*. One of the measures is based on a *currency denomination perspective* (dollar, euro, yen, etc.) (Figure 7). We compare the size of US dollar-denominated credit extended to the non-financial sector of the USA with the size of US dollar credit extended to the non-financial sector outside the USA. The figure shows that, in mid-2010, dollar credit to the non-financial sector worldwide amounted to 13 per cent of the dollar credit extended to the non-financial sector of the USA, compared to 10 per cent in mid-2000. If dollar credit to governments is excluded, the share of the international component was even higher (17 per cent) in 2010. We also find that US dollar credit to the rest of the world grew faster than credit to US residents. The latter grew around 9 per cent on average year on year between 2000 and 2007, amounting to USD 23 trillion or 167 per cent of US GDP. By contrast, the annual growth of dollar credit to the non-financial sector outside the USA was 30 per cent in the same period (BIS 2011).

Figure 7
Domestic and international US dollar credit

(credit to the non-financial sector)

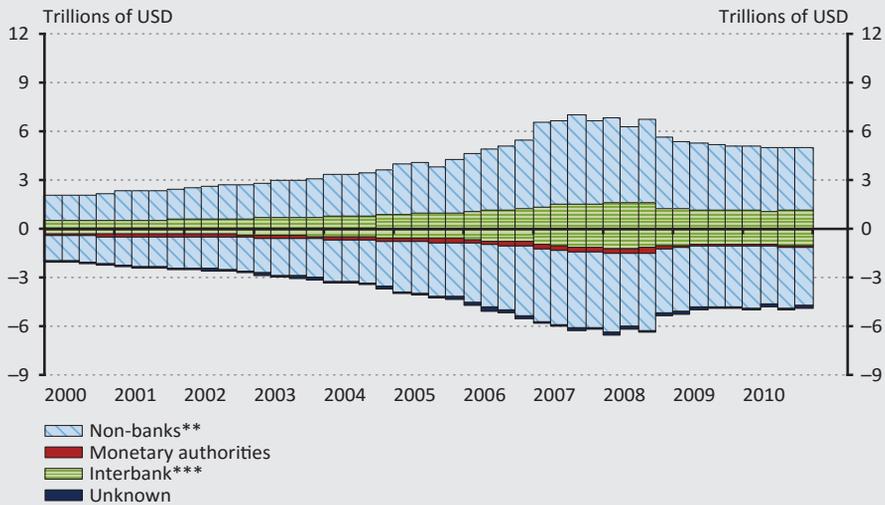


Note: * Non-financial sector debt of residents of the United States. ** Outstanding debt securities issued by non-financial residents outside the United States. *** Cross-border and local US dollar loans to non-bank residents outside the United States.

Source: BIS (2011)

Another two approaches to measuring liquidity are *credit analysis* – disregarding the currency denomination – and the examination of the *form of international balance sheet positions* in which the credit is received by the recipient economy. With respect to the former, based on the US dollar books of a sample of European banks, before the 2008–2009 crisis many large international banks had built up in their countries sizeable US dollar asset positions through FX swaps from short-term loans, and encouraged growing exposures to investment risk (BIS 2011). As regards the latter, global credit aggregates continued to expand throughout the 2001–2010 period, while cross-border credit contracted during the years of the crisis (Figure 8). The contraction in cross-border credit was even more pronounced in the USA and more moderate in the euro area and in European emerging economies, although the downward shift had begun earlier – in 2005–2006 – in the latter regions.

Figure 8
Internationally active European banks' on-balance sheet USD positions*
(gross, by counterparty sector)



Note: * Estimates are constructed by aggregating the worldwide on-balance sheet cross-border and local positions reported by internationally active banks headquartered in Germany, the Netherlands, Switzerland and the United Kingdom. ** International positions vis-à-vis non-banks plus local positions vis-à-vis US residents booked by banks' offices in the United States. *** Interbank transactions: estimated net interbank lending to other (unaffiliated) banks.

Source: Bloomberg; JPMorgan Chase; BIS consolidated statistics; BIS locational statistics by nationality

Based on the experiences of the past period, the cyclical nature of the evolution of global liquidity and the resulting shocks to financial stability indicate that liquidity exhibited substantial swings and volatility and that, owing to this reason, the sudden build-up of liquidity shortages and hence, the deceleration of economic growth, could not be prevented from time to time. The latter effect persisted particularly long during the period following the 2008–2009 crisis. Mitigating the excessive liquidity shortages became imperative in developed economies, and central banks' reciprocal FX swap arrangements served this purpose successfully (*Goldberg – Kennedy – Miu 2011*). Besides FX swap transactions, central banks may contribute to the international distribution of liquidity with various additional instruments, such as interbank repos or cross-border collateral arrangements (CBCAs). As part of the crisis management measures, the asset purchase programmes of central banks (Fed, European Central Bank, Bank of England, Bank of Japan, etc.) have also played a pivotal role in facilitating the expansion of official liquidity in recent years.

International financial institutions, in particular the IMF, are also important participants in mitigating liquidity shocks. The IMF has expanded the range of its lending instruments with the Flexible Credit Line (FCL) and the Precautionary Credit Line (PCL), both designed to meet the liquidity needs of countries with sound macroeconomic fundamentals for crisis-prevention measures. Another important instrument in the IMF's toolkit are Special Drawing Rights (SDR), which are primarily intended to mobilise official liquidity in respect of the key reserve currencies while they may also be used to boost global liquidity. It should be noted in this regard that, under the 14th General Review of Quotas, the overall quota was raised to SDR 477 billion (around USD 668 billion).

2.4. Stability

Financial stability – perhaps the most comprehensive requirement for the IMFS – has not been attained despite the fact that this deficiency jeopardises the achievement of central banks' primary objective. Efforts to strengthen stability are demanded by circumstances such as:

- the vulnerability of emerging and developing countries stemming from the fact that, on the one hand, foreign capital is present in these regions on such a large scale that the “*depth*” (Bernanke 2005) – level of development – of the financial sector (market) does not support its efficient operation and, on the other hand,
- that advanced economies have pursued diverging monetary policies in managing the arising macroeconomic and funding risks.

It is a complex task to define and set the objective of financial stability due to the multi-dimensional interpretation of its concept that ranges from the application of inflation targeting through asset prices, political conditions, loan sizes and changes in the financial cycle to integrated inflation targeting. The latter (new) concept is to be understood as the joint, coordinated application of monetary and macroprudential policies which, over the long run, may contribute to the sustainability of the financial system and hence, foster economic growth.⁹ This explains why it is so difficult to measure global stability: there is a need to define a point of reference (the equilibrium state, target levels or the prevailing trends) against which changes are compared in each dimension of stability. Moreover, the role of policies aimed directly at strengthening stability should be also determined.

In the course of crisis management, significant progress was made in this regard. It is an important achievement of the European asset quality review (AQR) that

⁹ Macroprudential policies typically address bank lending and liquidity adequacy. As a result of the crisis and the policy pursued by the major central banks, the proportions of financing have changed: in many countries, large corporations accessed new funding primarily through bond issuance, while the share of bank financing diminished. Macroprudential policy did not – and would not be able to – follow bond issuance. Stability can only be approximated, but not fully reached, by macroprudential policy.

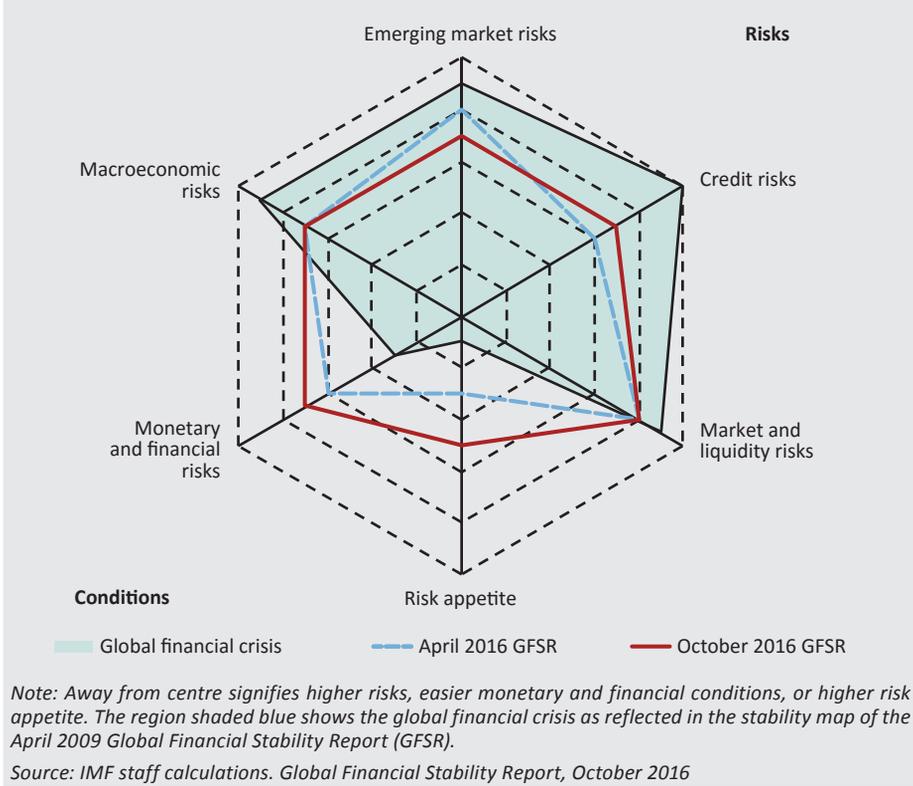
it dissipated the uncertainties around the quality of bank assets. In addition to stress tests, the establishment of the European Single Supervisory Mechanism (SSM) succeeded in mitigating the risks surrounding banks' balance sheets and improved investors' confidence in the sector. Another important macroprudential instrument of the policies aimed at fostering stability is the countercyclical capital buffer rate, which is intended to allocate extra cushions of capital for financial crisis situations and, in the case of excessive lending, for the mitigation of financial risks. The methodological guidelines of the European Systemic Risk Board (ESRB) assist in the country-specific application of the instrument. Moreover, the efforts of the Bank for International Settlements (BIS) in formulating a credit risk measurement system also constitute an important part of the range of instruments serving stability policies.

Although short and medium-term policies tend to be in focus in terms of policies fostering stability, long-term strategic measures have an equally important role, especially with respect to the convergence of emerging and developing countries where – based on the indicators calculated from the GDP proportionate data of consolidated bank balance sheet reports – the sophistication (depth) of the financial sector lags behind the level of development in advanced economies (*Table 4*).

Based on the indicators, between 1989 and 2009 the depth of the financial sector increased sharply in four advanced economies (Switzerland, Belgium, United Kingdom and USA), continued to improve in Japan, and in 2009 it reached an outstanding level (21.6 per cent) in Ireland within a short period of time. Among the emerging economies, such a level of development was only observed in Hong Kong, with Singapore ranking second in this regard. The global index stood at 4.2 per cent in 1989 and by 2009 it rose to 6.7 per cent. Emerging economies contributed to this figure by 0.3 per cent in 1989 and 1.2 per cent in 2009, while advanced economies contributed by 3.9 per cent and 5.5 per cent, respectively. The indexes show that – despite the progress achieved – further stability-strengthening monetary policy measures are needed in the financial sector of emerging countries to thwart the adverse effects of external shocks.

Short and medium-term analyses fostering global financial stability are based on a *global financial stability map* (Figure 9) (IMF 2016). According to the map, global stability improved somewhat in 2016. Macroeconomic risks remained unchanged. The continued monetary stimulus of central banks eased monetary and financial conditions, supporting a recovery in risk appetite. However, as a result of the recovery in commodity prices and external financial conditions and owing to the pickup in capital flows, market and liquidity risks remained elevated in an environment of extended realignment across major asset classes.

Figure 9
Global Financial Stability Map: Risks and Conditions



Despite lower short-term risks, medium-term risks were rising in 2016 as policymakers faced a wide range of vulnerabilities and new challenges. Credit risks, for example, were exacerbated by the sharp deterioration of banks' resilience in the low interest rate environment. In a broader sense, the biggest risk is posed by the USA's unbalanced politics and policy, which might lead to tighter-than-expected financial conditions, greater fluctuations and risk aversion. Another

challenge is posed by the worldwide trend of strengthening protectionism and hence, the deceleration in trade and economic growth. Consequently, attaining stability depends strongly on policymakers' ability to implement the right "mix" of national and global policies (*IMF 2017:1*).

2.5. International spillovers

This field merits thorough analysis not only because of the amplification of international spillover effects, but also because central banks sometimes do not appear to pay sufficient attention to harmonising these effects and the country-specific financial conditions in making monetary policy decisions. The various facets of this challenge are aptly illustrated by the shock effects emerging in recent years in relation to the global crisis.

A prominent manifestation of these effects is the convergence of *interest rates* at a rate that exceeded the pace of the changes in macroeconomic fundamentals. This is evidenced by the panel regression analysis conducted by (*Hofmann – Takáts 2015*) for a sample composed of 30 emerging and smaller advanced economies for the period of 2000–2014. The authors found (*Table 5*) that a 100 basis point change in the short-term (3-month) US interbank rate induces a 34 basis point change on average in the countries under review (*Column 1*). Spillovers were even stronger (59 basis points) for 10-year bond yields (*Column 2*). The VIX index – the other explanatory variable besides the US interest rates that is designed to capture global investment risk aversion – was another important influence on the interest rates concerned. A 100 basis point decline in the federal funds rate was associated with a 43 basis point downward shift in the policy rate of the sample countries compared to the level implied by the normative Taylor rule (*Column 3*). Moreover, when the examination of the effect of US policy rate was based not on the normative but descriptive Taylor-rule, then the impact was higher, 70 basis points on average, in emerging and smaller advanced economies (*Column 4*). Based on these findings, the USA's federal funds rate had a significant impact on the short and long-term interest rates of the 30 countries under review.

Table 5
Interest rate spillovers¹

Explanatory variable	Dependent variable			
	Change in 3-month interest rates ²	Change in 10-year bond yields ²	Interest rate deviation ³	Interest rate level ⁴
US rate	0.34***	0.59***	0.43***	0.70***
VIX	0.51***	0.21**	1.99***	1.54***
F-stat of US output and inflation ⁵	0.24	2.35*	20.80***	6.80***
F-stat of domestic output and inflation ⁵	17.18***	2.09	–	12.60***
R ²	0.25	0.26	0.45	0.82

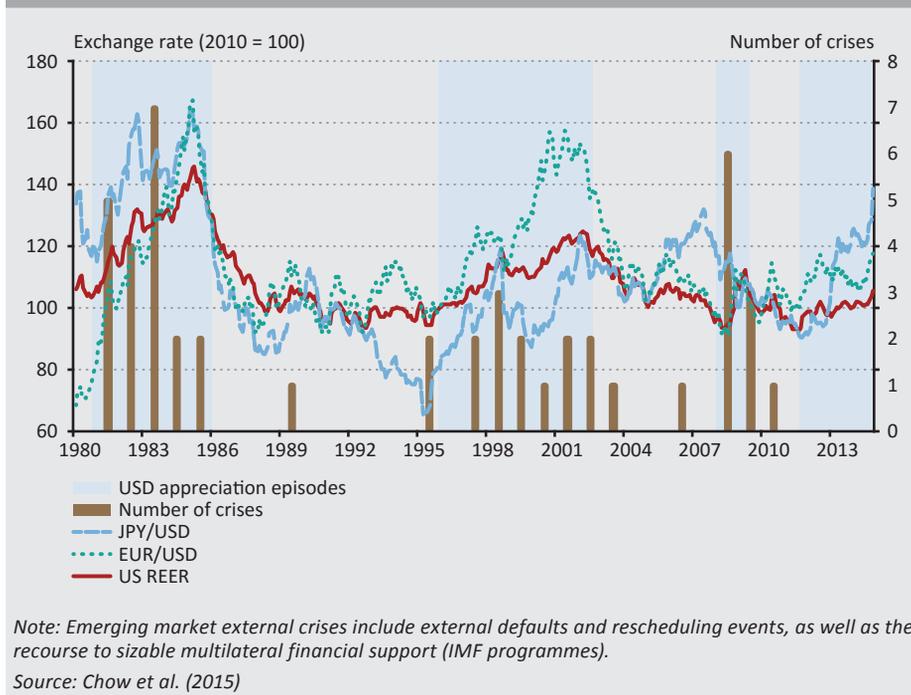
Note: ¹ Results from unbalanced fixed effects panel regressions for 30 emerging market and advanced economies (Australia, Brazil, Canada, Chile, China, Chinese Taipei, Colombia, the Czech Republic, Denmark, Hong Kong SAR, Hungary, India, Indonesia, Israel, Korea, Malaysia, Mexico, New Zealand, Norway, Peru, the Philippines, Poland, Russia, Singapore, South Africa, Sweden, Switzerland, Thailand, Turkey and the United Kingdom) for the sample period Q1 2000–Q4 2014. ² Panel estimation of $\Delta r_t^i = \beta_0 + \alpha_0^i + \beta_1 \Delta r_t^{us} + \beta_2 X_t^i + \varepsilon_t^i$ where Δr_t^i indicates the quarter-on-quarter change in economy i 's 3-month money market rate and 10-year bond yield and Δr_t^{us} is the change in the corresponding US rate; X includes the change in US real GDP growth and inflation, the log change in the VIX and the change in domestic real GDP growth and inflation. ³ Panel estimation of $i_t^i - Taylor_t^i = \beta_0 + \alpha_0^i + \beta_1 i_t^{us} + \beta_2 X_t^i + \varepsilon_t^i$ where i_t^i is the policy rate, $Taylor_t^i$ is the policy rate implied by a normative Taylor rule (calculated following Hofmann and Bogdanova, "Taylor rules and monetary policy: a global 'Great Deviation'?", BIS Quarterly Review, September 2012, pp 37–49), i_t^{us} is the federal funds rate and X includes US real GDP growth, US inflation and the (log) VIX. ⁴ Panel estimation of $i_t^i = \beta_0 + \alpha_0^i + \beta_1 i_t^{us} + \beta_2 X_t^i + \varepsilon_t^i$ where X includes US real GDP growth, US inflation and the (log) VIX as well as domestic inflation and the domestic output gap (calculated using a standard Hodrick-Prescott filter). ⁵ F-test of the null hypothesis that coefficients of the variables equal zero. ***/**/* denotes results significant at the 1/5/10 per cent level based on cluster-robust standard errors.

Source: Hofmann – Takáts (2015); BIS (2015)

Another example demonstrates spillovers from the appreciation of the US dollar in emerging and developing countries. This was examined by IMF staff (IMF 2015a: 12–16) for the periods 1980–1985, 1995–2001 and 2008–2009 when emerging economies faced financial crises (Figure 10). In the period 1995–2001, US dollar appreciation took place against a background of tighter US monetary policy relative to that prevailing in Europe and Japan. During the period, emerging countries improved their net international investment positions significantly with a parallel increase in foreign exchange reserves, and – also as a result of other important development factors – emerging and developing countries in general relied increasingly on domestic resources and reduced their dependence on foreign currency debt. Emerging economies were also able to increase the share

of domestic assets in public debt and to improve their international investment position from -40 per cent of GDP in 1995 to -25 per cent of GDP in 2013, with net debt improving by 20 percentage points of GDP and FX reserves by 10 percentage points of GDP.

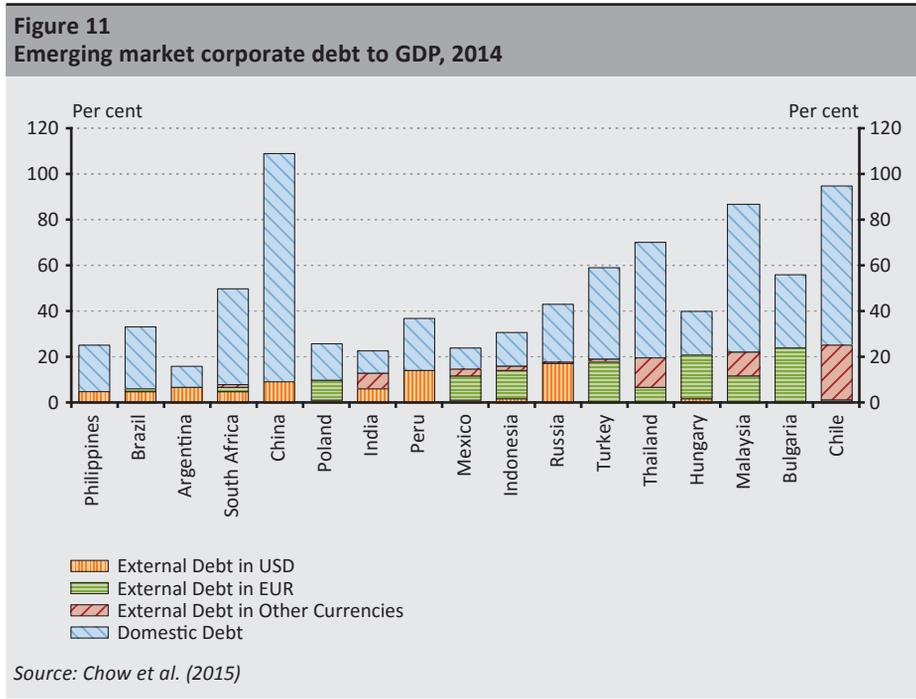
Figure 10
USD appreciation episodes



Despite these positive effects, appreciation of the US dollar significantly increases the vulnerability of some countries. While net external positions improved, gross positions made the economy more vulnerable in countries where foreign liabilities were especially large (Chile, Hungary, Malaysia, Poland and Thailand). These liabilities underpin that, in view of the intensification of financial globalisation and, in general, international effects, *a more frequent impact assessment of the implications of gross positions (“gross flows”) is called for in economic policy analysis, in addition to the net external positions monitored regularly by the countries and the professional literature.*

In addition, the currency composition of the FX debt position also involved vulnerabilities: naturally, the appreciation of USD affected those countries the most where dollar debt comprised the bulk of external liabilities (e.g. Turkey, China and Thailand).

In many countries, another negative implication of USD appreciation was a sharp increase in the *outstanding debt of non-financial corporations* in recent decades, with a parallel rise in foreign currency debt (Figure 11).



As shown in the Figure, an extremely large portion of Chinese corporate debt was denominated in the local currency. It should be borne in mind, however, that even USD-denominated Chinese debt amounted to over USD 1 trillion based on 2015 statistics. And while China’s USD reserves exceed this amount, these dynamics are concerning especially in light of the fact that the country’s foreign currency reserves have fallen by more than USD 1.2 trillion in recent years. Other emerging countries (primarily Chile, Malaysia, Thailand, Turkey, Russia, Bulgaria and Hungary) recorded high foreign currency debt ratios. In most of these countries, FX debt was largely denominated in US dollars, especially in Asia and Latin America but also in Turkey. In European emerging economies (Bulgaria, Hungary and Poland), the euro accounted for the lion’s share of total FX debt, which made these countries less exposed to the spillover effects of USD appreciation.

As shown by these two examples, *the monetary policy of the United States exerted a significant impact on the rest of the world*. Similarly, the monetary policies of other globally important countries (euro area, Japan and the United Kingdom) in

particular, may induce negative or positive effects (unfortunately mostly the former). The intensity and duration of these effects are influenced by three important factors:

- The openness of the communication cooperation and dialogue between the central banks of globally influential countries and the rest of the world. From this respect, the former group should assume a more active and initiatory role.
- The measures taken by individual countries to strengthen monetary policy transmission (e.g. in relation to non-performing loans and to the reduction or elimination of the output gap).
- The extent to which the applied policy mix – e.g. fiscal policy – supports economic growth and reduces the output gap. For example, in connection with the latter *Dabla-Norris et al. (2015)* pointed out that government expenditure on infrastructure development may boost economic growth and employment significantly in a low-demand environment. Of course, different policy mixes are needed in different countries and situations.

2.6. Institutional framework

The functioning of the IMFS is strongly affected by the operational efficiency of the institutional framework vested with international coordination. In the Bretton Woods system, the key institution in this framework was the IMF. The second amendment to its Articles of Agreement – effective from 1978 – redefined and expanded the responsibilities of the IMF in relation to the IMFS: accordingly, the IMF is expected to monitor the functioning of the system, to conduct analyses and surveys with a view to strengthening the system, and to maintain a forum of dialogue with regard to the system.

In addition to the IMF, from the mid-1970s to 1986–1987 the G5 and the G7 also played a prominent role in international coordination. Rather than making specific decisions, their role was mainly focused on formulating joint strategic opinions regarding the potential directions of economic development. This role was very successful in facilitating the spread of good practices. Some economists and policymakers (e.g. *Kenen 1988; Tobin 1987*) described this policy coordination as “regime preservation”, which prevents international economic and financial stability from being jeopardised by certain factors. An important achievement in this endeavour was the abovementioned *Plaza Accord* in 1985.

The 1997–1998 regional financial crisis in South East Asia and intensifying anti-globalist social movements prompted a striking turnaround in international coordination. One sign of this was China’s invitation to the 1999 G7/G8¹⁰ Cologne summit. The event was also made memorable by the fact that summit leaders

¹⁰ It was at this summit that China joined the G7, which therefore grew to become the G8.

voiced concerns about developments in global debt structure for the first time, especially in view of the growing significance of derivate transactions.

Another sign of the turnaround was the presentation of the IMF's 1998 report by the managing director, which underscored, as a critical issue in the IMF's policy, the need to *strengthen the architecture of the international monetary system* (*Address of the Managing Director 1998*). The Managing Director's 1999 Address, in turn, outlined the concept and feasibility of a "managed" global financial system where, as envisioned by De Larosière, logic and ethics prevail (*Address of the Managing Director 1999*). The report pointed out that while the liberalisation of capital movements had sizable benefits, it also created significant risks and emphasised that, due to their special traits, this principle should be applied differently in transition and emerging economies.

This notwithstanding, the IMF was strongly criticised for the passivity exhibited before the South East Asian crisis: the Fund failed to issue an explicit warning about the development of a meltdown. Although in line with its 2005 medium-term strategy, the IMF had launched – in justification of its activity, as it were – an experimental, multilateral economic policy consultation to further partnership between the participants in an effort to reduce global financial (current account) imbalances, due to inadequate support from the key partners (the United States and China) the experiment had failed. Indeed, if for no other reason than its professional expertise, the IMF would have been rightfully expected to work out alternative solutions to mitigate or eliminate the global financial asymmetry.

It was also due to these circumstances that instead of the IMF, *the G20 became the most important global forum for the coordination of economic policies* in the past one and a half decades, and that a handful of major emerging economies also joined the ranks of the key actors of global economic governance. One of the most memorable moments of this change was when the functioning of the IMFS became a central item on the G20's agenda in 2011, and the Palais Royal initiative – worked out in an effort to strengthen the IMF's activity by a group of world-renown economists and politicians under the lead of former IMF managing director Michel Camdessus – became a subject of extreme interest and heated debate (*Camdessus – Lamfalussy – Padoa-Schioppa 2011*). The authors put forward a number of recommendations on how to strengthen the IMF's activity (how to remedy "the lack of teeth" of its procedures) and on how Europe could (would) contribute to the needed reform of the international financial system.

The principal objective of the *international financial reform programme launched by the G20 in 2008* was to increase the resilience of the IMFS while preserving its open and integrated structure. The implementation of the regulatory reforms is

coordinated by the Financial Stability Board (FSB) established in 2009, with special focus on the improvement of four major areas:

- building resilient financial institutions (implementation of Basel III capital and liquidity standards, better risk analysis, etc.);
- ending the rescue of global systemically important – too-big-to-fail – banks;
- making derivatives markets safer and
- transforming shadow banking into resilient market-based finance – regulating the shadow banking system.

Another important objective of the reform programme is to intermediate financial flows and to facilitate the financial reforms of emerging markets and developing economies.

The FSB reports on the progress of the international financial reform programme's implementation in annual reports (*FSB 2015; 2016*). Based on the reports, the institutional framework and programmes developed by the G20 have already provided adequate incentives for the countries to comply with the rules and requirements of the IMFS.¹¹

3. Summary conclusions

Concurring with Jacques de Larosière's authoritative opinion, since the early 1970s the IMFS has been best described as a “*non-system*” rather than as a “*system*” (*De Larosière 2012*). Under such circumstances, the build-up of external imbalances perpetuated the accumulation of liquidity surpluses and – at least up until the outbreak of the 2008–2009 crisis – financial markets developed at a faster pace than the real economy. In this new world, everything has become so complex and diverse that the established intellectual premises of the IMF are no longer functioning in practice. Although on several occasions, the IMF attempted to launch a number of major reforms (*Kruger 2012; IMF 2015b*), due to the unravelling of financial markets and the lack of international support, they were never implemented.

Based on the analysis presented above on the operation of the IMFS, we can confirm that the *functioning of the system is inadequate and inefficient*. Below this finding is summarised through three relationship circles between the elements of the system.

¹¹ On a European level, the institution of the European Stabilisation Mechanism should also be mentioned, the establishment of which in 2012 represented an important step forward in the efforts to avert the spillover of a potential crisis within the euro area. It may serve as a “firewall”, as it were, to prevent the insolvency of a country from triggering panic reactions in other markets.

The first circle comprises capital flows, the volatility of exchange rates and the evolution of imbalances. Based on the development of the global *current account* imbalances shown in *Section 2.2*, we can conclude that countries running large deficits often lose the confidence of surplus countries, which may set in motion an exchange rate crisis between their currencies. An exception to this “rule” is the USA, which ran a substantial current account deficit for decades without giving rise to a similar crisis. As regards *cross-border capital flows* – the other important area of global disequilibrium – we need to point out that foreign currency debt may hurt borrowers significantly if their local currency is devalued and consequently, the repayment of their FX-denominated loans becomes far more expensive. This is exemplified by the 1997–1998 South East Asian exchange rate crisis. And if debtors are unable to repay their loans, the creditors will also suffer as their money will only be repaid, if at all, later on or only partially.

Despite these risks, capital flows were almost uniformly perceived as useful for the recipient country. Little mention was made about the challenges posed by the inflows to macroeconomic governance, the magnitude of the threats to financial stability or the impact on reserve accumulation or prudential measures. Also in view of the international spillover effects, policymakers failed to identify safer directions for capital flows or regulate the process within the framework of the IMFS.

Capital flows also contributed to the *extreme volatility of exchange rates*. Devaluation pressures – which improve the competitiveness of exports – posed a risk to indebted countries and to countries with large balance of payments deficits, because they may become the focus of policy decisions to such a degree that leads to a decline in internal demand. By contrast, the structural realignments in countries running balance of payments surpluses fuel even greater imbalances in deficit countries. And the often cited idea that the market (floating) exchange rate can resolve current account imbalances even in the short run has not been confirmed by practical experience. Hectic swings in the exchange rates, for the most part, did not even reflect the underlying economic trends and led to particularly large distortions in the economic development of small, open economies.

The second circle involves the attainment of the optimal level of international liquidity in a period of extreme volatility. From the aspect of the operation of the IMFS, too much or too little liquidity is no longer a theoretical but rather a practical, measurable (benchmark) issue. During the years preceding the global financial crisis, the Fed’s federal funds rate was lower than the level implied by the Taylor rule. This deviation suggests that these years were characterised by excess liquidity. This lends justification to the opinion (*Taylor 2007*) that the Fed’s monetary policy contributed to the overheatedness of the secondary mortgage market and hence, the outbreak of the global financial crisis. As the crisis unfolded, central banks attempted to resolve the drastic credit crunch with aggressive liquidity expansion,

which was reflected in steep interest rate cuts, i.e. quantitative easing (purchases of sovereign and corporate securities directly by the central bank). The resulting liquidity situation, however, demanded excessive self insurance, which materialised as large and costly capital formation. This process continued in the years following the crisis despite the fact that volatility-induced tensions had eased somewhat. Strengthening the global financial safety net and extending the scope of cooperation between central banks may assist in scaling down this method of self insurance.

The third circle is (economic and monetary policy) cooperation, the relationship between spillovers and diverging volatilities. Before the 2008–2009 crisis, this relationship, for the most part, was lackadaisical: international cooperation did not or only slightly facilitated the development of positive spillover effects and the mitigation of volatilities. This led to the build-up of dangerous systemic risks in the financial sector, which was also boosted by the widespread application of various technical innovations.

In *Section 2.5* we presented the correlation between the USA's federal funds rate and the policy rates of a number of significant countries, which indicated that the Fed's low interest rates played an important role in shaping the prices of products, services and assets in other countries. The spillover of this monetary effect incited banks and investors to take greater risks. This was reflected in the fact that lending and investment activities – primarily towards emerging smaller countries – grew at a faster pace than banks' capital. The sharp acceleration in international capital flows through the bond markets is especially evident in Latin American countries.

It is an important and progressive institutional change that willingness for cooperation between countries and financial institutions noticeably improved after the crisis. The *Mutual Assessment Process* (MAP) initiated by the G20 is a good example: under the MAP, with the participation of the IMF, participants analysed the countries' macroeconomic conditions and balance of payments positions in order to evaluate key imbalances.¹² However, there was no change in respect of the fact that the Fed's strongly influential monetary policy and deliberations on the recent interest rate hikes in the USA focused mainly on serving domestic economic requirements and interests. Through international spillovers, this continues to shape the monetary policy stance of other countries.

The functioning of the IMFS demonstrates that more comprehensive analysis work, global discussions and greater international efforts are required for a broad-based and complex enhancement of the system. Although numerous experts concluded that the system enhancements performed after the crisis had made the operation

¹² Note that in the Macroeconomic Imbalance Procedure applied by the European Commission external balance is analysed with more detailed indicators than those of the MAP process.

of the IMFS safer, additional radical changes are needed to avoid the imminent outbreak of new financial crises.

With that in mind, *the following four scenarios should be considered among the possible scenarios that may lay the foundations of the reform of the IMFS.*

- *Analysing the lessons of the 2008 global crisis in relation to the IMFS and aligning the resulting policy responses with the newly arising issues and challenges.* This requirement may be satisfied by the financial regulatory reform of the G20 (ensuring financial stability, promoting traditional bank financing and innovative financial sources, preventing systemic risks in the operation of financial institutions, strengthening the macroprudential policy framework). The implementation of these goals is progressing according to plan, and may serve as a starting point for a more comprehensive and thorough reform of the IMFS in the near future.
- *Reforming the IMFS with a focus on rule definition.* The automatic international regulatory mechanisms missing from the existing system would simultaneously serve the resolution of global imbalances and balance of payments deficits, the mitigation of negative spillover effects and the justified restriction of global capital movements.
- *Reforming the IMFS with the IMF in its changed role at the centre of global economy and finances.* This is in part based on the broad-based quota and governance reform adopted in 2010 and taking effect in 2016 which, while expanding the funds of the IMF, increases the weight of the G20 and also responds to the needs of emerging and developing countries. Another condition is the reduction of the USA's current excessive role.
- *Transforming the existing, USA-led, unipolar global economic system and the IMFS into a multipolar system.* So far, there have been no attempts – or indeed initiatives – in this regard on the part of any of the main power centres (USA, China, euro area). With regard to the radical reform of the IMFS, the transformation of the existing system into a multipolar system and the presentation of the positions of the new power centres in terms of the possible directions of the reform merit special attention. The latter is especially intriguing in light of the unprecedented clash of opinions between the United States and the European Union and the aspects of the already existing, new regional agreements (bilateral SWAP arrangements, Chian Mai Initiative, etc.).

References

- Address of the Managing Director (1998) to the Board of Governors.* Fifty-Third Annual Meeting. Washington, D.C. 6 October
- Address of the Managing Director (1999) to the Board of Governors.* Fifty-Fourth Annual Meeting. Washington, D.C.
- Báger, G. (2010): *Financial Globalization.* United Nations Institute for Training and Research. Geneva, p. 62
- Báger, G. (2011): *Magyarország integrációja a nemzetközi pénzügyi intézményekbe (Hungary's integration into the international financial institutions).* Akadémiai Kiadó. Budapest
- Bernanke, B. (2005): *The Global Saving Glut and the U.S. Current Account Deficit.* Speech delivered at the Sandridge Lecture, Virginia Association of Economists, Richmond, VA., 10 March. Online: <https://www.federalreserve.gov/boarddocs/speeches/2005/200503102/>
- BIS (2011): *Global liquidity – concept, measurement and policy implications.* Report submitted by an Ad-hoc Group established by the Committee on the Global Financial System. Bank for International Settlements, November
- BIS (2015): *85th Annual Report.* Bank for International Settlement. Basel, 28 June
- Boughton, J.M. (2009): *A New Bretton Woods?* International Monetary Fund, Finance and Development, Volume 46, Number 1
- Camdessus, M. – Lamfalussy, A. – Padoa-Schioppa, T. (2011): *Palais-Royal Initiative. Reform of the international monetary system: a cooperative approach for the twenty first century.* Online: http://global-currencies.org/smi/gb/telechar/news/Rapport_Camdessus-integral.pdf
- Catão, L.A.V. – Milesi-Ferretti, G.M. (2013): *External Liabilities and Crises.* IMF Working Paper 13/113, International Monetary Fund, Washington (Journal of International Economics)
- Coene, L. (2012): *A New International Monetary System?* In: Koeune, J.C. – Lamfalussy, A. (dir./eds.): *In Search of a New World Monetary Order.* P.I.E Peter Lang, Bruxelles, Bern, Berlin, Frankfurt am Main, New York, Oxford, Wien, pp. 161–166
- Chow, J. – Jaumotte, F. – Park, S.G. – Zhang, S. (2015): *Spillovers from Dollar Appreciation.* Online: <https://www.imf.org/external/pubs/ft/pdp/2015/pdp1502.pdf>
- Dabla-Norris, E. et al. (2015): *The New Normal: A Sector-Level Perspective on Productivity Trends in Advanced Economies.* IMF Staff Discussion Note, 15/03 (Washington: International Monetary Fund). Online: <http://www.imf.org/external/pubs/ft/sdn/2015/sdn1503.pdf>

- Edwards, S. (2005): *Is the Current Account Deficit Sustainable? And If not, How Costly Is Adjustment Likely to be?* NBER Working Paper No. 11541 (Cambridge, Massachusetts: National Bureau of Economic Research)
- De Larosière, J. (2012): *Can the IMF Become the Centrepiece of the International Monetary and Financial System?* In: Koeune, Jean-Claude and Lamfalussy, Alexandre (dir./eds.): *In Search of a New World Monetary Order*. P.I.E. Peter Lang, Bruxelles, Bern, Berlin, Frankfurt am Main, New York, Oxford, Wien, pp. 191–194
- Fleming, J.M. (1962): *Domestic Financial Politics Under Fixed and Under Floating Exchange Rate*. IMF Staff Papers 12, pp. 969–980
- FSB (2016): *Implementation and Effects of the G20 Financial Regulatory Reforms*. 31 August, 2nd Annual Report. Online: <http://www.fsb.org/wp-content/uploads/Report-on-implementation-and-effects-of-reforms.pdf>
- FSB (2015): *Implementation and effects of the G20 financial regulatory reforms*. Report of the Financial Stability Board to G20 Leaders. <http://www.fsb.org/wp-content/uploads/Report-on-implementation-and-effects-of-reforms-final.pdf>
- Goldberg, L.S. – Kennedy, C. – Miu, J. (2011): *Central Bank Dollar Swap Lines and Overseas Dollar Funding Costs*. FRBNY Economic Policy Review, May, pp. 3–20
- Hofmann, B. – Takáts, E. (2015): *International monetary spillovers*. BIS Quarterly Review, 13 September
- IMF (2007): *World Economic Outlook*. Washington, D.C., October
- IMF (2008): *World Economic Outlook*. Washington, D.C., April
- IMF (2014a): *Global Financial Stability Report. Risk Taking, Liquidity and Shadow Banking. Curbing Excess while Promoting Growth*. October 2014
- IMF (2014b): *World Economic Outlook*. IMF, October 2014
- IMF (2015a): *Spillover Report*. IMF
- IMF (2015b): *The Managing Director's Global Policy Agenda*. IMF
- IMF (2016): *Global Financial Stability Report*. IMF, October
- IMF (2017): *Global Financial Stability Report*. IMF, April
- Isard, P. (2005): *Globalization and the International Financial System. What's Wrong and What Can Be Done*. Cambridge University Press
- Kenen, P. (1988): *Managing Exchange Rates*. London, Routledge

- Kruger, A.O. (2012): *Sustainability and Reform of the International Monetary System*. In: Koeune, J.C. – Lamfalussy, A. (dir./eds.): *In Search of a New World Monetary Order*. P.I.E Peter Lang, Bruxelles, Bern, Berlin, Frankfurt am Main, New York, Oxford, Wien.
- Krugman, P.R. (2007): *Will There Be a Dollar Crisis?* *Economic Policy*, 22(51), pp. 435–467. <https://doi.org/10.1111/j.1468-0327.2007.00183.x>
- Kürthy, G. (2013): *Globális egyensúlytalanságok stock-flow konzisztens modellben (Global imbalances in a stock-flow consistent model)*. In: Pénz, világpénz, adó, befektetések (Money, world currency, tax and investments), pp. 45–78
- Mussa, M. (2004): *Exchange Rate Adjustments Needed to Reduce Global Payments Imbalances*. In: Bergsten, C.F. – Williamson, J. (ed.): *Dollar Adjustment: How Far? Against What?*
- Mundell, R. (1960): *The Monetary Dynamics of International Adjustment Under Fixed and Flexible Exchange Rates*. *Quarterly Journal of Economics*, 74, pp. 227–257. <https://doi.org/10.2307/1884252>
- Mundell, R. (1961a): *The International Disequilibrium System*. *Kyklos*, 14, Fasc. 2, pp. 153–172
- Mundell, R. (1961b): *Flexible Exchange Rates and Employment Policy*. *Canadian Journal of Economics and Political Science*, 27, pp. 509–517. <https://doi.org/10.2307/139437>
- Mundell, R. (1962): *The Appropriate Use of Monetary and Fiscal Policy for Internal and External Stability*. *IMF Staff Papers*, 12, pp. 70–79
- Sacomanni, F. (2012): *How to Deal with a Global Triffin Dilemma*. In: Koeune, J.C. – Lamfalussy, A. (dir./eds.): *In Search of a New World Monetary Order*. P.I.E Peter Lang, Bruxelles, Bern, Berlin, Frankfurt am Main, New York, Oxford, Wien
- Taylor, J.B. (2007): *Housing and Monetary Policy*. In: *Housing, Housing Finance, and Monetary Policy*, Proceedings of the Federal Reserve Bank of Kansas City Symposium, Jackson Hole, WY. September, pp. 463–476. <https://doi.org/10.3386/w13682>
- Thomas, C.P. – Marquez, J. – Fahle, S. (2008): *Measuring U.S. International Relative Prices: A WARP View of the World*. FRB International Finance Discussion Paper, No. 917 (Washington: Board of Governors of the Federal Reserve System)
- Tobin, J. (1987): *Agenda for International Coordination of Macroeconomic Policies*. In: Paul A. Volcker et al. (eds.): *International Monetary Cooperation: Essays in Honor of Henry C. Wallich*, *Essays in International Finance*, No. 169. Princeton, NJ: Princeton University, International Finance Section, Department of Economics, pp. 61–69

Equity-based Crowdfunding

Mónika Kuti – Zsolt Bedő – Dorottya Geiszl

This paper presents the latest results of international research on equity-based crowdfunding, and thus – according to the authors’ current knowledge – fills a gap in the Hungarian literature. Equity-based crowdfunding offers a fundraising opportunity through online platforms for small and medium-sized enterprises seeking alternative forms of funding for growth opportunities and innovation in this digital age. Investors interested in risk-adjusted financial return have several methods available to handle capital market friction and agent costs in the context of “crowdfunding”. Information signals on entrepreneurial quality are essential for seasoned investors when making capital allocation decisions in a crowdfunding campaign. In addition to the national legislative framework, platforms have many different rules, techniques and cultural norms that attempt to handle capital market imperfections.

Journal of Economic Literature (JEL) codes: F36, G19, O16

Keywords: equity crowdfunding, financial innovation, start-up financing

1. Introduction

Resources intermediated via information technologies open the door for new types of business models and roles. Online exchange, sharing and cooperation between Internet users allows entrepreneurs to recognise opportunities, control resources and create capacity through crowdsourcing, marketing (social brand and word-of-mouth) and crowdfunding (*Shneor – Flåten 2015*).

This paper focuses specifically on equity-based crowdfunding from among the many different business models available and presents its complex theoretical framework. The appearance of Hungary’s first crowdfunding portal for equity

Mónika Kuti is an assistant professor at the Faculty of Economics of the University of Pecs.

E-mail: kutim@ktk.pte.hu

Zsolt Bedő is an assistant professor at the Faculty of Economics of the University of Pecs.

E-mail: zsoltbodyo@ktk.pte.hu

Dorottya Geiszl is a student at the Faculty of Economics of the University of Pecs.

E-mail: dorottya.geiszl@gmail.com

This paper was written with support from the EFOP-3.6.1-16-2016-00004 Institutional Developments Fostering Intelligence Specialisation tender.

The Hungarian manuscript was received on 19 March 2017.

DOI: <http://doi.org/10.25201/FER.16.4.187200>

purposes (tokeportal.hu) attests to the topic's local relevance and its inherent business opportunities.

2. Equity-based crowdfunding

The shift in corporate funding towards cyberspace, the paradigm change seen in recent years and the shift of collective capital formation towards the digital realm is paving new paths. Equity financing offers the opportunity for early-stage fundraising and for bridging the funding gap for young, innovative startups. These financial innovations open the door for startups to obtain funding even if capital markets are not functioning. In the wake of the global economic crisis, business angels and venture capitalists have increased their investment activity and their preference has shifted to later-stage investments (*Block – Sandner 2009*). This trend has highlighted the timeliness and opportunities offered by equity-based funding, opening up new directions for research in the international literature. Equity-based crowdfunding (hereinafter: ECF) is likely to create great challenges for business angels and venture capitalists in the near future (*Vulkan et al. 2015*). Equity-based crowdfunding consists of entrepreneurs publishing open calls online for the sale of their equity shares in the hopes of attracting a large number of investors. The calls and investments take place through online platforms which provides tools such as the legal background, pre-screening and the handling of financial transactions. Equity-based crowdfunding is a model where investors acquire a stake in the company in the form of a share or share-like agreement (such as profit-sharing). Equity-based crowdfunding involves the sale of securities and is therefore significantly influenced by the legal environment of the host country (*Bradford 2012*).

In terms of the choice between traditional bank financing and crowdfunding, the latter offers financial benefits such as market feedback, product validation and community impacts. Imperfect information and moral hazard also affect companies' choice between the available business models of crowdfunding. In case of equity-based crowdfunding versus donation-based crowdfunding, financial return is important for investors alongside companies' future capacity to generate cash flow. In reward-based crowdfunding, backers tend to look at whether the idea's pitcher will be able to create and deliver the prepurchased product. For large-scale projects, the campaign-starting companies prefer equity crowdfunding and investors also factor in long-term profit; by contrast, the minimum threshold may be difficult to achieve with rewards-based crowdfunding in campaigns that require a high amount, so this format is generally preferred for smaller value, high-quality projects. If a product can reach a large customer base, reward-based crowdfunding enabling pre-purchases is preferred, while profit-sharing-based models are preferred for special niche products (*Miglo 2016*).

3. Capital market imperfections

The problem of information asymmetry is a central issue in equity-based crowdfunding. Researchers are actively investigating the correlations between the disclosure of information and successful fundraising. Startups send different signals to angel investors and venture capitalists than to small investors. Small investors are less sophisticated in terms of their financial knowledge and experience compared to venture capitalists, who have extensive knowledge on startups and the founding teams' valuation (Freear *et al.* 1994). Not every signal on startup quality is able to effectively handle information asymmetry. The effectiveness of signals depends on how perceivable they are and their cost, in other words how much investors take notice of them and understand the messages, and also depends on the costs of producing the signals not exceeding their benefits, in other words, they must be structured in such a way that dishonest signals do not pay (Connelly *et al.* 2011). Ahlers *et al.* (2015) investigate what project signals and entrepreneur quality attributes most spurred investors to pledge their financial resources in the context of equity-based crowdfunding. They found that projects that signal higher entrepreneurial quality and have perceptible characteristics are more likely to attract funding; the less detailed the information on the link between the offered business share and financial forecasts, i.e. the higher the uncertainty, the more careful potential investors will be when evaluating the startup; social capital and intellectual capital barely have an impact on campaign success. The hard signals issued by entrepreneur are discounted by investors while the hard information from investors significantly shapes campaign success; meanwhile, soft information from both entrepreneurs and investors reinforce network impacts (Estrin – Khavul 2016).

The fraction of equity retained by the entrepreneur or overcapitalisation is typically interpreted as a sign of entrepreneurial intent and a greater fundraising opportunity (Leland – Pyle 1977). If the long-term objective of fundraising is growth in the context of equity crowdfunding, the startup must strive to retain control in the course of issuance. The proposed proportion of equity is generally disclosed on the platforms. Retaining equity can essentially be interpreted as a perceivable signal (Ahlers *et al.* 2015). On the one hand, entrepreneurs who offer a larger portion of their startup through the platform are less likely to attract potential investors; on the other hand, the success rate is higher for entrepreneurs with larger social networks, as this decreases uncertainty and increases visibility (Vismara 2015b).

Researchers are also investigating information asymmetry and the question of location. Seed-stage equity-based and platform-based interregional trading creates opportunities for startups, investors and legislators, contributing to more effective capital allocation. Angel investments are geographically localised; early-stage investors focus on local transactions. Although there is home bias in equity-based crowdfunding that characterises investors who are similar to business angels, there

is no such behaviour among sufficiently diversified investors (*Hornuf – Schmidt 2016*). A common explanation for the absence of remote angel investments is information asymmetry. Although the Internet breaks down certain types of information constraints, it is incapable of transmitting face-to-face interactions such as the founding team's grit, determination, interpersonal dynamics, and trustworthiness (*Agrawal et al. 2015*). The face-to-face personal communication characteristic of angel investors and venture capitalists (*Moritz et al. 2014*) is replaced on these platforms by pseudo-personal communication.

Equity-based crowdfunding syndicates made up of leading investors are able to tackle the issue of information asymmetry. These business formations present the transactions to masses of potential backers (*Agrawal et al. 2015*) and link the face-to-face due diligence and monitoring that is characteristic of the off-line realm with online global access. The introduction of syndicates in equity crowdfunding is a relatively new phenomenon, available on AngelList since 2013. This platform allows individuals, angel groups and private equity funds to form syndicates and boasts a far higher rate of successful syndicated transactions compared to non-syndicated transactions. Venture capitalists therefore tend to prefer the syndicate strategy to increase the number of incoming transactions, improve brand reputation and provide sound arguments to their key partners with a view to concluding effective agreements (*Coppey 2016*). A syndicate is essentially a pre-screening mechanism prior to the subsequent non-syndicated investment stages. Syndicate leaders are generally well-connected through investment channels and business opportunities, which creates conflict of interest with unsophisticated small investors seeking to invest in startups. There is thus a trade-off in syndicated equity crowdfunding between access to deals and conflict of interest (*Itenberg – Smith 2017*). The authors demonstrated that syndicated investment offer rounds are riskier, yield slower return, and startups in which a syndicate leader invested earlier perform worse.

In the case of equity-based crowdfunding, a kind of *counterselection* resulting in a sub-optimal balance can be observed, i.e. a process in which the price and quantity of products and services change within a given market because one counterparty possesses information that the other party cannot obtain even at a reasonable cost. According to *Agrawal et al. (2015)*, neither ordinary investors nor startup founders are able to correct information asymmetry at a reasonable cost in the context of crowdfunding. Ordinary investors typically lack the time, money and expertise to be involved in the daily management and supervision of the startups they fund. On the other hand, startup founders have a better understanding of what they are doing and their market prospects, but have limited resources to communicate with the investors, to widely disseminate information and to prepare detailed financial reports. *Moral hazard* is when the party which incurs the costs has to assume disproportionate risk. Once a crowdfunding campaign ends, backers do not have

effective means to force startup founders to perform in line with the promised conditions. At this point, the acquired equity belongs to the startup, the bargaining position has shifted between opportunistic startups and investors, while ordinary investors assume the risk of loss.

The examination of campaign updates introduced a dynamic approach into research following the static analysis of information asymmetry. Startups can increase the probability of their equity-based crowdfunding campaign succeeding by disclosing up-to-date information. It is thus useful to know which types of updates are effective. Updates related to the number of investments made by the crowd and the amount of investments collected by startups have a positive impact on campaign success with a few days' delay (*Block et al. 2016*). According to the authors, this impact increases with the number of words featured in the updates, but decreases in parallel with the number of updates. They pointed out that updates related to the startup team (qualifications, age and personal interests), the business model (expected market, business idea, future business orientation and expansion expectations), co-creation projects and product development no longer have a major impact on the level of success. By contrast, updates on new funds (e.g. from angel investors, venture capitalists or tenders), the development of the business related to increases in sales revenues and favourable customer feedback as well as the promotional campaigns led by the startup, with networking taking place on social media, have a positive impact. Businesses release significantly more updates during the funding campaign than afterwards and publish more information if the campaign did not reach the funding objective or if the remaining funding period is short (*Dorfleitner et al. 2017*). According to Dorfleitner et al., businesses use some linguistic tools in their updates that strengthen group identity and group cohesion, and the probability of updates is higher when there is strong competition among the crowdfunding campaigns running in parallel.

The success rate is higher for those businesses that are also supported by professional investors (business angels and venture capitalists) and which obtained some grant money, or have protected their intellectual capital in the form of patents, trademarks or copyrights – pointed out *Ralcheva and Roosenboom (2016)*, who also emphasised that if the business hired a consulting body, appointed nonexecutive directors or involved consultancy firms, or if the owner intends to keep a larger share of the business after the funding, a successful campaign is more probable. According to some other studies, the campaign's success is determined by the pre-filtered campaign characteristics and the private and public networks, i.e. the investment decision criteria applied by traditional venture capitalists and angel investors are less relevant (*Lukkarinen et al. 2016*).

Early supporters play a key role in the success of crowdfunding campaigns on donation-based platforms (*Burtch et al. 2013*), reward-based platforms (*Colombo*

et al. 2015) and debt-based platforms (*Zhang – Liu 2012*). Empirical evidence to support this in the case of equity-based crowdfunding has been missing for a long time. On equity-based crowdfunding platforms, the names of individual investors are public and immediately accessible for every single offering which, in turn, affects the behaviour of future investors. Creating momentum in an early stage is a success factor, as the information can be evaluated based on the list of previous investments of early supporters which may convince future investors. Those investors who made their profile public often have higher qualifications, more project specific industry experience and a longer list of investments made through the platform. Additionally, their offering share compared to other investors is larger during the initial days of the campaign and then declines over time (*Vismara 2015a*). According to the author, information cascades are being formed as the presence of investors with a public profile attracts additional investors as the result of the “success brings more success” process. The psychological concept of observation-based social learning (*Bandura 1977*) appears in the financial literature as information cascades (*Welch 1992*). Two conditions must exist for the information cascade to occur: uncertainty and sequentiality (*Bikhchandani et al. 1992*). Uncertainty appears more on the demand side of equity-based crowdfunding, where investors are less capable of overcoming information asymmetry. In this case, the signals sent by additional investors, i.e. the mass impact, becomes significant. In the course of crowdfunding, subsequent investors learn from observing the behaviour of previous supporters.

The dynamics of crowdfunding is L-shaped on the platforms operating on “first come first served” basis, while it is U-shaped on the platforms operating on “sealed offer and secondary price auction” basis. (*Hornuf – Armin 2015*). Therefore, we have evidence supporting the idea that investors make their decisions based on the information updates provided by the entrepreneur and based on the investment behaviour and comments of other crowd investors. Campaign attributes, the level of investor sophistication, the process of granting funds, herd behaviour and stock exchange volatility determine the extent to which investors are willing to pay for the cash flow rights of startup businesses (*Hornuf – Neuenkirch 2015*). It is difficult to tell to what extent crowdfunders become small business angels: sometimes they invest together with professional investors, but sometimes they are the competitors of business angels (*Hornuf – Schwienbacher 2014*).

In terms of their composition, crowd investors cannot be considered as a homogeneous community. *Wallmeroth (2016)* revealed that strategic investors investing a larger part of their own assets are less likely to become investors who later return to the platform and are more likely to invest in businesses that will not go bankrupt after the campaign. In his opinion, it is likely that investors offering large amounts are sophisticated, professional investors as opposed to the crowd of less sophisticated investors sporadically investing small amounts. *According to*

Brown and Davis (2017), equity crowdfunding is characterised in general by the “all or nothing” funding style, the scarcity of investor profit and the crowd of non-cooperating investors.

In the case of equity crowdfunding, nonfinancial motivations are not typical, and investors are more likely to redeem after the campaign their intention to make a financial commitment expressed during the campaign than in the case of the reward-based model (*Cholakova – Clarysse 2014*). As for crowdfunding providing financial return, crowdfunders are typically young men with degrees motivated by the acquisition of business share and by excitement (*Daskalakis – Yue 2017*). On the markets of equity crowdfunding, 93 per cent of investors are men, which can be explained more by the gender difference related to risk avoidance than by the differences in terms of excessive self-confidence (*Hervé et al. 2016*). One interesting observation is that the balanced presence of genders is the highest for equity-based crowdfunding compared with all other forms of business funding, and furthermore, that gender determines the nature of the relationship between demand and offer of capital: According to *Vismara et al. (2016)*, projects of woman-led businesses have twice as many female investors than those helmed by men. The authors showed that although the great majority of investors are men, women invest 34 per cent more in general.

The literature also addresses the performance indicators after the campaign. One year after a successful campaign, the businesses’ revenues and the growth rate of their sales increased, which shows a positive link with the large number of investors during the campaign, while the profit dropped in general. According to *Décarre and Wetterhag (2014)*, these trends resemble the one seen when business ventures are investing in companies. They also emphasise that a PR effect can be felt, resulting from the fact that these businesses also use the campaigns for marketing and promotional purposes. Based on their research of equity crowdfunding platforms in the UK between 2011 and 2015, *Signori and Vismara (2016)* found that only 10 per cent of the businesses failed, and 30 per cent of them were able to involve additional funding with one or more seasonal share issuances, while investors’ average expected return was 8.8 per cent.

4. Equity-based crowdfunding platforms

Social networks and online platforms create new opportunities to raise funds for businesses and enable nonprofessional investors to place their funds without the use of financial intermediary systems. The acceleration and scaling of early funding opportunities within business funding schemes via Internet platforms can be exploited, taking account of the given regulatory system, which creates a balance between capital formation and investor protection. The platform is an architecture

which enables the exploitation of the signals and network effects in a reduced transactional cost environment (*Estrin – Khavul 2016*). The transactional interface for ECF must provide the legal background and the pre-screening of businesses, and allow financial transactions to be made (*Ahlers et al. 2015*). Intragroup and intergroup moderation is possible on the platform acting as an intermediary interface (*Belleflamme et al. 2015; Viotto 2015*). ECF is a multiple-phase process where targeted communication is needed for business ideas to be convincing on the platform and for investors.

According to *Löher (2016)*, the three important areas of the intermediary role include a) project pre-screening, where the platform relies on its own networks thanks to access to deal offers, which has the objective of reducing the cost of deal-seeking; b) the use of standardised contracts that comply with the legal background; and c) communication to reduce information asymmetry and promote a uniform campaign strategy. These platforms essentially search, assess and structure investment opportunities.

Brown et al. (2016) formulated the following correlations in terms of crowdfunding and platform efficiency:

1. The equity-based crowdfunding platform must simultaneously ensure that entrepreneurs can lawfully raise capital, and that investors can lawfully access investment opportunities.
2. It is more naïve investors narrow-mindedly following signals and basing their decisions on mass wisdom and not sophisticated investors who are needed for the equity-based crowdfunding platform to be efficient.

Lower quality businesses with a smaller social network are forced to turn to the crowd if they are unable to obtain funds from friends, family members, business angels or from venture capitalists through traditional early-stage funding. To reduce counter selection and moral hazard, in addition to legislation, various platform rules, self-regulation forms, technical features and cultural norms were born, which influence what quality and what types of startups are attracted by the platform. Investment ceilings, fundraising ceilings, limited disclosure obligations, platform investments, crowd audits, additional disclosures and performance transparency, and milestone-based financing are the rules and methods that minimise the chance of market failure (*Catalini et al. 2016*). The differentiating characteristics of high quality startup businesses and the renowned portals functioning as intermediaries may represent the way out of the Lemons Problem (*Ibrahim 2015*).

The platforms channel the flow of investments with standardised contracting practice. The platforms make various business offers for the rights to shareholder's equity, profit sharing and future share capital. The offering of non-equity securities is aimed at short-term profit sharing not including shareholder rights. In parallel, the platform also discloses a timing for the future investment rounds and strives to prevent that businesses maintain a costly, long-term relationship with a great number of small investors. The rights to future equity securities offered on the platform last until the occurrence of some future investment event. Future equity capital offers postpone the acquisition of shareholder's rights until the point when the high-growth startup becomes the target of acquisition, IPO or risk capital financing when crowd investors receive their ownership share. In the meantime, crowd investors do not have shareholder's rights, which would entail a number of costly, time-consuming risks and obligations. The link between the features of the early stage startup financing and the transparent investor agreement comes from the Simple Agreement for Future Equity, SAFE, as a deferred equity instrument. Another solution for the same issue is the KISS or Keep It Simple Security, which is essentially a convertible bond according to which the startup business promises the investor to repay the capital with some interest. KISS turns into a property right once the startup reaches certain milestones, such as acquisition, IPO or the involvement of risk capital. Therefore, the difference between KISS and SAFE is that if the investment is not converted into a property right, then in case of SAFE, the investors lose their capital while in the case of KISS, they gain it back through debt service. According to *Wroldsen (2017)*, there are great opportunities in these two simplified contracting forms, but interestingly investors benefit more from the collective protection provided by social media than from the formalised contractual rights.

On the German crowdfunding market, the subordinated loan providing profit sharing became dominant as an investment form. According to *Klöhn et al. (2016)*, the most recent trends include that the minimum maturity time went from 4 years to 6-8 years, lately offering fixed interest payment, and investors are also entitled to profit sharing. The number of contracts granting the right to share in exit revenues increased recently, while veto rights have been discontinued, and dilution is now possible for next round investments. On German equity crowdfunding platforms, crowd investors greatly depend on venture capital funds because they do not hold any share capital in the businesses. Such venture capital funds have a vested interest in crowding out small investors through their mezzanine financing contracts; in response to this trend, the platforms created contracts that promote the coexistence of these two groups (*Hornuf – Schmitt 2016*).

5. Conclusions

In recent years, equity-based crowdfunding platforms started to emerge once the legislative and regulatory environment had been created. Compared to other crowdfunding business models, investors are driven by financial return and entrepreneurs come up with larger investment projects compared to the average project size of reward-based crowdfunding. A number of success factors have been identified, including signs of strong entrepreneurial skills, the internal proportions of business share offering, the retention of ownership, the availability of hard information on investors, the size of the entrepreneurs' social networks and updates identifying future growth opportunities. When creating equity-based crowdfunding platforms, attention must be paid to the concordance of pre-screening, deal offering and the processes defined by the standardised contracts, compliance with the regulatory environment and the consideration of switching effects between network effects strengthening platform reputation. Off-line investment agents have a beneficial effect on the platforms operating in the given region. Striving to establish best practices requires constant innovation which can be observed in the case of experiments with syndicate strategies and also in the case of the evolution of the various securities and rights serving the reduction of conflict of interest between crowd investors and sophisticated investors.

The ever increasing development of digital technologies, blockchain and bitcoins with the so-called Initial Coin Offering deals propose new directions in the world of business investments through platforms the analysis of which goes far beyond the framework of this study.

References

- Agrawal, A. – Catalini, C. – Goldfarb, A. (2015): *Are Syndicates the Killer App of Equity Crowdfunding?* MIT Sloan Research Paper No. 5126-15; Rotman School of Management Working Paper No. 2569988. <http://dx.doi.org/10.2139/ssrn.2569988>
- Ahlers, G.K.C. – Cumming, D.J. – Guenther, C. – Schweizer, D. (2015): *Signaling in Equity Crowdfunding*. *Entrepreneurship Theory and Practice*, 39(4): 955–980. <https://doi.org/10.1111/etap.12157>
- Bandura, A. (1977): *Social Learning Theory*. Englewood Cliffs, NJ: Prentice Hall.
- Belleflamme, P. – Omrani, N. – Peitz, M. (2015): *The Economics of Crowdfunding Platforms*. *Information Economics & Policy*, 33 (Dec.): 11–28. <https://doi.org/10.1016/j.infoecopol.2015.08.003>

- Bikhchandani, S. – Hirshleifer, D. – Welch, I. (1992): *A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades*. Journal of Political Economy, Vol. 100, No. 5 (October): 992–1026. <https://doi.org/10.1086/261849>
- Block, J.H. – Hornuf, L. – Moritz, A. (2016): *Which Updates During an Equity Crowdfunding Campaign Increase Crowd Participation?* <https://doi.org/10.2139/ssrn.2781715>
- Block, J. – Sandner, P. (2009): *What is the Effect of the Financial Crisis on Venture Capital Financing? Empirical Evidence from US Internet Start-ups*. Venture Capital – An International Journal of Entrepreneurial Finance, Vol. 11: 295–309. <https://doi.org/10.1080/13691060903184803>
- Bradford, S.C. (2012): *Crowdfunding and the Federal Securities Laws*. Columbia Business Law Review, Vol. 2012, No. 1.
- Brown, D.C. – Davies, S.W. (2017): *Financing Efficiency of Securities-Based Crowdfunding*. <https://doi.org/10.2139/ssrn.2692828>
- Brown, D.C. – Davies, S.W. (2016): *Equity Crowdfunding: Harnessing the Wisdom of the Crowd*. <https://pdfs.semanticscholar.org/6382/9449bc90e899e84ee649a31fd2a87d326476.pdf>
- Burtch, G. – Ghose, A. – Wattal, S. (2013): *An Empirical Examination of the Antecedents and Consequences of Contribution Patterns in Crowd-funded Markets*. Information Systems Research, Vol. 24, No. 3: 499–519. <https://doi.org/10.1287/isre.1120.0468>
- Catalini, C. – Fazio, C. – Murray, F. (2016): *Can Equity Crowdfunding Democratize Access to Capital and Investment Opportunities?* MIT Innovation Initiative Lab for Innovation Science and Policy Report. <https://doi.org/10.2139/ssrn.2780551>
- Cholakova, M. – Clarysse, B. (2015): *Does the Possibility to Make Equity Investments in Crowdfunding Projects Crowd Out rewards-based Investments?* Entrepreneurship Theory and Practice, Vol. 39: 145–172. <https://doi.org/10.1111/etap.12139>
- Colombo, M.G. – Franzoni, C. – Rossi Lamasra, C. (2015): *Internal Social Capital and the Attraction of Early Contributions in Crowdfunding*. Entrepreneurship Theory and Practice, Vol. 39, No. 1: 75–102. <https://doi.org/10.1111/etap.12118>
- Connelly, B.L. – Certo, S.T. – Ireland, R.D. – Reutzel, C.R. (2011): *Signaling Theory: A Review and Assessment*. Journal of Management Vol. 37: 39–67. <https://doi.org/10.1177/0149206310388419>
- Coppey, L. (2016): *From Value-Added VCs to Equity Crowdfunding Syndicates: the New Platforms of the Venture Capital Industry*. MIT Sloan Management Department. <http://hdl.handle.net/1721.1/104539>.

- Daskalakis, N. – Yue, W. (2017): *User's Perceptions of Motivations and Risks in Crowdfunding with Financial Returns*. <https://doi.org/10.2139/ssrn.2968912>
- Décarre, M. – Wetterhag, E. (2014): *Uncovering the Outcomes of Equity Crowdfunding: Post-Funding Outcomes of Equity Crowdfunded Firms in Europe*. <https://doi.org/10.2139/ssrn.2545276>
- Dorfleitner, G. – Hornuf, L. – Weber, M. (2017): *Dynamics of Investor Communication in Equity Crowdfunding*. Max Planck Institute for Innovation & Competition Research Paper No. 17–06. <https://doi.org/10.2139/ssrn.2962951>
- Estrin, S. – Khavul S. (2016): *Academy of Management of Proceedings* (Meeting Abstract Supplement) 13036.
- Freear, J. – Sohl, J.E. – Wetzel Jr., W.E. (1994): *Angels and Non-Angels: Are there Differences?* Journal of Business Venturing, Vol. 9: 109–123. [https://doi.org/10.1016/0883-9026\(94\)90004-3](https://doi.org/10.1016/0883-9026(94)90004-3)
- Hervé, F. – Manthé, E. – Sannajust, A. – Schvienbacher, A. (2016): *Investor Motivations in Investment-Based Crowdfunding*. <https://econpapers.repec.org/paper/haljournal/hal-01452026.htm>
- Hornuf, L – Schmitt, M. (2016): *Success and Failure in Equity Crowdfunding*. CESifo DICE Report Vol. 14, No. 2: 16–22.
- Hornuf, L. – Neuenkirch, M. (2015): *Pricing the Value of Cash Flow Rights in Crowdfunding: An Analysis of Investment Backers*.
- Hornuf, L. – Schmitt, M. (2016): *Does a Local Bias Exist in Equity Crowdfunding?* Max Planck Institute for Innovation & Competition Research Paper No. 16–07. <https://doi.org/10.2139/ssrn.2801170>
- Hornuf, L. – Schwienbacher, A. (2014): *Crowdfunding – Angel Investing for the Masses?* Handbook of Research on Venture Capital: Volume 3. Business Angels, Forthcoming. <http://ssrn.com/abstract=2401515>. Downloaded: 18 August 2016.
- Hornuf, L. – Schwienbacher, A. (2015): *Funding Dynamics in Crowdfunding*. <https://ideas.repec.org/p/zbw/vfsc15/112969.html>. Downloaded: 18 August 2016.
- Ibrahim, D.M. (2015): *Equity Crowdfunding: A Market for Lemons?* Minnesota Law Review, Vol. 100; William & Mary Law School Research Paper No. 09–292. <https://doi.org/10.2139/ssrn.2539786>
- Itenberg, O. – Smith, E.E. (2017): *Syndicated Equity Crowdfunding: The Trade-Off between Deal Access and Conflicts of Interest*. Simon Business School Working Paper No. FR 17–06. <https://doi.org/10.2139/ssrn.2933822>

- Klöhn, L. – Hornuf, L. – Schilling, T. (2016): *Financial Contracting in Crowdfunding: Lessons from the German Market*. <https://doi.org/10.2139/ssrn.2839041>
- Leland, H.E. – Pyle, D. (1977): *Informational Asymmetries, Financial Structure and Financial Intermediation*. *Journal of Finance*, Vol. 32: 371–387. <https://doi.org/10.2307/2326770>
- Löher, J. (2016): *The Interaction of Equity Crowdfunding Platforms and Ventures: An Analysis of the Preselection Process*. *Venture Capital*, Forthcoming. <https://ssrn.com/abstract=2875673>.
- Lukkarinen, A. – Teich, J.E. – Wallenius, H. – Wallenius, J. (2016): *Success Drivers of Online Equity Crowdfunding Campaigns*. *Decision Support Systems*, Vol. 87: 26–38. <https://doi.org/10.1016/j.dss.2016.04.006>
- Miglo, A. (2016): *Crowdfunding: Balancing Imperfect Information and Moral Hazard Considerations*. <https://doi.org/10.2139/ssrn.2793631>
- Moritz, A. – Block, J.H. – Lutz, E. (2014): *Investor Communication in Crowdfunding: A Qualitative-Empirical Study*. <http://ssrn.com/abstract=2462282>. Downloaded: 18 August 2016.
- Ralcheva, A. – Roosenboom, P. (2016): *On the Road to Success in Equity Crowdfunding*. <https://doi.org/10.2139/ssrn.2727742>
- Shneor, R. – Flåten, B. (2015): *Opportunities for Entrepreneurial Development and Growth through Online Communities, Collaboration and Co-Creation*, In: Kaufmann, H.R. and Shams, R.: *Entrepreneurial challenges in the 21st century*, Chapter 11, Basingstoke UK: Palgrave Macmillan. <http://ssrn.com/abstract=2684712>. Downloaded: 18 August 2016.
- Signori, A. – Vismara, S. (2016): *Returns on Investments in Equity Crowdfunding*. <https://doi.org/10.2139/ssrn.2765488>
- Vismara, S. (2015a): *Information Cascades Among Investors in Equity Crowdfunding*. <https://doi.org/10.2139/ssrn.2589619>
- Vismara, S. (2015b): *Equity Retention and Social Network Theory in Equity Crowdfunding*. <https://doi.org/10.2139/ssrn.2654325>
- Vismara, S. – Benaroi, D. – Carne, F. (2016): *Gender in Entrepreneurial Finance: Matching Investors and Entrepreneurs in Equity Crowdfunding*. Forthcoming in Link, A. (eds.): *Gender and Entrepreneurial Activity*. Cheltenham, UK: Edward Elgar. <https://ssrn.com/abstract=2833946>
- Vulkan, N. – Astebro, T.B. – Fernandez, S.M. (2015): *Equity Crowdfunding: A New Phenomena*. Saïd Business School WP 2015-21. <https://doi.org/10.2139/ssrn.2700236>

Wallmeroth, J. (2016): *Investor Behavior in Equity Crowdfunding*. <https://doi.org/10.2139/ssrn.2881394>

Welch, I. (1992): *Sequential Sales, Learning, and Cascades*. *Journal of Finance*, Vol. 47: 695–732. <https://doi.org/10.1111/j.1540-6261.1992.tb04406.x>

Wroldsen, J. (2017): *Crowdfunding Investment Contracts*. 11 *Virginia Law & Business Review* (Spring, Forthcoming). <https://ssrn.com/abstract=2844771>

Zhang, J. – Liu, P. (2012): *Rational Herding in Microloan Markets*. *Management Science*, Vol. 58, No. 5: 892–912. <https://doi.org/10.1287/mnsc.1110.1459>

Presentation of the Hungarian Publications of the Magyar Nemzeti Bank Published between 1950 and 1990 from a Visual Communication Perspective

Ákos Vajas

1. Introduction

Upon hearing the name ‘Magyar Nemzeti Bank’, most people unwittingly think of the main building near Szabadság tér in Budapest. This comes as no surprise, since it is a distinctive, one-of-a-kind building, and its architectural style, the groups of statues on its façade, and the windows adorning the interior are all astonishing works of leading architects, sculptors and applied artists. The building has housed the Bank’s professional activities for many decades and has symbolic significance in Hungarian and international financial and economic policy. Despite the fact that the Bank’s main building has been presented in many books through the painstakingly meticulous work of various authors, the professionals playing a part in the Bank’s past, their community and engagement as well as the Bank’s publications, i.e. the journals circulated internally or externally that provided a publishing opportunity for everyday information and experts’ serious professional works, are hardly mentioned. The various publications are interesting and valuable not only because of their visual appearance, but also because they represent a real rarity as regards to their content. If we go beyond the surface, we can take a peek behind the curtain and get to know, understand and in a sense relive all of the events in the Bank’s distant past (e.g. personal recollections from during and after the Second World War, the 1956 Revolution and its aftermath within the Bank). The accounts as well as the illustrations, caricatures and drawings found here are of exceptional quality and poke fun at the everyday events at an organisation and a discipline (economics) as a whole. Furthermore, sometimes we can also see the artists’ honest self-reflection about the social and political structure and atmosphere between 1945 and 1990. This short description, which is far from comprehensive, was compiled on the one hand to pique readers’ interest in visual communication, a topic rarely in the focus in

Vajas Ákos is a project coordinator at the Magyar Nemzeti Bank. E-mail: vajasa@mnbb.hu

The Hungarian manuscript was received on 11 September 2017.

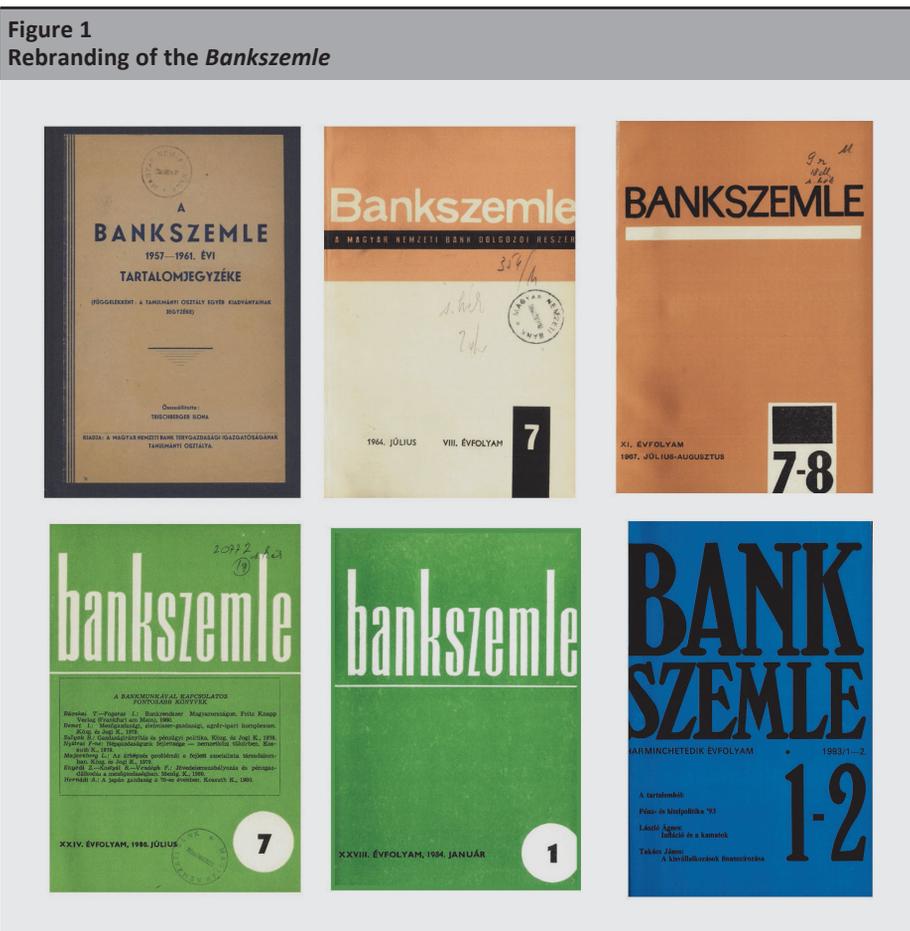
DOI: <http://doi.org/10.25201/FER.16.4.201214>

economics, and on the other hand to pay tribute to the *Bankszemle* (Bank Review), and commemorate it on the 60th anniversary of its first publication.

2. Bankszemle

The *Bankszemle*,¹ first published in June 1957 (Figure 1) was the scientific journal of the Magyar Nemzeti Bank, first prepared for internal use and circulation only. The publication was intended for professionals and published in 350 copies, which was gradually increased to 1,000. On account of its private circulation, it was written and read mainly by the Bank's staff. Later, in parallel with rebranding and page-setting innovations, circulation of the journal was widened, thereby reaching several financial institutions, organisations and companies, and professionals from outside

Figure 1
Rebranding of the *Bankszemle*



¹ The *Bankszemle* was published until 2001, when it was replaced by the *Financial and Economic Review* as a major Hungarian scientific journal in the fields of economics and finance.

the Bank were also allowed to publish in it. As a result, the number of subscriptions to the journal jumped from around 1,000 to over 3,000, which had made it a major journal in the country. We can see that on the outside, the publication changed somewhat over the decades, but its topics remained the same. The periodical's new look and especially the page-setting of the publication's inside pages were intended to enhance the quality and prestige of the professional content published.² It appeared as an 80-page publication for a long time, and then the number of pages was reduced to 64, thereby making the whole work leaner. Thanks to new page-setting technologies, it was possible to increase the length of the contents, and the readability of the texts was aided by modification of the page layout (e.g. two-column pages).

3. The *Forint* magazine

The *Forint*, the journal of the staff of the Magyar Nemzeti Bank, was first published on 15 August 1950.³ Its forerunner was called *Bankújság* (Bank Journal),⁴ and its sister papers were the *Fillér*⁵ and the *Bankó* (Bank Note) (Figure 2–3).⁶ Due to paper shortages, the *Forint*'s publication was interrupted in the autumn of 1951, and it only became available again in 1957.⁷ This internal-circulation magazine was one of the main platforms for the Bank's active and retired employees for social interaction and voicing their opinion. The paper entitled *Magyar Nemzeti Bank Dolgozóinak Lapja* (Journal of the Staff of the Magyar Nemzeti Bank), which was occasionally called *Üzemi Híradó* (Operational News), is one of the most authentic sources for obtaining profound insight into everyday life at the Bank in the past.

² The editor-in-chief, Dr János Zádori, said the following about the necessity of rebranding in the January 1986 edition of the *Forint* magazine (Page 11): “The *Bankszemle* had become obsolete, it was the only small-circulation periodical in Hungary. In such a journal with second-rate typography and bland cover, even top-quality works could not make an impact.”

³ In its heyday, around 1975, approximately 3,600 copies of the journal were published, which was later slowly and gradually reduced. As a result, the initially subscription-based system was replaced by a free-or-charge model at the turn of the 1990s, which was to stop the plummeting circulation of the journal.

⁴ The paper existed between 21 December 1949 and the end of July 1950 (*Botos – Botos, 2004*).

⁵ It was published by the Hungarian Communist Youth Association (KISZ), and its contents and quality were a far cry from those of the *Forint*.

⁶ This was the official publication of the Hungarian Banknote Printing Company and was first published in April 1958.

⁷ After June 1953, the social information provision function of the *Forint* was taken over by the *Híradó*, which was published in a stencilled form.

Figure 2
Transformation of the covers of the *Bankújság* between 1949 and 1950, and the *Forint* between 1950 and 1989
 (Part 1)

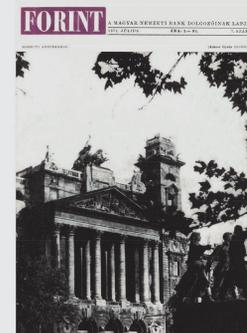
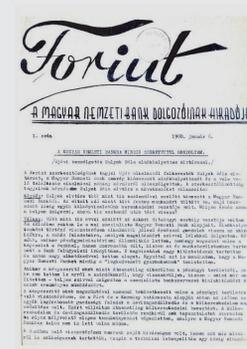


Figure 3
Transformation of the covers of the *Forint* between 1950 and 1989
(Part 2)



The journal was produced by an external printing house (Kossuth Nyomda) from 1969, which was replaced by printing within the Magyar Nemzeti Bank in 1979. This transition shortened printing times and brought about technological innovations in connection with production. New page-setting opportunities arose, which enabled the expansion and renewal of the contents of the journal on the one hand, and facilitated and enhanced the readability and structuring of the content on the other hand. These innovations had a positive impact, not only on the text but also on the visual content. One of the most striking examples for this was the change in the content and visual appearance of the journal's cover and header, which was fine-tuned over time, and became increasingly varied and exciting. Text-based and monochrome covers were replaced by ones with a colour background, and they were illustrated initially with drawings and then with black-and-white photographs, and after the 1990s with colour photographs. Thanks to page-setting

technology, a balance was struck between the text and the images in the journal, as the illustrations within the text gained increasing prominence, thereby supporting and complementing the message of the written content. Moreover, *caricatures* shaping readers' opinion also appeared as independent, unique visual means of communication.

These focused on economics and general topics related to economics and finance (e.g. the "Közgazdasági kislexikon" (Concise Encyclopaedia of Economics) series or the "Értelmező bankszótár" (Dictionary of Banking Terms) series) on the one hand, and were representations of the various problems and situations in the everyday life of the Bank (e.g. lack of space in the Bank, the crowded rooms, the elevator that was perpetually out of order) on the other hand. Completely independent from these, we can of course find drawings addressing general social issues (e.g. the "Fintor" (Grimace) series). One of the special features of the journal was the illustrations on half a page or an entire page that were published quite frequently and that may bring to mind the currently popular, storytelling images.

Of course, the *Forint* also contained political statements, notices and reports characteristic of the age, mainly in the first half of the paper. The journal's special columns stood out from this (e.g. "Bankokos Mancsi", "Fintor", "Fullánk", "Bankenciklopédia", "Értelmező bankszótár" (Know-All Mary in the Bank, Grimace, Sting, Banking Encyclopaedia, Dictionary of Banking Terms)) and humorous writings in verse (e.g. "Tűnődő Bankos Sebestyén" (George, the Pondering Banker)). The one-off or serial publication of these made the paper distinctive, unique and also more personal and casual for the readers. In addition to the traditional pieces, reports and portrayals depicting the everyday life and the various events in the Bank as well as the institution's sporting and cultural life, which were later enriched by photographs, articles with a more critical and scathing tone were also included (e.g. cutting red tape, comprehensible communication, bland and expensive food in the canteen, overcrowding, wasting working time).

The circulation and print run of the journal was strongly influenced by the active headcount in the Bank, which started slowly diminishing in 1972, and then from 1986 this decline continued at an increasing pace. Despite the continuous attempts to renew the journal, the downward trend could not be halted, which was a recurring problem until publication was terminated.

3.1. The secret graphic artist of the *Forint* magazine

The overwhelming majority of the illustrations and drawings between 1945 and 1978 were prepared by *Dr Herbert Enyvvári*, a lawyer at the Magyar Nemzeti Bank. However, his life was much more adventurous than just this. Born in 1916, drawing and graduating from the Academy of Fine Arts were his passions from his childhood. After his school leaving exam in 1934, he enrolled to the Legal and Political Science faculty of the Pázmány Péter Catholic University at his father's

insistence. He continued to draw all throughout his years at the university, getting orders from several publishers (e.g. Athenaeum, Dante, Cserépfalvi, Révai), and collaborating with Zsigmond Móricz, Frigyes Karinthy among others, and he also worked together with folk authors Géza Fejes and József Darvas. The books he illustrated (e.g. the tales of Gusztáv Bernáth or *Háry János legújabb kalandjai* (The Latest Adventures of János Háry) by István Tamás) made him popular not only in Hungary but also at the international level. He received his degree in 1939 and then enlisted in the army. He only returned home in 1943, due to a shrapnel wound sustained on the front. At his father's insistence, he started looking for a respectable job and began working at the Magyar Nemzeti Bank. That same year, he commenced work on the illustration to Cervantes' *Don Quijote* published in the translation of Miklós Radnóti. A deep and close friendship developed between the poet and the illustrator. In the summer of 1944, he asked for permission to return to the front, where he was wounded again and then captured during the Soviet advance. He only returned to the Bank in 1945 as an expert in labour law, where he was later also in charge of preparing the internal rules of procedure. That was when he started drawing fervently in the *Forint*. He married his wife, Éva Rákóczy,⁸ whom he met in the Bank, in 1948. Let us see the perhaps most telling recollection about the personality of Herbert Enyvvári: "During his life, he always resisted the oppressive regimes' apparatus with all available means, even in peacetime after the war. During communism, he stood with two bags of potatoes in his hands during the compulsory official assemblies in the bank (!) so that he did not have to clap" (Miklauzič, 2013b: 37). Herbert Enyvvári's caricatures and illustrations were the cornerstones of the *Forint's* character.⁹ Although many people¹⁰ followed him in this position, his subtle drawings, unique graphic style, humour and exceptional attitude would have been impossible to imitate.

⁸ Éva Rákóczy, a descendant of the aristocratic Rákóczy family, also joined the Bank in 1945. She was first employed in the Tax Division and then in the Legal Department. One of her ancestors was János Rákóczy, "the secretary and closest confidante of Lajos Kossuth, János Damjanich's brother in arms, a fervent supporter of Sándor Petőfi, the lifelong friend to Imre Madách and Mór Jókai, the intellectual partner of Hungarian patriots, a commanding figure of the army's leaders, supporting the 1848–1849 freedom struggle with an independent cavalry troop" (Miklauzič, 2013a: 5).

⁹ Digitized publications can be viewed in the Library of Magyar Nemzeti Bank.

¹⁰ The first to take over the illustration work was *András Káspári*, another staff member at the central bank (occasionally in 1978–1979, then mostly between 1984 and 1990), and there was a short interim period (1982–1983) when the illustrations in the journal were mainly prepared by *Dr Ferenc Kesztyűs*. From the 1990s, the *Forint* started publishing the works (or reproductions) of professional caricaturists, mostly *László Dluhopolszky (DLUHO)* and *Jenő Dallos*.

4. Other publications

The *Referáló Lapok* (Reports) series published between 1965 and 1980 (Figure 4) included summaries of articles that appeared in the general economics and financial papers of the socialist countries, covering different topics in each issue. The *Tanulmányok* (Studies) series had the same image in 1960 when it was launched, but it was “rebranded” around 1964 (Figure 5).

Figure 4
Covers of the *Referáló Lapok* series in 1977

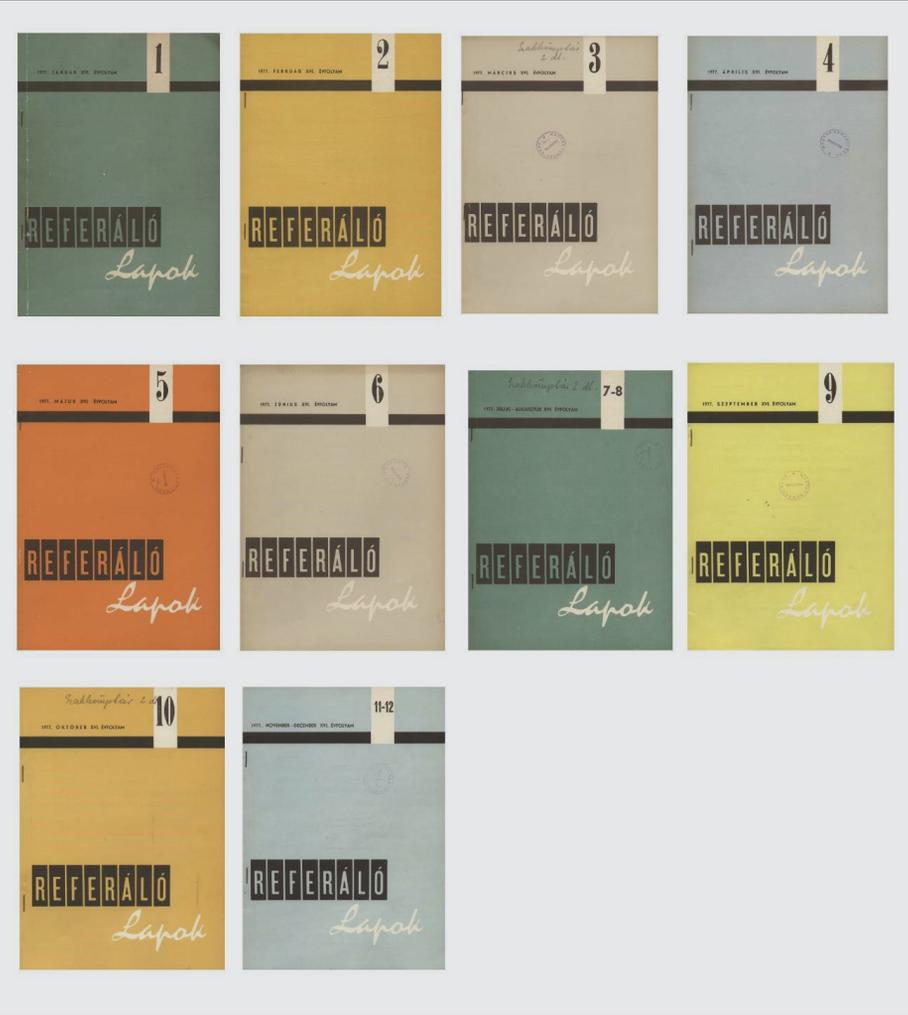


Figure 5
Final covers of the *Tanulmányok* series
(around 1963–1964)



The covers in both series use different colours, but are still characterised by a recurring look at certain intervals. On account of the variety of covers and the typographical solutions, to the eyes of the modern reader these were perhaps the two series of the Bank with the most interesting content and the most uniform visual image.¹¹

The *Közgazdasági Irodalmi Szemle* (Economic Literature Review) (Figure 6) was launched by the Economic Department. The series was published between September 1940 and June 1944 (Botos, 1999: 256–257), with the aim of supporting and facilitating the professional awareness of the Bank’s staff through reviews of

Figure 6
Rebranding of the covers of the *Könyvismertetések*, the *Folyóiratszemle* and the *Közgazdasági Irodalmi Szemle*

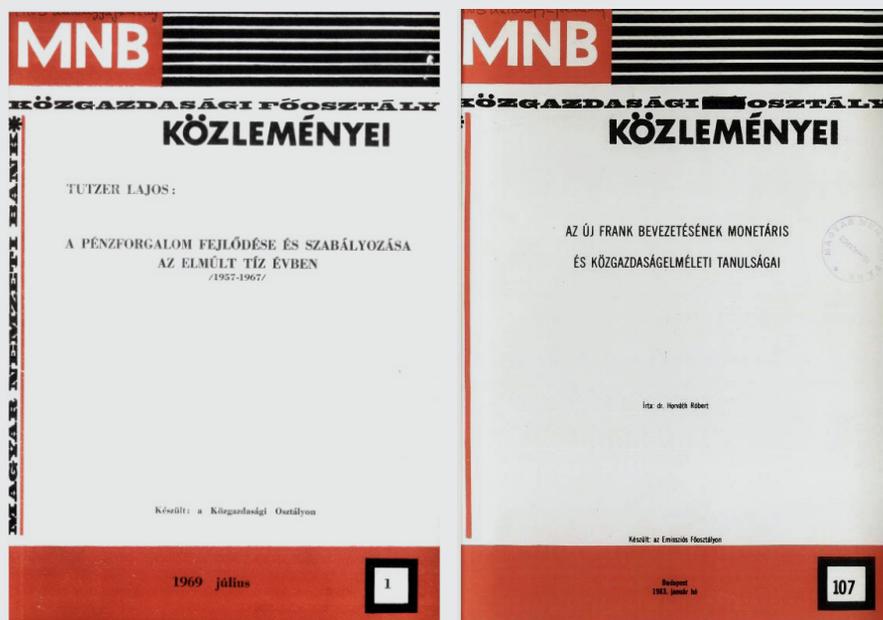


¹¹ There are remarkable similarities between the style of the covers and the character of the currently very popular handbook series of the *Harvard Business Review* (HBR) Press. Both use a sleek design and simple colours and typography for the covers of their publications. The “only” difference is that the HBR series also includes demonstrative illustrations.

foreign literature. After processing the articles published in Hungarian journals¹² with an explicit focus on banking, the reviews were divided into seven topics¹³ for the readers. By contrast, the *Folyóíratszemle* (Journal Review) (Figure 6) provided detailed information on the monetary and credit management methods of socialist countries and the theoretical articles published in official journals. As suggested by the title, the *Könyvismertetések* (Book Reviews) (Figure 6) sought to present books aimed at a professional audience.

The *Közgazdasági Főosztály Közleményei* (Publications of the Economics Department) series (Figure 7) was published between July 1969 and August 1989, with various longer professional and scientific studies, treatises and summaries by the Bank's staff. It was very close, or, we could say, a forerunner in style to the

Figure 7
Covers of the *Közgazdasági Főosztály Közleményei* in 1969 (Issue 1) and in 1983 (Issue 107)



Note: The reason why Issue 107 from 1983 is shown here is that from then on, later issues were available only in black and white in the digital archive of the Magyar Nemzeti Bank (Nainuwa). However, except for the colours, everything appeared as in the previous issues. Nevertheless, in order to illustrate the visual continuity, the last colour-cover issue available in the archive is included here.

¹² In the *Bankszemle* (Bank Review), the *Gazdaság* (Economy), the *Közgazdasági Szemle* (Economic Review), the *Külkereskedelem* (Foreign Trade), the *Pénzügyi Szemle* (Financial Review) and the *Társadalmi Szemle* (Social Review).

¹³ The issues of general finance, income regulation, budget, lending, payments, FX management, foreign trade and investment finance.

MNB Füzetek launched at the turn of 1995 and 1996,¹⁴ the idea of which originated from György Szapáry,¹⁵ and to the *MNB Occasional Papers* launched in the autumn of 2005.¹⁶ During the 30 years of this series, a total of 117 issues appeared, the special feature of which was not the cover of the series, which was unusually unique and incongruous with the world of banking, just like its typography, but the fact that it was able to appear with the same look and structure all throughout its existence.

We must also mention another publication, the series published by the Education Department,¹⁷ established in 1952 (*Figure 8*). The department was tasked with the uniform supervision of the internal and partly external professional trainings of the Bank at that time (*Botos – Botos, 2004: 138*). The publications for mostly internal circulation, produced in various print runs, providing comprehensive and concise professional advice and task descriptions to the staff in a topic or position, were probably published because of this from 1953. The publications were published by Pénzügyi Kiadó, and the covers were using soft colours (grey, white, dark brown).

5. Summary

Nowadays, the importance of the visual image and especially visual communication is undeniable. The method, channel, structure and “external look” of communication are critical, despite seeming insignificant and non-essential, and they should form an integral part of professional and scientific life. This is because by nature, visual images and visual communication raise awareness and guide us, which may facilitate the processing of the content geared towards professionals. In a world where the available information multiplies day by day, it is much harder, but not impossible, to stand out from the “crowd” of professional communication. Well-structured organisations, journals and series with a uniform visual image can communicate continuity, stability and reliability towards the outside world. Over the longer term, accessing, structuring and categorising the visual image of an organisation’s professional materials paves the way for using the visual foundations as a platform in order to make the organisation’s content geared towards professionals stand out even more from the crowd, thereby drawing attention to the organisation.

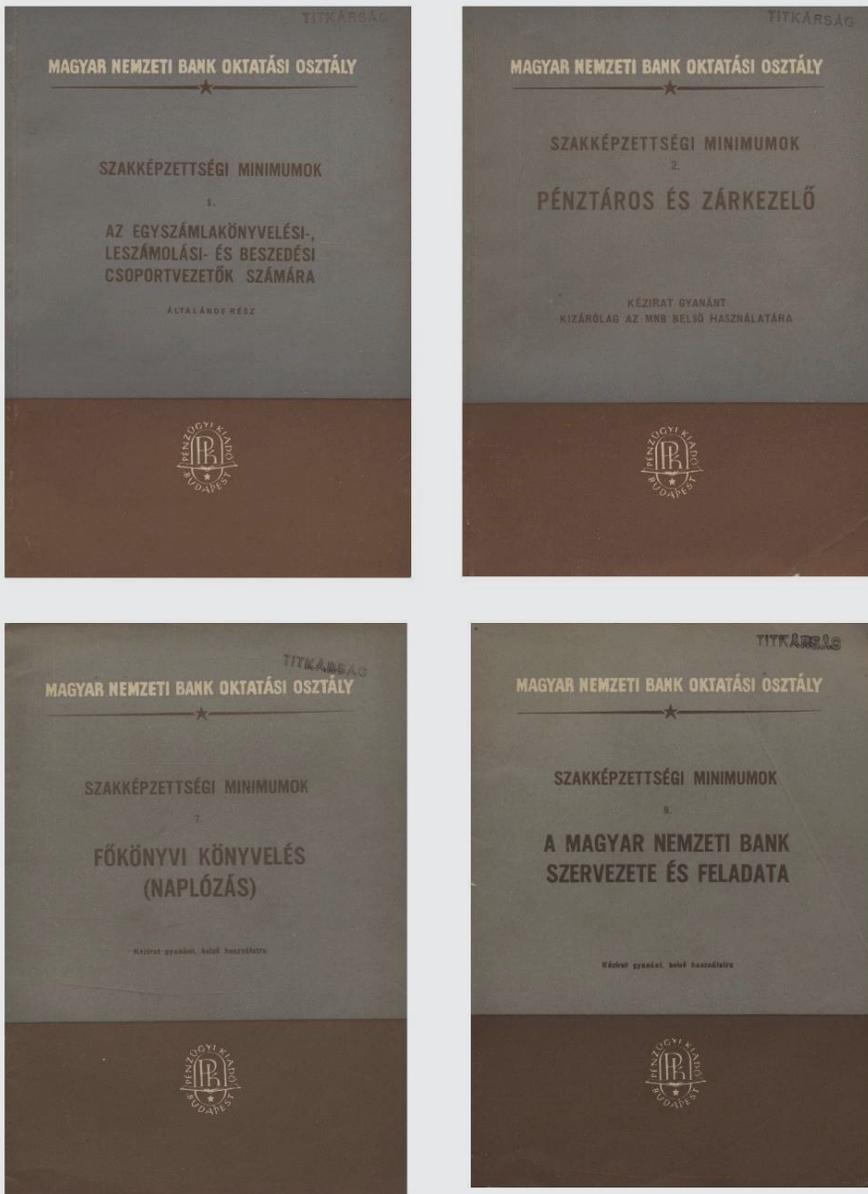
¹⁴ The *MNB Füzetek* was first published only in Hungarian, and then – in line with similar international publications – only in English, and, accordingly, it was renamed *MNB Working Papers* and is still published today.

¹⁵ Deputy Governor of the Magyar Nemzeti Bank between 1993 and 1999, and 2001 and 2007, and currently chief advisor to the Governor of the Magyar Nemzeti Bank.

¹⁶ The *MNB Occasional Papers* combined the *MNB Háttér tanulmányok* (MNB Background Studies) series and the *MNB Műhelytanulmányok* (MNB Workshop Studies) series published before the autumn of 2005 and is still published by the Magyar Nemzeti Bank.

¹⁷ The Education Department, which was in charge of the publication, was created in November 1952. It sought to provide assistance in the professional training of the colleagues in the national network of regional offices, county branches and branch offices established between 1950 and 1956. The department operated for three years and was then merged into the HR Department in 1955.

Figure 8
Covers of the publications by the Education Department of the Magyar Nemzeti Bank in 1953



I believe that the present description may be of interest not only because of its special nature, but also because it may point out aspects for the future that can help the further development of the Bank's and its publications' external and internal communication.

References

- Botos, J. (1999): *A Magyar Nemzeti Bank története II. kötet – Az önálló jegybank 1924–1948 (History of the Magyar Nemzeti Bank, Volume II – The Independent Central Bank 1924–1948)*. PressCon Kiadó Kft, p. 399.
- Botos, J. – Botos, K. (2004): *A Magyar Nemzeti Bank története III. kötet – A jegybank útja rendszerváltozástól rendszerváltozásig 1948–1989 (History of the Magyar Nemzeti Bank, Volume III – The Central Bank from One Political Transition Until the Next 1948–1989)*. Tarsoly Kiadó, Budapest, p. 753.
- Miklauzič, I. (2013a): *A nemesi Rákóczy-család. Történelmi életutak (The Aristocratic Rákóczy Family. The Lives of Historical Figures)* Mécses Sorozat, Székely Ház Közhasznú Alapítvány, p. 137.
- Miklauzič, I. (2013b): *A Mikó – Nyiredy – Arlow család. Prémiszövők, méneskariak és a rendíthetetlenek kopjafája (The Mikó-Nyiredi-Arlows. Fur Weavers, Cavalry Officers and the Wooden Memorial Columns of the Dauntless)*. Mécses Sorozat, Székely Ház Közhasznú Alapítvány, p. 196.

The Revolution of Money

Gábor Kürthy

Tamás Bánfi:

A pénz forradalma. A pénzteremtés elmélete és gyakorlata

Cenzus Bt., Budapest, 2016, 287 pages

ISBN: 978-963-12-7223-9

“Banks lent more of the deposits held with them to residential customers than usual. The increasing demand for consumption had some serious consequences: the price level increased and the level of savings dropped, which made funding investments impossible. Owing to the drop in investments, unemployment rose and tax revenues shrank, which increased the government deficit and placed government debt on a growth path. The funding of pensions was also jeopardised. The central bank, which is independent of the government, tried to control this harmful process by raising the interest rate.”

The above – imaginary – quotation could be included in the case studies of any mainstream textbook on macroeconomics, although it is full of entrenched concepts and relationships that we use and teach without any serious review. And we are wrong. The book by Tamás Bánfi entitled *A pénz forradalma – A pénzteremtés elmélete és gyakorlata* (The Revolution of Money – The Theory and Practice of Money Creation), explores, considers and reconsiders the relationships between money creation, measurement of value, government finances (deficit, debt, pension financing), and the relationships between investments and savings, unemployment, fiscal and monetary policy. Each chapter can be considered a study in its own right, yet each is related to money and the creation of money.

Banks create money by accounting transactions, lending and crediting the amount of money on the account of the customer. To understand and bring others to understand this statement is not an easy task, even in academic and professional financial circles. People who are constantly faced with arguments denying or countering money creation or neglecting its importance should not be surprised that a scientific collection on money begins with a chapter introducing the creation

Kürthy Gábor is the Head of the Department of Finance at the Corvinus University of Budapest, associate professor.

E-mail: gabor.kurthy@uni-corvinus.hu

of money (*The theory and practice of money creation*). One should begin reading the book with this, and we should be impressed at how simple this operation is.

Chapter two (*The measurement unit of the numéraire is the currency standard*) introduces us to the theoretical and sometimes practical (e.g. 27 July 1946 in Hungary) zero point of money creation, the time of the introduction of a new currency. In hyperinflation, money loses its functions one by one: it is no longer a store of value, a medium of exchange, and it can also no longer be used to measure prices. Therefore, upon introduction of the new currency, the fiat money, the monetary standard should be determined. The monetary standard can be determined fully arbitrarily, but it is practical to choose a standard that minimises the money circulation, information and other transaction costs.

Chapter three (*A theoretical history debate of money creation*) explores money creation from a historical aspect. The long period when monetary functions were fulfilled by substitutes for money besides and instead of commodity money connects the eras of commodity money and fiat money. Classical paper money and classical banknotes basically substituted precious metals in their functions of means of exchange and means of payment, but after some time their increasing amount started to determine prices as well, while the price-determining role of gold became questionable. And if this is so, then – having fulfilled all the functions of money – they are not merely substitutes for money, but rather are the monies themselves.

Chapter four (*Savings and investment equality does not prevail according to Keynesian definitions*) is a debate partly on definitions and partly on the history of theory, while at the same time it also has important practical, economic policy messages. The mistaken belief found in professional circles that investments require the preliminary (or at best, simultaneous) accumulation of savings, can only be verified if money creation is denied. It is possible to finance deficits in government finances or business current deficits by new, fiat money without these funds having been collected by someone previously, and without having price increases as a result of such financing. The investment-savings equation may still apply: fiat money finances investments, and thus it will still be received by someone as income. Unspent amounts will remain savings, by definition, until the income has been spent on consumption. But even if it does get spent, it will become (as yet unspent) income – and simultaneously – savings somewhere else.

Chapter five (*The blind spot (or badge of shame?) of economics: the uninterpretable and unmeasurable price level*) should be started (almost) from the end, so we can see how many authors use the concept of “price level” in so many places, without going deeper into its exact interpretation. This interpretation is not only a theoretical challenge. The measurement, forecasting, reduction and/or maintenance of inflation is a practical, defining element in monetary policy, and these activities

cannot be adapted to a falsely created, inconsistent benchmark. Several alternative options are possible: on the one hand, we could select a good, whose price changes allow us to follow nominal changes of price values; on the other hand, we could use wage level instead of price level, since it measures a homogeneous type of good (or at least much more homogeneous than the world of goods), and its changes are closely related to changes of the rest of the prices.

Human communities have had an increasingly hard time controlling those harmful dynamics – increasing inequalities in income and wealth, destruction of nature, economic, social and political crises – that are generated by selfishness and the pursuit of self-interests channelled into market demand and supply. It makes the struggle for survival as a community even more difficult when we use inaccurate concepts and unclear relations to describe and model the economy, and importantly, because incorrect conclusions are drawn from the models, there is no harmony between fiscal and monetary policies (Chapter six: *Relationship between fiscal and monetary policy in the 20th century and at the beginning of the 21st century*). It is worth following the historical road that begins with the bank that finances the state and ends with a central bank independent of government finances, with the function of targeting inflation, dragging and pulling interest rates – and thereby international capital movements. And we should also consider – in light of the harmful processes mentioned above – what could function as the new basis for this relationship.

In order to reach a fiscal and monetary relationship that is more workable than at present, it is necessary to understand government deficit and government debt from a theoretical and historical (!) aspect (Chapter 7: *Government deficit–government debt*). It is widely believed that controlling government deficit and government debt is one of the important – if not the most important – keys to economic and financial stability. The endeavour to reduce the deficit and decrease the debt follows from this concept. Yet, what we can see from economic and social developments tends to imply that there is a growing need for the presence of the state, while its funding is jeopardised from several sides (demographics, migrations). In other words, the budget deficit and the resulting indebtedness are exactly the conditions for social stability and sustainability. On the other hand, the government deficit played an important role at the time of the appearance of fiat monies, and the market of government debt is still the most important benchmark of money markets.

In the 20th century, not only was the relationship between the state, and between the central bank and the economy transformed, the structure of society and the family also underwent a transformation which was not independent of the establishment of the system of state institutions (e.g. the pension system). Families became smaller, and women appeared in the labour market. One thing that has not changed, however, is the interpretation of “unemployment” in theoretical

macroeconomics. Chapter eight of the book (*The necessity of revise of employment and unemployment in macroeconomics*) argues that – with the spread of the dual-income family model and single-adult households – it is no longer possible to classify persons of working age as active and inactive, as was possible at the time of the single-income family model. From that it follows that unemployment and inactivity cannot be interpreted and measured separately, instead of these, employment should be placed in the focus of macroeconomic thinking and aims of economic policy.

Pensions are the “invention” of the last two centuries, and presumably this invention has significantly contributed to the transformation of Western society; the established social and family structure makes it necessary to retain the pension systems, since neither self-care nor family care can be ensured. However, the current pension system cannot be sustained and requires major transformations. The acceptance of the proposals for change includes reconsidering the relationship between the central bank and the state, because according to the contents of Chapter nine (*A paradigm shift in the pension system*), as a temporary arrangement in a new, state-funded old age pension system the basis for later entitlements of the beneficiaries would be financed by money created by the central bank. Of all the proposals perhaps, this one is the strongest, and it would be instantly denied by economists trained to believe in the indisputable independence of the central bank. By comparison, the proposal on the termination of state support for voluntary pension funds or the statement that the old age pension system should be independent of employment are insignificant.

Each chapter of the book *The Revolution of Money* has some exciting contents: for economists, financial professionals, bankers, practitioners or theoretical connoisseurs of monetary and fiscal policy. I believe some readers will be not only surprised and even rather astonished by the statements and proposals set forth in this book. This is because they have learned different things using other methods: they use and read the concept of “price level” every day and they are convinced that the bank lends the deposits it holds. In fact, it is probably not easy to review these entrenched relationships and abandon them. But it is certainly worth a try: reading the book of Tamás Bánfi offers a good starting point for that.

Has the US Economy Lost Its Long-term Growth Potential?

Péter Tóth

Robert J. Gordon:

Rise and Fall of the American Growth

Princeton University Press, 2016, 784 pp.

ISBN: 9780691147727

Economic productivity and 20th century economic history can be viewed from a techno-pessimist and a techno-optimist perspective. According to techno-optimists, such as Klaus Schwab, founder of the World Economic Forum, the “Fourth Industrial Revolution” based on digitalisation will have unprecedented economic consequences. Robert J. Gordon is a pessimistic economist, who believes that techno-optimists exaggerate the impact of digitalisation. He maintains that this is confirmed by the decades after the American Civil War, during which the US was completely transformed due to the discovery of electricity, the invention of the internal combustion engine and the spread of motor vehicles. Gordon argues that the Golden Age of American economic growth is over.

The 1870s were the dawn of the United States. Over the next six decades, innovations and discoveries transformed all walks of everyday life. By 1929, all urban homes had electricity, gas pipes, telephone, sewer systems and clean running water. By 1929, horses had disappeared almost completely from the streets, replaced by motorised vehicles. In 1929, the opportunities for the public’s entertainment included the radio, listening to music on a phonograph and motion pictures. By 1929, maternal mortality had dropped to almost zero, and medicine had almost reached its current standard. The number of hours spent working decreased, and electric appliances made household chores considerably easier. This transformation progressed slowly, and then after 1900 it gained traction due to electrification and the spread of motorised vehicles. The 1870s laid the foundations of our modern age. Electric lighting, the first reliable internal combustion engine and wireless telecommunication were all invented in 1879, within a three-month period, and they were followed by the appearance of the telephone and the phonograph within a decade. The quality of food consumption, i.e. varied diets, developed at a slow pace, but the quantitative change was significant. Average calories consumed in 1870

Péter Tóth is the head of the New York office of the Magyar Nemzeti Bank. E-mail: tothp@mnb.org.hu

exceeded those in the 1920s. Clothing did not change much from a qualitative and quantitative perspective between 1870 and 1930, but the situation of housewives improved considerably. Thanks to the increasing efficiency of textile manufacturing and the spread of mail-order catalogues, the price of clothes plunged, which, coupled with the increasing income of households, enabled women to buy clothes, thereby saving them much work. The slow growth in food and clothing consumption is not a paradox because households spent the money they saved on those items on purchasing the latest discoveries. Nonetheless, it is highly likely that the growth in food and clothing consumption was underestimated due to the distortion of the price index. The price index does not take into account that as chain stores appeared, consumers were able to buy food significantly cheaper than earlier. According to the author's calculations, in 1911 chain stores were cheaper than traditional stores by 22 per cent on average. We can see a similar distorting effect in clothing prices, and in this case mail-order, catalogue-based merchants were the cheaper alternative to tailors. In addition to the distorting effect of the price index affecting food and clothing, the considerable improvement in living conditions shows that a change in real consumption does not necessarily reflect the public's living standards. Can we assign economic value to the transition to fully equipped homes? The answer lies in the connection between prices and rents of homes and the various levels of amenities. Homes equipped with a bathroom were 82 per cent more expensive in the period under review than those without one, and homes with central heating cost 28 per cent more on average than unheated flats. No studies have been prepared concerning the price of homes furnished with electricity, but Gordon believes that they must have had a premium of at least 28 per cent on the market. *The revolutionary change in the living standards provided by US residential buildings is one of the main topics of this book. These were innovations that can only arise once in humanity's history.* Although it took over 50 years for the above-mentioned solutions providing all the necessary amenities to spread throughout the entire population, the process was completed by the mid-20th century. Sustained economic growth required a steady flow of innovation; most electric household appliances targeting consumers were invented before 1940, and it was only a matter of time for them to reach all households. The author argues that with the exception of air conditioning, no invention after 1940 was able to radically transform people's daily lives. *Homes with full amenities and vehicles using internal combustion engines boosted the living standards of the American population so tremendously that the author believes that this cannot be repeated.* The dramatic change in transportation is portrayed by Gordon through three major channels. The first channel is the development of the uncoordinated, low-quality railway system before 1870. By 1940, train speeds had increased from 20–25 miles per hour to 70–75 miles per hour, and due to infrastructure developments, trains were now able to cover large distances. The introduction of air-conditioned trains in 1940 contributed to the improvement of travel comfort. The second channel is linked to the development

of transportation within city limits. Until 1870, urban transportation was provided by horse-drawn carriages, at a speed of 3–6 miles per hour. By 1902, horse-drawn vehicles of public transportation had been replaced almost completely by trams and trolleys. On account of the frequent traffic jams in large cities, underground and elevated rapid transit railways were built shortly after the introduction of trams, and they were first used within the city limits, then they became suitable for travelling between cities. The third and most important channel was the appearance of cars, which transformed both urban and rural life between 1910 and 1930. Henry Ford's Model T could be purchased for a quarter or half of the average annual income in the 1920s, as a result of which this car was available to the entire public. By 1926, 93 per cent of farmers from Iowa and other northern states had a car. *Between 1870 and 1940, the US population rose from the period of isolation thanks to innovations in communication and the entertainment industry.* In 1938, one-third of the total US population, i.e. over 40 million people, listened to the broadcast about the competition between Sea Biscuit and War Admiral, at one of the most famous horse races of all time. Between 1870 and 1940, a drop in infant mortality and contagious diseases can be observed. *The increase in life expectancy was twice as high in the first half of the 20th century as in the second.* The favourable health developments can be mainly attributed to the establishment of urban sewerage systems and clean running water. According to estimates, 75 per cent of the decrease in infant mortality was due to the provision of clean running water. Furthermore, growing hygiene awareness also has to be mentioned, just like the preservation of food and the introduction of regulations pertaining to the contamination of food, as these all contributed to the reduction of contagious diseases. Between 1890 and 1930, the environment and conditions of work experienced an unprecedented change. The number of hours worked per week and occupational accidents dropped, a much larger portion of the young generation could take part in education as child labour was abolished, and the innovations used in households made the life of housewives easier. The Great Depression and the Second World War interrupted these developments, but 25 years after the war, the impact of improving working conditions could be felt throughout the population. The reduction in the marginal disutility of work occurred in three dimensions, one of which was the decline in working hours per week from 60 to 40.¹ The second dimension was a shift in the distribution of labour from demanding and dangerous blue-collar work towards white-collar work. The third and most important dimension was the invention of household appliances that eased the burden on housewives.

Economic growth started decelerating in 1940 and especially in 1970. This claim can be considered evident if we examine the changes in the quality of consumer goods and services providing basic necessities, i.e. food, clothing and shelter.

¹ In the steel industry, the average number of hours worked per week fell from 72 to 40.

The better part of the development in food and clothing occurred between 1870 and 1940. The spread of frozen food after 1940 merely provided another alternative to canned food that had been eaten for decades. As households' income increased, the share of spending on food declined from 45 per cent to 13 per cent between 1870 and 2012. The number of cars increased rapidly until 1970, and then this rise decelerated between 1970 and 1990 and halted completely after 1990. Miles travelled by car increased until 2000 at a diminishing pace, and then shrank continuously between 2000 and 2014, which, according to Gordon, can be probably attributed to the drop in the share of the population holding a driving licence. In contrast to the areas examined earlier, no slowdown could be observed in entertainment after 1970. Among the components of the US living standards, entertainment opportunities increased the most between 1940 and 2014. Today, households can access services linked to entertainment that did not even exist in 1940. As a result of the development in information technology, ideas were born that fundamentally reformed human communication and data storage habits. Gordon Moore, the co-founder of Intel, gave a very accurate estimate in 1965. He believed that the memory of computer chips would double every two years. Between 1990 and 2006, computing capacities doubled every 16 months, however, the pace of development declined after 2006, and currently the memory of IT devices doubles every 4–6 years. According to the author, this is because there is no demand for the engineering expense to reduce the size of chips, since their current size and capacity are perfectly suitable for completing the tasks on desktop and laptop computers. The emergence of the Internet has transformed users' communication and shopping habits. Information has become free, and new communication channels, such as Facebook, Twitter and Skype, have appeared. Despite the rising costs of healthcare, the positive impact of the development of the technologies used in medicine falls short of the effects seen between 1940 and 1970. Large-scale healthcare spending on medical research and treatment will undoubtedly continue to raise the life expectancy of the US population at a slow pace, but it is unlikely to cause radical changes in the living standards of Americans. Since 1940, working and living conditions have improved considerably, but at a significantly slower pace than before 1940. Between 1940 and 1970, the trend that started in the early 20th century continued: tough physical labour was replaced by safer and less demanding activities. By 1970, household appliances had become almost universal among the population, which lifted the burden of household chores from women. Although between 1940 and 1970 modern working conditions became established, the transformation of women's role in society and the labour market started only in the mid-1960s. Rising educational attainment was the main reason behind the spread of white-collar work. Child labour was abolished in the US before the Second World War, and by 1970 three-quarters of the total population had a secondary school diploma. As a result of the passing of the social security legislation in the 1930s, most workers retired at 62, which, coupled with

the rising life expectancy, placed extreme pressure on the US pension system. In order to maintain the financial sustainability of the social security system, most employers started using defined-contribution 401(k) plans instead of the annuity-based plans in the 1980s. The author identifies two factors that have contributed in an unparalleled manner to US economic growth since the 1940s. *The emergence of women on the labour market and decreasing discrimination against the African American population paved the way for the efficient allocation of investments in human capital.* One study found that this labour market transformation contributed 15–20 per cent to the economic growth between 1960 and 1990. The unique nature of the transition can be seen in the fact that the female labour market participation rate has been on the decline since the 2000s, and the difference between the average wages of whites and African Americans has been stagnating since 1990. The other reason behind the decelerating economic expansion is the slower growth of the population's educational attainment. The share of those with a secondary school diploma has been stagnant since the 1970s, and even though the number of graduates increases steadily, many young graduates are unable to find work on the labour market.

Gordon argues that the Great Depression of 1929 and the Second World War contributed directly to the largest economic growth in US history. Without the Great Depression, the New Deal would not have been born, just like the NIRA or Wagner Act, which facilitated the establishment of trade unions. Partly due to unionisation, real wages started rising, and the average hours worked per week decreased from 60 to 40. As a result of higher wages and shorter working hours, economic productivity soared in the early 1930s, before the US entered the Second World War. Owing to the surge in real wages, private sector investments increased, which launched the process of substituting labour with capital. *The other positive effect of the Great Depression was the reform of business models. The cost-cutting measures entailing mass layoffs were inevitable due to diminishing production and profitability. The new and more efficient business models contributed to the productivity growth observed after the Great Depression.* The pressure exerted by the Second World War on the economy had a less speculative effect on productivity. Several economists analysed at length the speed and efficiency of building the Liberty freighter ships. Much to the surprise of the economists, the remarkable labour productivity growth persisted after the war. The drop in military expenditure did not reduce labour productivity in the military industry for a few years after the end of the war. War fostered innovation, and the productivity-boosting effect of new technologies took hold even in peacetime.

According to experts forecasting the impact of innovations, economic output will be increasingly produced by robots. Techno-optimists believe that as a result of future innovations, labour productivity will substantially increase, which will offset the

persistent mass unemployment caused by the machines. *Techno-pessimists forecast that the emergence of technologies will not bring about radical productivity growth, and – similar to the past decades – it will only entail slow economic growth and a gradual decrease in unemployment. Statistics confirm techno-pessimists' view.* Unemployment in the US has dropped from 10 per cent to below 5 per cent since October 2009, and annual labour productivity growth is 0.5 per cent, which is well short of the 2.3 per cent productivity growth seen in the dot-com era between 1994 and 2004. Gordon predicts that labour productivity will increase by 1.2 per cent on average annually between 2015 and 2040, which is far below the standards of 1920–1970 and 1994–2004, but exceeds the average annual percentage growth in the past 11 years, and the 0.5 per cent increase observed in the last five years. The author estimates that in the coming 25 years, the median disposable income of the US population will grow by 0.3 per cent on average each year, which is significantly lower than the 2.25 per cent average between 1920 and 1970, and the 1.46 per cent between 1970 and 2014.

Impact of the Financial Crisis on the European Constitutional System

Dániel Papp

*Kaarlo Tuori – Klaus Tuori:
The Eurozone Crisis: A Constitutional Analysis
Cambridge University Press, 2014
ISBN: 978-1107056558*

The authors would like to make the reader see the crisis as more than just a financial calamity affecting government debt and the banking sector. They place the issue in a broader context and urge us to view the crisis from a broader public order and constitutional perspective, encompassing European integration as a whole. This is because the euro-area crisis had a massive effect on the entire European constitutional system. The authors thus argue that the problem can only be analysed taking into account the common European values guaranteed by the Treaties. In line with this, the authors undertake not only to paint a comprehensive picture of the economic reasons behind the crisis and the crisis-management solutions employed, but also to present a historical and conceptual analysis of the EU Treaties in the light of the crisis.

The book lists the economic causes of the crisis, such as the collapse of the US mortgage market, which suddenly turned the highly interconnected nature of the financial market, which had been regarded as beneficial in the EU, into a predicament, creating a multitude of contagion channels that enabled the rapid spread of the crisis. In addition to these, the book also pays special attention to abstract causes such as the fact that financial risk-taking and responsibility for those risks have not always been in direct proportion. The authors believe that in many cases, this could occur because governments intervened in the system by providing state guarantees, which is in large part attributable to the fact that household debt management remained within the purview of the Member States. The other shortcoming of the EU presented in the book is that the Member States disregarded the changing conditions on the global market when drafting the current Treaty, especially the rapidly growing clout of Asian countries.

The most striking example of the lack of consistency between the Treaties and the rescue packages created in response to the crisis can be seen in the “*no-bailout*” clause of the EU Treaties. According to the clause, the EU does not bear the responsibility for the obligations of the Member States in any form. However, this is in stark contrast to the EU’s behaviour during the crisis. The authors believe that this was possible only because the rescue packages were obtained by the recipients under the auspices of the European Stability Mechanism, which is a legal person in international law, separate from the Member States, therefore from a formal law perspective it is uncertain whether it is covered by the “*no-bailout*” clause enshrined in the Treaty. This cunning legal argument was implicitly accepted by the subjects of EU law, since in a teleological approach, the aim of the “*no-bailout*” clause and the rescue packages was the same: maintaining the stability of the euro area. Based on this, we cannot say that in a teleological sense the rescue packages would be in conflict with Article 125(1) of the EU Treaty. The authors argue that the fact that this is not admitted explicitly by major EU institutions is because it would entail serious moral hazard, as in that case Member States could spend without limitation, secure in the knowledge that the EU would come to their rescue if needs be. The other important consequence of the crisis with respect to the EU’s constitutional order is that the issue of Member State sovereignty has become relativised and conditional on the financial assistance provided by the EU or the EU Member States. This relativisation could be best observed in connection with the Greek rescue package where the Member States providing the assistance set out strict conditions for the disbursement of the funds, and these conditions radically affected Greek society.

The authors believe that a straightforward solution to the above-mentioned problems would be to considerably strengthen the direct democratic legitimacy of the EU institutions, which is also part of the recent reform plans of French President Emmanuel Macron. Traditionally, the main institution in this thought experiment is the European Parliament. In order to provide a genuine solution and avoid a situation where Member States would feel that the EU’s economic decisions are imposed upon them, the decision-making mechanism should be reformed as well, since currently the European Parliament plays no role in most major economic decisions and rescue packages. The fact that the European Parliament is so weak is a deeply rooted problem. It is perhaps also attributable to the fact that the development of the European demos, just like the European identity, is in an embryonic stage.

The main crisis management measures were implemented without the substantial engagement of the European Parliament and the EU institutional system subject to the Treaties, within a new framework based on individual intergovernmental agreements. The crisis generated a multitude of official, semi-official and unofficial bodies which do not have formal powers, yet exert significant influence. All of

the bailout packages were developed by organisations such as the Eurogroup, the dedicated working group alongside the president of the European Council, or the Euro Summit. However, in some sense these institutions are not subject to the system of checks and balances in the EU. One good example for this is the transparency guaranteed in the Treaties and the right to access documents, which are generally implemented in the EU institutional system, however, they take hold much less in the case of the mechanisms that are outside of this system, for example in the European Stability Mechanism. One serious problem arising from this is that – due to the fact that the responses to the crisis had no solid legal foundations – the cornerstones of the EU are called into question. One such issue is the question of the rule of law. This is because the rule of law requires that all the elements it comprises be implemented. If an important element such as the right to legal remedy is violated, there is no real rule of law. In the case of the financial rescue packages, judicial control is very much illusory and could not be enforced at all due to the pressure on the Member States.

According to the authors, the deepening of the crisis was halted in 2012, and by that time there was light at the end of the tunnel. The European institutions used the relative tranquillity to analyse whether they were able to provide appropriate responses in the most pressing times and to examine the lessons learnt from the measures taken, and they tried to ascertain how to avoid another crisis in the future. In light of this, a committee was established, headed by Herman Van Rompuy, with the participation of the European Council, the Commission, the Eurogroup and the ECB, to prepare a comprehensive reform plan for the EU's future.

The committee basically had three goals: to deepen economic cooperation, to remedy the democratic deficit caused by the crisis, and to coordinate the two-speed EU. Deepening economic cooperation was mainly to be achieved through harmonisation of the financial intermediary system, and thus they devised the basic idea behind the banking union that has since become clearer and has provided some practical experiences. Two of the originally planned four phases of the banking union plan have already been realised. In addition to the Capital Requirements Regulation (CRR) and the Capital Requirements Directive (CRD IV) pertaining to the operation of the banking system and also containing substantive law norms, the Single Supervisory Mechanism (SSM) overseeing the finances of euro-area countries was also launched in 2014. With regard to the future, the authors believe that everything points towards the other two pillars being fully realised, thereby creating a full-blown common European resolution framework and deposit insurance system.

The special relevance of the work is the authors' conclusion concerning the two-speed integration of the EU, where they assert that this phenomenon has been more pronounced than ever since the crisis. Nevertheless, this is not regarded to

be clearly negative but rather an opportunity in some sense, because they believe that if the goal is to deepen integration, this can be achieved much faster and more efficiently by agreements between the Member States than by amending the EU Treaties covering all countries, since the latter are used, due to the unanimity requirement, by certain Member States for promoting their own interests, thereby obstructing the process. All in all, the authors are pessimistic about the deepening of the EU's integration.

Report on the Competitiveness and Responsible Corporate Governance Section Workshops of the Hungarian Economic Association’s 2017 Itinerant Conference

Péter Domokos – Róbert Hausmann – Gábor Szarka

The Hungarian Economic Association (MKT) hosted its 55th Itinerant Conference, the largest annual conference of Hungarian economists with deep-rooted traditions, between 7 and 9 September 2017 in Eger. The Conference included 17 workshops in addition to the opening and plenary meetings. The conference’s keynote speaker was New York University professor *Thomas J. Sargent*, winner of the Nobel Prize for Economics. In this paper, we summarise the workshops of the *Competitiveness* section and *Responsible corporate governance* section.

The MKT’s competitiveness section held its workshop at the Itinerant Conference for the second time this year, with the priority objective of drawing attention to structural policy. The competitiveness section examined the topic from both theoretical and practical perspectives, providing a meaningful contribution to the development of economic thinking on sustained convergence in Hungary.

At this year’s workshop, the competitiveness section hosted three foreign and three Hungarian speakers who spoke about the definition of competitiveness and the associated methodological questions, and also investigated the competitiveness of the CEE region and Asia from various aspects. The presentations on the different dimensions and the diverse findings confirm that competitiveness cannot be given a single definition and that there is no single recipe for successful economic and social convergence, even if certain international experiences can often be adapted. However, there was consensus among the speakers on the fact that knowledge and innovation-based development is the best way to increase competitiveness in a sustained manner. Hungary’s macroeconomic figures provide an adequate basis for achieving a competitiveness turnaround and thus achieving sustained convergence.

Tan Khee Giap, Co-Director of the Asia Competitiveness Institute and professor based in Singapore, spoke about the central role of the state in avoiding the middle income trap, primarily through its economic stimulating, social balancing and

Péter Domokos is a junior analyst at the Magyar Nemzeti Bank. E-mail: domokosp@mnbb.hu
Róbert Hausmann is a junior analyst at the Magyar Nemzeti Bank. E-mail: hausmannr@mnbb.hu
Gábor Szarka is Professional Secretary to the CEO of the Budapest Stock Exchange (Budapesti Értéktőzsdé Zrt.)
E-mail: SzarkaG@BSE.HU

political stabilising functions. An efficient state is also essential according to the IMD's chief economist, *Christos Cabolis*. He also drew attention to factors such as infrastructure, economic openness, the availability of skilled labour and an ideal labour market environment. Polish professor *Arkadiusz M. Kowalski* identified the development of independent technologies, in addition to stable economic fundamentals, as the key to maintaining and improving competitiveness. This, however, calls for a technological turning-point in Hungary according to professor *László Monostori*. The National Competitiveness Council established this year may contribute greatly to a competitiveness turnaround in Hungary according to Deputy State Secretary of the Ministry for National Economy, *László Turóczy*. MNB experts *Gergely Baksay* and *Gábor Horváth* gave a presentation on the relationship between the banking system and competitiveness, presenting the central bank's newly developed international banking system competitiveness index. University professor *Zoltán Sipos* identified the quality and measurement of education and close cooperation between universities with a new generation of students and the corporate sector as the cornerstone for improving Hungary's competitiveness.

Tan Khee Giap, Co-Director of the Asia Competitiveness Institute spoke about the successful economic model of Asia and Singapore, the competitiveness of far Eastern countries and the research organisation's methodology for measuring competitiveness. Giap, who holds a PhD from the United Kingdom, is a professor at the National University of Singapore and also Chairman of the Singapore National Committee for Pacific Economic Cooperation. He has also served as a consultant to international agencies such as the ASEAN and the United Nations, is the lead author of 20 books and is a member of the editorial advisory committee of the British Competitiveness Review. His research interests include the Cost of Living Index and the Global Liveable Cities Index and competitiveness analysis on China, India and ASEAN (Association of Southeast Asian Nations) economies.

According to Giap, East Asia's economic development model is based on overcoming (1) production, (2) infrastructure and (3) financial challenges which helps overcome the middle income trap and may be the right path for Eastern European countries. He deems that successfully achieving these three criteria and thereby achieving economic convergence calls for a strong and responsible role for the state. According to his presentation, the role of the state must focus primarily on its economic stimulating, social balancing and political stabilising functions. The professor emphasised the importance of SMEs in improving competitiveness, mainly through their capacity to create jobs, flexibility, cost-effectiveness and innovative capacity. Tan Khee Giap presented the Asia Competitiveness Institute's current efforts to develop a productivity and efficiency assessment index for SMEs that will be capable of measuring their competitiveness. He specifically emphasised that SMEs must be given state support in making digital progress, as failure to do so

will give an advantage to multinational corporations. On this note, he mentioned Hungary's Funding for Growth Scheme as an exemplary initiative, emphasising the significance of the MNB's measures to support SMEs. In the context of an economic history overview, he noted that Singapore's economic structural transformation was mainly built on its people, progressing upwards within the production chain and the reinforcement of high-tech industries. Education and innovation have become the foundations of competitiveness.

The main advantage of the competitiveness index created by the Asia Competitiveness Institute is that it uses a regionally interpretable objective weighting method that enables the comparison of countries in space and time. The ranking, which includes 100 basic indicators, classifies the dimensions of competitiveness into four categories and 12 pillars in which the four areas have identical weights.

Christos Cabolis, chief economist at the IMD World Competitiveness Center, gave a presentation on the methodology for measuring competitiveness and Hungary's competitiveness. Cabolis has held numerous academic positions at various universities, including Yale, and has published on the topic of competitiveness in numerous publications. His current research areas include the evaluation of credit rating agencies, the interaction between competition and specialisation of venture capital companies, and optimal bank size.

Cabolis stressed that the state plays an important role in creating an environment in which companies are able to produce high value added products. He added that participating in the high value added phases of the value chain and efficiency are one of the cornerstones of corporate profitability that contribute to both job creation and sustained convergence. He gave a particularly positive assessment of Hungary's experiences in terms of job creation. Successful corporate governance also plays a pivotal role in company growth, but regular renewal and innovation along with identifying talent are also important factors.

According to Cabolis, economic openness plays a major role in national competitiveness and correlates with competitiveness. The more open an economy, the better it places in competitiveness rankings. A 2017 analysis by the IMD World Competitiveness Center identified Hong Kong, Switzerland and Singapore as the three most competitive states in the world. According to the IMD survey, the factors that contribute the most to Hungary's competitiveness are skilled labour, the labour environment and commercial relations. On the other hand, opportunities for development were identified in areas such as governmental and business efficiency. The IMD digital competitiveness ranking was also showcased at the end of the presentation, with its main pillars consisting of knowledge, technology and future

readiness. Improving flexibility and the capacity to adapt and further developing education may play a pivotal role in addressing digital challenges.

Arkadiusz M. Kowalski, professor of the World Economy Research Institute at the Warsaw School of Economics, investigated the competitiveness of the Visegrad Group using a multiple dimension econometric model. Kowalski has previously worked at Poland's Ministry of Economy in the domain of innovation policy and business clusters, has written numerous papers on competitiveness and international economics and also contributed to compiling the Polish competitiveness report. His presentation focused on the past economic performance of the Visegrad Group over the past nearly 30 years, identifying GDP per capita as the main indicator of this performance. In post-communist countries, Poland has achieved the greatest convergence since 1989, seeing its per capita GDP (based on PPP) go from 38% to 65% of the EU15 average. Kowalski identifies the components of economic convergence as (1) economic growth, (2) price stability, (3) full employment and (4) fiscal and (5) external balance, which he sums up as the "Pentagon of economic performance". The economist made a distinction between macro-level and micro-level competitiveness: the former is based on national attributes and is defined by economic policy, while the latter can be interpreted through the quality of the business environment and the competitiveness of firms.

According to the speaker, the influx of foreign direct investment contributed to the economic convergence of the V4 economies, the dynamic multiplier effects of which also had a positive impact on economic and social dimensions. Although FDI-based investments contributed to the region's technological convergence, the development of own technologies is necessary to maintain competitiveness and a knowledge and innovation-driven economic model needs to be created. Defining research and development directions (by identifying beneficiary sectors) is the state's central (top-down) function, but is also based on bottom-up initiatives. Regional clusters are the practical embodiment of innovative (R&D based) sectoral collaboration, and their spread will form the foundations of a competitive economy in the future. Kowalski argues that Hungary is in a particularly good position in terms of innovation within the Visegrad Group.

László Monostori, president of the Industry 4.0 National Technology Platform and Director of the Hungarian Academy of Sciences' Institute for Computer Science and Control (SZTAKI), held a presentation on the significant technological turning-point necessary to achieve a turnaround in competitiveness. Monostori is a full member of the Hungarian Academy of Sciences and university professor of the Faculty of Mechanical Engineering of the Budapest University of Technology and Economics. Winner of the Széchenyi Award and recipient of the Knight of Cross from the Order of Merit of the Hungarian Republic, László Monostori's research areas include intelligent production processes and systems, real time cooperative

corporations and change and disruption management within the manufacturing structure.

Industry 4.0, or the fourth industrial revolution includes the creation and spread of cyberphysical systems. As part of this process, the integration of reality and virtual reality, unattainable in the past, will be used to implement a new level of organisation and regulation of the entire value chain within product lifecycles. The cycle is based on the real-time availability of information. New business models emerge as a result, which strive to serve increasingly individualised customer needs. Monostori argued that Hungary has much to do in order to take advantage of the opportunities offered by Industry 4.0, mainly in the realm of education to develop digital competencies, renewing companies' equipment and providing direct and indirect support to SMEs.

To foster the national adaptation of the fourth industrial revolution, every industrially developed country is devising strategies. These include Germany's Industrie 4.0 strategy, China's Made in China 2025 programme and Japan's Industry 4.0, Society 5.0, as well as the Robot Revolution Initiatives. In Hungary, the Industry 4.0 National Technology Platform is aimed at fostering the spread of the fourth industrial revolution mainly by providing a framework for the 2016 Irinyi Plan and the objectives defined in the Industry 4.0 strategy. Monostori said that in the context of the discussed process, poles of excellence in production information technology and production management will be created with the task of accelerating the innovation process, creating digital industrial solutions, training a new generation of qualified experts and fostering the emergence of a sustainable and competitive production ecosystem.

Representing the Hungarian government, László Turóczy, Deputy State Secretary for economic planning and competitiveness of the Ministry for National Economy (NGM), gave a presentation on the Hungarian aspects of competitiveness and the activities of the National Competitiveness Council which was established this year. In his introduction, Deputy State Secretary Turóczy stressed that competitiveness can be identified over the long run based on the level of productivity, which defines the economy's growth rate. While the traditional economic policy tools for stimulating growth seem to have run out, the emphasis is increasingly shifting to competitiveness-based economic policy. The NGM's Deputy State Secretary stressed that low wages in Hungary no longer give it a competitive edge and sustained convergence calls for the creation and spread of higher value added activities.

In order to achieve convergence objectives, the Hungarian government established the National Competitiveness Council in March 2017 to deepen professional dialogue and professionally prepare political decision-making in the affected area. Deputy State Secretary Turóczy stressed that the National Competitiveness

Council is a consultative body based on the professional apparatus of the NGM (and of its members) and the professional authority of its members that formulates non-binding recommendations for improving competitiveness in Hungary. The body has so far assessed and crafted recommendations on the business regulatory environment, the supply of competitive labour and economic and social digitalisation. Going forward, after fine-tuning the economy, the National Competitiveness Council will also address structural matters (such as education and productivity) in the course of which it will involve external experts in the professional forum.

Gergely Baksay and Gábor Horváth gave a presentation on the measurability of banking system competitiveness and reported on the new banking system competitiveness index created by the central bank. In the first half of the presentation, Gergely Baksay, the MNB's director in charge of competitiveness explained how the banking system affects competitiveness and the perspectives for investigating this. Banks can contribute to the sustained convergence of the national economy through the efficient and sustainable allocation of funds in the long term. For this, customer needs must be met on the one hand, while remaining attractive to investors (banking system shareholders) on the other hand. Banking system competitiveness can thus be interpreted in two distinct segments: the user side and the investor side. In some cases however, there may be a conflict of interest between these two sides. The customer wants low interest rate spreads, numerous branch offices and as much competition as possible, while profit-oriented shareholders want less competition, higher interest rate spreads and cheaper operation (for instance, fewer branch offices). The MNB's experts deemed that this conflict is resolved over time and the two approaches can become mutually reinforcing in the long term, and a competitive banking system serves the interests of both groups. For instance, the user base may grow in response to a banking system that is more focused on customer interests, which increases return and enables more investments. Although the role of financial intermediation is covered by international competitiveness surveys, it is only assessed partially based on specific criteria. However, a broader international comparison is important for the MNB, and so it developed a banking system competitiveness index for the entire European Union based on objective fundamentals and factoring in both user and investor aspects using the available statistical data and surveys.

Chief economic expert Gábor Horváth explained that the index factors in availability, quality and pricing from the user side, and the aims to measure the banking system's capacity to attract capital from the investor side through profitability, growth opportunities, the operating environment, stability and the availability of new technologies. If we look at Hungary's position compared to other banking systems within the EU, the main shortcomings on the financing side can be identified

in the realm of household loan prices, digitalisation and product penetration. In terms of its capacity to attract capital, improvements are mainly necessary in the banking system's technological and cost efficiency ratios. Improvements in efficiency and a change in mentality would be the strongest drivers of competitiveness gains in the Hungarian banking system. In addition, unexpected innovation can disrupt the status quo at any time, so every sector must shape its services flexibly, looking ahead and factoring in the needs of younger generations.

In the final presentation including the competitiveness section workshop, Zoltán Sipos, professor and head of department at Zsigmond Király University, presented the results of a study investigating the competitiveness of private higher educational institutions that he co-authored with the rector of Zsigmond Király University, Péter Szatmári. Sipos' main research areas are international marketing strategies, postmodern marketing in tourism and the development of international business methods. His key publications address the international business relationship network.

In his presentation, he criticised the current measurement practices of higher education competitiveness in Hungary, which have a flawed approach to the group of higher education institutions in Hungary and fail to sufficiently factor in outcome indicators. He also added that in his opinion, the Hungarian methodology does not allow for international comparison and because only the leading Hungarian universities appear in international rankings, we have no realistic picture of the other Hungarian institutions. Hungarian practice primarily looks at input indicators (such as the number of applicants and admission score thresholds) while only factoring in research results (such as publications and impact factor) among output criteria, while ignoring graduates' knowledge, competencies and employment opportunities.

The author formulated recommendations for improving the standard of Hungarian higher education based on his experiences drawn from private higher education. In his opinion, institutions must react flexibly and quickly to challenges, particularly those associated with generation Z, remain open to new opportunities and innovation and maintain continuous operative ties with market players, which can be achieved through dual training. These require a stable higher education development strategy ensuring, through a reform of the accreditation system and through performance measurement based on learning outcomes, that graduates have competencies tailored to the needs of the market.

The Responsible corporate governance section held an independent workshop for the first time this year. The workshop's topic was presenting the challenges to corporate growth through the example of generational changes in family businesses and corporate governance practices associated with presence on

the stock exchange. *Bianka Parragh*, member of the Magyar Nemzeti Bank's Monetary Council, chaired the workshop and explained the importance of the values represented by the renewed section in her keynote address: the spread of corporate governance practices contributes significantly to improving the country's competitiveness, as transparent corporate governance conducted according to high standard principles improves efficiency and boosts company value.

During the first part of the workshop, the speakers addressed the issues associated with the generational change in family businesses which are not traded on the stock market. *Tamás Madlena*, deputy CEO of the Budapest Stock Exchange (BSE) in charge of business development, presented the BSE's market development efforts in the SME segment. While a large number of family-owned businesses are present on the stock market in Western Europe, this ratio is negligible in Hungary. The BSE has hosted almost 100 bilateral company meetings based on which it has focused on education (the ELITE programme in partnership with the London Stock Exchange) and preparations for the stock market in the medium term. The capital market ecosystem can be established in the long run and with a systematic approach, in which the MNB, as the owner, is a supporting partner.

After Tamás Madlena's presentation, *Dávid Boross*, co-president of the Family Business Network Hungary (FBN-H) and also the second generation owner and manager of the horticultural company Oázis Kertészeti, spoke about the importance of generational changes for family businesses. The topic is particularly important because many of the founders of family-owned businesses accounting for approximately half of Hungarian GDP will reach the age of 60-70 during this decade. Only a third of businesses will survive the first generational change. FBN-H provides assistance for conscious planning and a successful generational change.

Kálmán Nagy, partner at Concorde MB, gave a presentation about the weight of medium and large corporations within the Hungarian economy (approximately 2,000 companies with EBITDA of over HUF 250 million) and the significance of the wealth to be legated (approx. HUF 2 trillion), which may be properly managed by the trustee. This market is currently underdeveloped and inactive in Hungary. If confidence, predictability and experience are improved adequately, trusts may spread in Hungary.

During the panel discussion that followed the presentations, *Márton Michaletzky*, Head of Issuers Acquisitions Division of the BSE, and *András Szabó*, founding owner of iData, joined the speakers to discuss the elements of successful generational change with the audience's active participation. Dávid Boross stressed the importance of the family Constitution in which family and business rules are separated and the conditions of operation are defined.

During the second part of the conference, the financial director of MasterPlast, *Róbert Nádasi* presented the company's history to highlight the conscious development that has always been aimed at bringing the company to the stock market. He particularly emphasised the significance of an external professional board of directors separate from the owners and the role of publicity associated with being present on the stock exchange in helping the company's owners shape its long-term strategy and future prospects and understanding of their company with the help of regular disclosures and investor questions and feedback.

Levente Zsembery, chairman of the Hungarian Private Equity and Venture Capital Association (HVCA), spoke about the objectives and expectations of venture capitalists. Venture capitalists are looking for a credible story (adequate market, product, management, motivational structure) for which they provide smart money: capital, a relationship network, support, feedback and financial literacy training. The presentation's concluding thought was to highlight the corporate governance shortfalls of Hungarian companies.

During the panel discussion that followed the presentations, the Marketing Director of the BSE, *Balázs Bozsik* and OTP Bank's Head of Legal Services, *Zsolt Wieland*, joined the speakers. The participants analysed the development opportunities for the Hungarian capital market which urgently calls for government support, the establishment of trust and the selection of the right companies, along with the right owner mentality.

The Itinerant Conference addressed relevant topics, hosted interesting and credible professional speakers and actively involved the audience. The topic that elicited the greatest response from the audience was generational change for family businesses. The section introduced itself successfully and plans to present new programmes in the near future, in the wake of which the section's members and activities are expected to increase.

The concepts addressed at the workshop can be summarised in the following points and recommendations:

- Fostering the spread of the best Hungarian corporate governance practices which can improve competitiveness within the national economy: this can be achieved by formulating recommendations, defining rating systems or spreading international best practices to serve as examples.
- A successful change in generations by Hungarian small and medium-sized enterprises is a national economic interest that can be a defining factor of competitiveness even in the short and medium run. The matter's economic impact should be better evaluated and an action plan created (drawing attention to the importance of the matter, creating government programmes, communication).

- It is essential to handle certain sectors simultaneously involved in capital and knowledge-intensive activities on a priority basis. Support in the form of monitoring and targeted programmes in relation to the need for R&D&I and the involvement of the available resources would be called for. This group of corporations is most closely affected by technological progress and easing the conservative capital structure would open the door for faster growth and incentivise innovation.
- The importance of developing the capital market. Venture capital financing and the creation of the necessary ecosystem requires development and contributes to dynamic economic growth, achieving a more balanced capital structure and successful generational change for SMEs, which requires the targeted transfer of capital market knowledge to a targeted group of companies.
- The BSE plays an essential role in education and creating the ecosystem.
- The BSE's role as advisor and partner and more conscious communication on this subject.

Report on the Digital Currencies International Conference

Gábor Horváth

On 6 November 2017, the Magyar Nemzeti Bank (MNB) Department of the Corvinus University of Budapest hosted an international conference on digital currencies at the Bálna event hall in Budapest. The event's keynote speaker was *Michael Kumhof*, Senior Research Advisor at the Bank of England, well-known expert on the topic and a member of the editorial team of *Ledger*, the first proofread academic journal specialising on virtual money and blockchain technology. *Ken Lo*, owner and CEO of the ANX Hong Kong Bitcoin exchange, arrived from the private sector and made the event unique with his presentation. There was substantial interest in the conference and the topic, as was clearly reflected by the nearly 450 participants. The first section of the conference focused on central bank digital currency (CBDC), while the second section was centred on Bitcoin and other altcoins as well as blockchain technology.

Dániel Palotai, executive director at the Magyar Nemzeti Bank (MNB), opened the conference; in his welcome speech, he highlighted that innovations and fintech companies may restructure the current financial system in many aspects. Bitcoin and the Distributed Ledger Technology (DLT) serving it enable direct settlement between the participants without a central or intermediary party. Despite the increasing popularity of virtual money, experts disagree: some of them believe that Bitcoin and other alternative, encrypted digital currencies may not play the role of money over the long run since they are too volatile, risky and energy intensive. However, the global trend – according to which the market capitalisation of Bitcoin and other altcoins has substantially increased over the past period – cannot be disregarded. The question is: Will more mature versions of these alternatives, which have outgrown their childhood diseases, represent a challenge later on to fiat money and ultimately also a challenge to central banks? This is the key question not only for central banks and governments, but also for every credit institution.

In his opening address, *Kristóf Lehmann*, head of the MNB Department, expressed his joy at seeing the diverse range of participants which was partly due to the current interest in the topic, and partly to the appeal of the internationally recognised names among the speakers. Making reference to a recently published paper of BIS, Lehmann pointed out that it is not likely that the Bitcoin and other

Gábor Horváth is a senior economic expert at the Magyar Nemzeti Bank E-mail: horvathga@mnbb.hu

digital currencies will be able to take over the role of the currencies issued by states, but he underlined that the distributed ledger technology behind virtual money could be deployed in other areas. Most recently, a few large central banks (BoE, Fed, Riksbank) announced that they are analysing and testing the application opportunities of the technology and even the possible introduction of virtual central bank currencies. In order to clearly see what this is all about, it would be worthwhile to analyse and group digital currencies based on various attributes within the dimensions also discussed by Kristóf Lehmann. For example, based on who the issuer is (central bank or private), the form of the currency (electronic or physical), the participants who have access to it (every participant or only a limited circle of participants), and the type of settlement mechanism (centralised or decentralised).

Based on this classification, if we assume that the central bank is the issuer and the form of the currency is electronic, then we can still differentiate four different central bank currencies available in digital form: (1) currently existing reserves of commercial banks kept with the central bank, (2) the currently non-existing deposit accounts of households and firms kept with the central bank, (3) the settlement operating on distributed ledger basis exclusively available for commercial banks, and (4) digital currency available for the public, registered on distributed ledger basis, but issued by the central bank (this latter one would be, for example, the Swedish e-krona or FedCoin according to current plans).

Michael Kumhof previously worked for the IMF, where he was responsible for the development of the dynamic stochastic general equilibrium (DSGE) model, and was one of the authors of a study entitled “Chicago Plan Revisited”. With his co-author John Barrdear, Kumhof analysed the introduction of central bank digital currencies available for households and companies in their paper titled “Macroeconomics of Central Bank issued Digital Currency”. We might say that there is no other researcher in the world who is better prepared than Michael Kumhof when it comes to presenting the economic effects of central bank digital currencies.

In his address, Kumhof presented the economic effects of the introduction of central bank digital currency under the circumstances as detailed in the abovementioned study.¹ In the presented model, a stock of central bank digital currency equivalent to 30 per cent of the GDP is introduced in such a way that the central bank purchases a stock of government bonds from households, equivalent to the quantity of the issued digital currency. This reduces real interest rates and thereby boosts aggregate demand in parallel with decreasing tax burdens, thanks to which the potential issuing finally increases by 3 per cent in the model. In addition, the inflation target and the counterpointing of business cycles become more achievable through the countercyclical pricing of the central bank digital currency. The presenter pointed

¹ <http://www.bankofengland.co.uk/research/Documents/workingpapers/2016/swp605.pdf>

out that under the conditions they had calibrated, the introduction of the central bank digital currency does not modify the banks' money creating capacity, so the financial system will not be significantly restructured merely by the introduction of the digital currency. The nature of the introduction, a controlled transition and the analysis of novel financial stability risks call for further detailed research.

In connection, *Lajos Bartha*, Director of Financial Infrastructures at the MNB, noted in his presentation that currently everyone is curious to see what kind of challenge FinTech will represent for banks. However, considering the fact that an increasing number of central banks globally raise the idea of central bank digital currency – it is possible that in the end not the innovative FinTech companies, but rather the FinTech central banks will represent the biggest challenge for the traditional banking system. Bartha started off with a money theory and money history summary: he presented the various roles money should take, and the development path through history of legal tender as a result of various innovations. He highlighted that until now central bank currencies in digital form were only accessible by the banking system, i.e. they were not available to households and companies. One of the possible paths of development is that central bank currencies available in digital form will become accessible for the public, i.e. central bank digital currencies will be introduced. Bartha pointed out that a central bank digital currency available to the public can be implemented in several ways; by expanding the traditional account management systems and by the same token, anonymity should also be ensured along with digitalisation, which introduces a further aspect into our taxonomy based on BIS. The presentation listed different motivation factors driving the introduction of a central bank digital currency: placing innovation at the forefront; the availability of risk free digital money; increased financial awareness entailed by the even broader use of electronic payment methods; the support of more efficient monetary policy. All of these factors explain why a number of central banks are contemplating the possibility of introducing a central bank digital currency.

According to Bartha, however, we should also see that this possible introduction raises a number of questions and risks from the perspective of financial infrastructure as well, such as how could a central bank provide the related services, how could the cyber security resilience of the central infrastructure be ensured, how could the central bank meet those legal requirements (for example client due diligence) that the central bank itself requires from commercial banks. Bartha pointed out that the introduction could be implemented in at least three different manners: on a traditional account basis, on a “value basis”, with solutions similar to prepaid cards, where not only the digital form but also anonymity can be ensured, or based on the distributed ledger technology also applied for Bitcoin, among others. According to the presentation, the pioneer Bitcoin has limitations in terms of money functions: the intermediary function is substantially restricted by the low

penetration in the acceptance network, the high exchange rate volatility limits the settlement unit function, while the value retention function is constrained not only by the previous factor, but also by the lack of an institutional background. Some factors limiting the role of the central bank digital currency in payment turnover also arise, which are partly specific Hungarian issues: although the technology is already generally available, cash turnover is still overwhelming in Hungary especially in the countryside; and if the market of payment services becomes one-tiered, who or what shall motivate further innovation or development?

In relation to the final proposal of Lajos Bartha, in the last presentation of the first section, *Gergely Szabó*, economic research expert at the MNB, talked about the potential impact of central bank digital currencies on money creation and the financial system. He demonstrated that a change in format, i.e. a central bank currency available to the public in digital form may potentially bring about some changes in content. After a short endogenous money theory introduction, Szabó said that over the past years a number of reputable researchers, including Michael Kumhof and several of his co-authors, demonstrated money creation in the banking system through lending. He also pointed out that the banks' money creating capacity is attributable for the most part to the fact that only the money created by banks is available in digital form to households and companies. Because the central bank currency is not competitive in terms of form, the money created by the banks is converted less and less to central bank money (to cash). Not necessarily, but all of this may change if a central bank currency became competitive also in terms of form, similarly to the money issued by commercial banks. This is because in this case part of the deposits created in the course of lending could leave the banking system and migrate into central bank money.² But this may place the banking system in a position – especially if commercial banks have no or only limited central bank resource availability – whereby it can only land money if it has preliminary savings, term deposits and investments. Therefore in an extreme case – should sight deposits kept by the banks completely migrate into central bank digital money and there would be no central bank availability – the introduction of central bank digital currency may lead to a system similar to a sovereign money system, where the money is created by the central bank while the banking system is the intermediary of the existing savings. Szabó emphasised that the impact of central bank digital currency on money creation, lending and the financial system and the inherent challenges have already been evaluated by several central banks. Szabó also called attention to the fact that a change in money creation is not a necessary component

² The speakers' opinions are not uniform in that respect. In his model, Kumhof assumed that bank account money will first be fixed in the form of government bonds at the State Treasury and can reach the central bank from there to be sold against digital money. But it is important to see that in Kumhof's model, money continues to be created by commercial banks in the course of lending, and the central bank is available for the population as an option through the above mentioned transaction, i.e. the central bank does not create new purchasing power (in the "pecuniary" sense), it merely provides special liquidity for society.

of the introduction of central bank digital currency; the central bank may decide on the conditions of the central bank digital currency and the central bank's availability and may shape these in such a way that their impact on money creation is limited. Finally – as observed by previous speakers – Szabó emphasised that exploring the effects of the introduction of central bank digital currency requires extremely complex research work; some central banks, for example the Bank of England and Riksbank had made significant progress, but for the time being the critical part of research work is yet to come.

The second section focused on Bitcoin and other altcoins. The first speaker of this section was *Antal Kuthy*, a recognised expert in the field of cryptography, founder of E-Group, a supplier to China Union Pay among others, who presented in detail the technology behind Bitcoin. He demonstrated that one of the strengths of Bitcoin is that it combines in a single applicable use many, already partially existing novelties available in different areas. The technology combines the building of consensus, distributed registration and encrypted procedures. Kuthy emphasised that the technology is not identical with cryptocurrencies but goes far beyond them. The registration of money is only a narrow field among the possible fields of utilisation. Blockchains may store not only numbers representing money, but also any other information such as contracts acting as programs. Kuthy stressed the Ethereum Project which has the objective of broadening the areas of application, i.e. providing a framework that can be applied by any utilisation working on distributed ledger basis.

Kuthy's presentation was followed by *Ken Lo*, the CEO and shareholder of the Hong Kong-based ANX Bitcoin exchange, with over 1 million members from more than 50 countries. Ken Lo emphasised that there is currently a significant "hype" surrounding not only Bitcoin but all other altcoins (Ethereum, Ripple and the other 1,200 cryptocurrencies). The exchange rate is extremely volatile, and huge price fluctuations are possible within a few days or possibly even within just a few hours. A few years ago cryptocurrencies were considered as a toy of some IT geeks, but this has changed by now. The fact that the capitalisation of cryptocurrencies already exceeds USD 200 billion makes it impossible not to consider them. There is substantial media coverage which supports the popularity of cryptocurrencies, but the fact that the CME (Chicago Mercantile Exchange), the US options and futures exchange, will list Bitcoin among its tradable instruments is maybe the most significant piece of news. Currently, the majority of large investors do not have access to Bitcoin, but through CME practically every major investor in the world will have access to this market which may result in substantial additional buying power. In addition to cryptocurrencies, Ken Lo also presented a new application area of the technology in the financing of innovative investments. While in the past the main financiers of innovative and risky investments were venture capital firms, now the

ICO (initial coin offering) may offer a new way in this field and in crowdfunding. The owner of an idea may be able to raise funds simply and globally by issuing its own currency (coin) and use such currency to finance its business. The advantage of this solution is that it can be resolved in a simple manner without any agents, the coin is easily accessible and can be easily sold later on the secondary market. However, this new form of fundraising does not relieve the investors from having to evaluate the actual project, as it may still carry high risks irrespective of the form of financing. And without market makers, completely unrealistic base prices and broad exchange rate movements may occur.

The second section was closed by *Michal Vodrážka*, Director at the Czech National Bank in charge of financial infrastructure. In his presentation, he described the official standpoint of the Czech regulation on Bitcoin and other altcoins. Vodrážka mentioned that for the time being this regulation treats cryptocurrencies as a rather insignificant matter, but that they are closely monitoring their evolution. The Czech National Bank does not consider virtual currencies as actual money, credit institutions are not allowed to trade in such instruments and their banking license do not extend to them. The regulation, governing the prevention of money laundering and terrorist financing analyses the related activities in detail. However, the speaker remained sceptical regarding the necessity of a major regulation on virtual currencies on top of the mentioned money-laundering issues, in addition to having drawn the attention to the serious risks related to virtual currencies. Although a small community is extremely committed to virtual currencies, they play a marginal role in transactions for the time being: in the Czech Republic only 160 stores accept these instruments.

The conference proved to be extremely successful based on the feedback of both the participants and the speakers. The presented topics were discussed from several aspects, thus generating substantial added value and new knowledge. The complexity of the topic and its increasing popularity suggest that this event – together with the research – may well have a continuation going forward.

The Future of Old-Age Provision – Demographic Change, Financial Services and Residential Property. International Expert Meeting Summary

Erzsébet Czinger

On 23 June, 2017 an international expert meeting entitled “The future of old age provision – demographic change, financial services and residential property” was held at the Andrassy University Budapest. The aim of this stakeholder meeting was to discuss Equity Release Schemes (ERS) and their future possibilities, with a special focus on Hungary.

At the meeting, the perspectives of providers, customers, regulatory and supervisory bodies and the academic sector were presented, and thus the topic was discussed from the point of view of nearly all stakeholders involved. Despite the efforts of the organizers, representatives of the life insurance sector, which – according to the current Hungarian legislation – may be the future potential providers of ERS products in the Hungarian market, were absent.

The meeting started with the opening address of *Martina Eckardt* from the Andrassy University Budapest (AUB), who welcomed the guests and stressed that the pension system in Europe is an important challenge which must be solved. The same point was emphasised by *Dietmar Meyer*, Rector of the AUB, who in his welcome address underlined that new approaches and new solutions are essential for Europe to be able to support the elderly.

The next presenter, *Jörg Dötsch* from the AUB, provided a short overview of the project “Integrating Residential Property with Private Pensions in the EU”¹ and presented its partnership, research objectives, initial outcomes and some findings on the Hungarian market. Following this, *John Maher* from the Waterford Institute of Technology, Ireland focused in his presentation on the retirement income pillars, the potential capital and income generation/ release from residential property and different proposals that have been examined in the project (e.g. unitised, leveraged, collective acquisition of residence rights with encashment of those rights post retirement, or, mortgage acquisition of residence with pension released

*Erzsébet Czinger is a Scientific Associate at the Andrassy Universität Budapest.
E-mail: erzsebet.czinger@andrassyuni.hu*

¹ This project is funded by the EU Commission, DG Employment, Social Affairs and Inclusion, VS/2015/0218.

from house, post retirement). Moreover, he outlined some overall parameters and proposals to be discussed, such as desirability and feasibility, obstacles and enablers, beneficiaries/ population cohorts, prerequisites for viability of the proposed solutions, scale of intervention and the relative priority of interventions. He also emphasised that if we want a good model, the scale is very important.

The next speaker, *József Hegedűs*, CEO of the Metropolitan Research Institute, Budapest, gave an overview of the topic from the point of view of a sociologist. He referred to the integrating mechanism concept as articulated by Karl Polanyi (Polanyi's work suggests that the market cannot be sustained without state regulation and a particular type of developed, ingrained social culture). Hegedűs also noted the impact that Airbnb is having on the rental market in Hungary (shifting supply from conventional residential renting to short-term tourism-type renting). Hegedűs envisaged choices regarding three sources of solutions to housing and income/ living standards in retirement: state, market, family. Family solutions include a life support contract in exchange for tenancy or occupancy rights. Hegedűs characterised housing as a special commodity, which thus merits particular consideration and treatment in a societal context. He also stated that privatising housing had the propensity for reinforcing or accelerating social inequality – look east for example to Russia. Moreover, he contrasted active earners (asset poor & cash rich) with retirees (asset rich & cash poor). Hegedűs gave some insights on a 1998 study performed in Hungary with respect to housing equity, which found that a housing equity release product was not feasible. He also referred to a DEMHOW study,² which concluded that releasing housing equity was competing with family care solutions. Lastly, he mentioned a deepening gap between social classes emerging in Hungarian society and highlighted the distribution of income, housing wealth and arrears over different quintiles.

Following this, *Imre Hild*, co-founder of Hild Life Annuity and founder and managing CEO of OTP Life Annuity presented the experiences of equity release programmes in Hungary from the providers' perspective. In his presentation, *The Story of the Life Annuity for Real Estate Model 2004–2009 in Hungary*, he gave an overview of the main features of life annuity for real estate programmes in Hungary and highlighted why such programmes could be viable and attractive in the country. He categorised the reasons into three main groups: structural reasons (e.g. in 2004 there were 170,000 pensioners without children and 40–50 per cent of pensions was spent on maintenance and upkeep), financial reasons (e.g. for most of the target market the apartment was acquired at EUR 1,000 and the average pension was EUR 300 in 2004) and social reasons (e.g. Life Annuity for Real Estate has been known as a private contract since 1950s and it could change the lifestyle of the elderly without

² For more information on the results of the DEMHOW project, see http://cordis.europa.eu/project/rcn/88908_en.html.

admitting hardship to neighbours). Summarising the experiences of the 2.5–3 years of operation, he explained that the average age of the contracting parties had been 71 years old; the average real estate value had been HUF 12.4 million; among the contracted properties 60 per cent had been apartments, 30 per cent had been houses and 10 per cent had been weekend houses; the upfront payment had been 25 per cent of the value of the property at the beginning, which had risen to 40 per cent due to increasing competition between different providers; and the average monthly annuity indexed by inflation had been HUF 30,000. As for the overall results of the 2.5–3 years of operation, Hild indicated that approximately 5,000 clients / 4,000 properties had been involved, HUF 13–14 billion in upfront payment had been made and HUF 50 billion in real estate value had been involved (this latest was a total by all three providers that were present in the Hungarian market from 2004 until 2008, when the ERS programmes came to a halt in Hungary due to a real estate crisis triggered by the global financial crisis). Lastly, among the lessons learnt, Hild mentioned that the key had been to speak with the customer voice; the contract had had to be a simple, no catch contract; security meant more than cash and that life annuity programmes had saved lives and overall, it had proved to be more than a simple financial programme, since it had turned into a social programme at least from the customers' point of view. However, only part of the social problems raised by their clients could be fixed by a financial product like ERS, Hild stated.

The last presenter, *László Kalmár*, CEO of Europ Assistance Hungary (a company active in the field of providing care and insurance services for life annuity programme participants), presented the topic from the customers' perspective. Apart from the presentation of his company's activities, he gave an overview of the number of open cases since January 2014 and he also emphasised the importance of being able to speak with a customer voice and also taking care of some needs of the customers which were not covered by the contract.

Finally, the stakeholder event was concluded with an open discussion, where participants also considered the future of ERS programmes in Hungary. It was mentioned that since 2015 only insurance companies were allowed to offer ERS products in the country, due to a change in the statutory requirements for operating such a business. However, there has not been much interest from the insurance companies so far, according to *Ferenc Szebelédi*, head of the Insurance Supervision Department at the Magyar Nemzeti Bank, representing the point of view of the regulatory and supervisory authorities. All in all, participating stakeholders agreed that despite the present situation, the ERS market could evolve over time.

INSTRUCTION FOR AUTHORS

Manuscripts should be submitted in accordance with the following rules.

- The length of the manuscripts should be limited to 40 000 characters (including spaces) but a \pm 50 per cent deviation is accepted. Manuscripts should be written in Hungarian and/or English.
- Papers always begin with an abstract which should not exceed 800–1000 characters. In the abstract a brief summary is to be given in which the main hypotheses and points are highlighted.
- At the bottom of the title page a footnote is to be given. The footnote contains every necessary information related to the paper (acknowledgement, relevant information etc.). This is followed by the name of the institution and position the author works at, e-mail address in Hungarian and English.
- Journal of Economic Literature (JEL) classification numbers should be given (three at least).
- Manuscripts should be written in clear, concise and grammatically correct Hungarian and/or English. Chapters and subchapters should be bold.
- Manuscripts should contain the list of references with the first and surname of the authors (in case of non-Hungarians the initials of the first name is required), the year of publication, the exact title of the book, the publisher, the place of publication. In case of papers, the exact title of the journal, the year, the volume, and the pages should be indicated. References in the text should contain the surname and the year separated by comma. When citing, the exact page be indicated.
- Tables and figures are to be numbered continuously (chapters and subchapters should not contain restarted numbering). Every table and figure should have a title and the units of quantitative values are to be indicated. Tables and figures are to be made by MS Word and Excel in Hungarian and English. Notes and sources are to be put directly at the bottom of the tables, figures.
- Equations should be aligned to the right and should be numbered continuously in parenthesis. (Chapters and subchapters should not contain restarted numbering.)
- Manuscripts are to be sent to the Editorial Office of the FER only. Papers are peer-reviewed by two independent and anonymous reviewers.
- Manuscripts should be sent as attachment by e-mail in MS Word file. Figures and tables should be sent in MS Excel file both in Hungarian and English.
- In case of further questions related to the manuscript visit the following website: <http://english.hitelintezetiszemle.hu/letoltes/authors-guide-en-1.pdf>

Thank you!

The Editorial Office of the Financial and Economic Review

H-1054 Budapest, Szabadság tér 9.

Phone: +36-1-428-2600

E-mail: szemle@hitelintezetiszemle.hu



Financial and Economic Review