PREFACE

In November 2010 Ph D students from the Kaposvár University and from the Budapest University of Technology and Economics have participated on conferences organized by the Doctoral Schools of both universities, where it was expected to deliver the lectures in English. The present issue "Analysis of Monetary Institutions and Space" of the *Public Economic Series* contains the best papers presented on these workshops.

The topics discussed by the participants have touched a lot of fields of economics, from economic history through finance up to extremely actual problems of spatial or regional economics. Considering the papers from methodological point of view, the institutional approach can be found as well as the application of advanced econometric methods, time series analysis, etc.

Organizers are firmly convinced that the event was successful. The opportunity to discuss parts, chapters or the approach of the Ph D theses with colleagues will certainly inspire all participants to continue their activities and to present their results to a broad scientific audience.

Last, but not least, we would like to express our thanks to *Monika Marton* and *Dr. Felix Stübben* for their technical assistance in preparing this new volume of PES.

Budapest, 2011

László Balogh
Dietmar Meyer
Heinz-Dieter Wenzel

EMIGRATION, REMITTANCES AND THE IMPACT ON ECONOMIC GROWTH

ADELA SHERA

1. Introduction

The role of immigrant remittances in economic growth and development continues to be an important issue for policymakers and researchers. Remittances represent a sustainable flow of financial resources, predominantly from developed economies to developing economies. They are an important source of foreign exchange earnings for the country of origin. Inflow of remittances affects economic growth positively by reducing current account deficit, improving the balance of payment position and reducing dependence on external borrowing (Iqbal and Sttar, 2005).

Inflows of remittances increase the economic growth and reduce the poverty by stimulating the income of the recipient country, reducing credit constraints, accelerating investment, enhancing human development through financing better education and health [Calaro (2008); Jongwanich (2007); Stark and Lucas (1988); Taylor (1992); Faini (2002); Gupta et al.(2009)]. However Chami et al. (2003) find that remittances have negative impacts on economic growth of recipient country because a significant flow of remittances reduce labor force participation and work efforts which lowers output. Thus, the impact of remittances on economic growth and development of recipient country has been controversial.

In case of Albania, a number of studies have been undertaken at micro as well as macro level that directly or indirectly focused on the impact of remittances on growth (Burney, 1987; Arif, 1999; Adams, 1998; Malik et al., 1993; Nishat et al., 1993; Burki, 1991; Kozel and Alderman, 1990; Amjad, 1986; Nishat and Bilgrami, 1991).

The general conclusion of these studies suggest that remittances would have positive effects on economy of Albania in terms of aggregate consumption, investment, reduction in current account deficit, external debt burden and improve education/skills of the households. Furthermore, labor migration is

considered to be a useful source of foreign exchange earnings (Naseem, 2004). Siddidui and Kemal (2006) explored the impact of decline in remittances on welfare and poverty in Albania.
During the current decade since remittances of 1991-1992 has increased in Albania sharply. This massive inflow of remittances contributes in reducing current account deficit, increasing foreign exchange reserves, stabilizing exchange rate and reducing poverty.

Emigrants' remittances represent the most positive aspect of emigration in Albania. During the period from 1991 to 2000, emigration remittances, mainly arriving from Greece and Italy, have been increasing progressively. According to the evaluations of the Bank of Albania, they actually amount to about US$530 million annually, or about 18 percent of GDP. That is why the remittances have been a fundamental factor in determining one of the characteristics of Albania transition, extroversion of the economy (extroversion) (Samson 1996), that is, that local consumption is higher than GDP, while the difference is compensated for by the remittances from emigration and foreign aid.

The theoretical literature on remittances is large, but most theories have tended to follow a common theme. Early approaches to the theory of remittances identified and described various costs and benefits to remitting. The earliest economists such as Whitelaw (1974) mention altruistic motivations for remittances. Lucas and Stark (1985, p. 902) argue that "Certainly the most obvious motive for remitting is pure altruism – the care of emigrant for those left behind. Indeed, more recent theories have focused on the idea that there can be self – interested reasons for remitting well, which nevertheless center of the family. The remittances sent by the migrant are used to care for the migrant's interests.

On the other hand, Poirine (1977) and Ilahi and Jafarey (1999) model the family a bank that finances migration for some members. The borrowers remit funds in order to repay the loans, which are put toward more loans to further the interests of other individual family members. Elbadawi and Rocha (1992) present a detailed review and insightful analyses of the literature on the causes of immigrant remittances, which applies well to all remittances. Wahba (1991) divides remittances into "fixed", which go toward family support and "discretionary" remittances which are investment flows. The fixed remittances depend on the family size and income level. There are numerous papers which analyze the effects of remittances, which they show that in different countries the

remittances are used in different ways and sectors. For example in Pakistan the most part if remittances are spent on consumption, followed by residential investments. In Albania is the same, the most part of them are used for consumption and investment, but in houses.

This way of spending remittances can be interpreted as revealing the remitter's concern that her family has sufficient food, clothing and shelter.

Figure 1.: Using of remittances
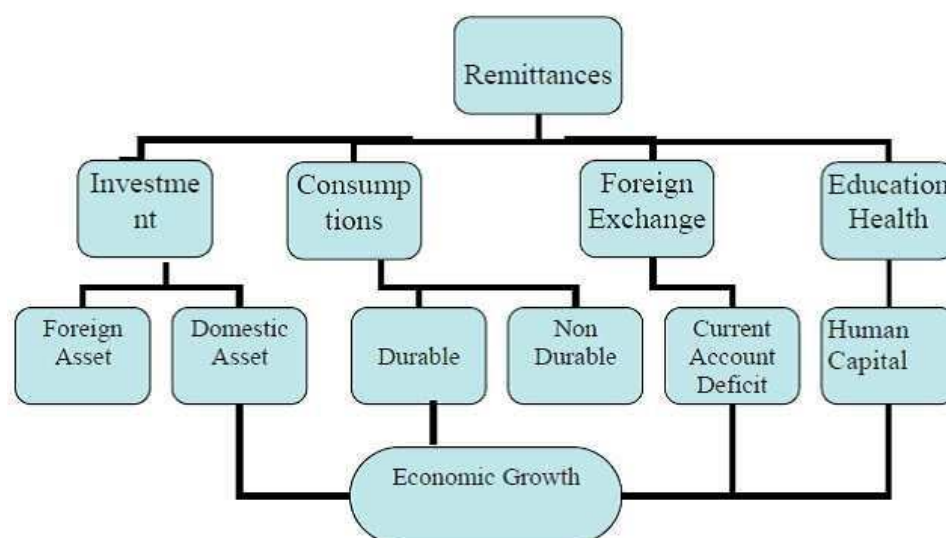
Source: Own analyses

2. Emigration

The emigration of the people from one country to another is one of the most important features for the countries in transition or developing countries. This has an overall impact to the economy and the welfare of the individuals and families.

Although the decision to migrate is based on numerous economic, psychological, social and political factors, economists have focused on wage and unemployment differentials. For risk-averse workers, destination employment probabilities may be a more important determinant of migration than wage rates, at least in the short run (see Treyz, 1993 and Hatton, 1995).

6

Borjas (1991), similar to Molle and van Mourik (1989), has pointed out the importance of political and psychological factors in the home countries as determinants of migration. UN/ECE (1995) claims that cultural and other non economic factors appear to be acting as powerful barriers to emigration from the former Soviet Republics to Western Europe or North America despite vast wage and unemployment differences. For example Albania is one of the countries, which has a higher number of population which has emigrated. Emigration impact may various significantly across these dimensions of social and economic changes.

Table 1.: Effects of emigration

Micro Values for desires and expectancy	MESO collectives and social network	Macro level opportunity structures
Individual values and expectancy. Improving and securing survive , health, status, comfort, stimulation, autonomy, affiliation and morality.	*Social ties* strong ties families and household weak ties networks of potential movers brokers and players *Symbolic ties* Kin, ethnic, national, political and religious, symbolic organizations, *Content of ties transactions* Obligations, reciprocity, and solidarity, information, control, and access to resources of others.	*Economics* Income and employment differentials *Politics* Regulation of spatial motilities through nation status and international regime Politic repression ethnic national, and religious conflicts *Cultural settings* Dominant norms and discourses Demography and ecology Population growth Availability of arable land Level of technology

Source: *Literature review Own Results*

2.1. Main phases in emigration

The emigration optimists: Neoclassical views

In the develop mentalists area of the 1950-1960s it was widely assumed that, through a policy of large scale capital transfer and industrialization, poor countries would be able to jump on the bandwagon of rapid economic development and modernization. Many developing countries became involved in the process (such as the source countries of the guest workers in the

Mediterranean) became involved in the immigration process amidst these expectations of the drawing a new area (Papademitriou 1985: 212). From this perspective (return) emigrants are perceived as important agents of change, innovators and investors. The general expectation was that the flow of remittances as well as the experience, skills and knowledge that migrants would acquire abroad before returning would greatly help developing countries in their economy take-off.[1] Neoclassical economists also tend to see migration in a positive light but it is important to notice that it has no place for remittances (Taylor 1999:65).

Table 2.: Main phases in emigration

Period	Research community	Policy field
Before 1973	Development and emigration optimism	Development optimism; capital and knowledge transfers by migrants would help developing countries in development take off
1973-1990	Development and emigration pessimism (dependency and brain drain)	Growing skepticism, concerns about brain drain; after experiments with return migration policies focused on integration in receiving countries emigration largely out of sight in development fields
1990-2001	Readjustments to more subtle views under influence of increasing empirical research	Persistent skepticism tightening of immigration policies
After 2001	Boom in publications mixed but generally positive views	Brain gain remittances and Diaspora involvement, further tightening of immigration policies

Source: *Dominant phases of emigration: Own results Literature review*

The emigration pessimists

The 1973 oil crisis heralded a period of world wide economic downturn, industrial restructuring and increasing unemployment. Also the great age of international migration had ended. A number of academic studies seemed to support the hypothesis that emigration sustains or even reinforces problems of underdevelopment instead of the reverse.[2] These emigration pessimists have argued that emigration provokes the withdrawal of human capital and the

---

[1] Pennix 1982, Bejler 1970, Kindleberger 1965

[2] Almeida 1973, Lipton 1981, Rhoades 1992, Binford 2003

breakdown of traditions. This would lead to the development of the passive, non-productive and remittance-dependent communities.

Also they argue that remittances are used only for consumption and "consumptive investments" and rarely invested in productive enterprises, and this would lead to provoke inflationary pressures (Russell 1992) and soaring land prices (Appleyard 1989, Rubenstein 1992).

3. Remittances and their determinants

Many migrants send remittances back to their home country. Some do it in favor of themselves and some do it in favor of their family and friends in the home country. In order to understand remittances there are two main approaches for analyzing remittances: the "portfolio" approach and the "altruism" approach (IMF, 2005, p.78). These approaches present two main channels for remitting behavior.

The theoretical debate about the determinants of remittances was triggered by Lucas and Stark (1985) with their ground-breaking paper "Motivations to remit: Evidence from Botswana", which is still the basis of the current discussion and extensions. Lucas and Stark studied remittances on a household level and hypothesized the main determinants to be "pure altruism", "pure self-interest" and "tempered altruism or enlightened self-interest". Any kind of contractual arrangements between the migrant and household left behind can be in the latter category, for example co-insurance, exchange-motives, loan repayment.

The theoretical motives and their effects on remittances are summarized in table 3.

We give a more detailed discussion of the motives below.

Table 3.: Theoretical Determinants of Remittances

Effect of … on level of remittances	household income	migrant income	household shock	migrant risk level	education level of migrant	intent to return	no. of migrants in HH	time
Pure altruism	-	+	+			+	-	-
Pure self-interest	+	+						
Co-insurance	-		+	+				
Loan repayment	+ / -	+			+			+, later -
Exchange motives	+ / -	+			+			
Strategic behaviour	-	+	+					

Source: *Lucas and Stark 1985, page 185*

It is natural to assume that remittances are sent to the family left behind due to altruistic feelings of the migrant. This can be modeled in a Becker type setting where the migrant derives positive utility from the consumption of the family. The migrant thus cares about poverty, shocks, etc. of the family and consequently sends remittances. In this case, there is a positive relationship between adverse conditions of the receiving household and remittances sent, see Table 1.1. Remittances should increase with migrant income (the migrant has more to share) and altruism and decrease with recipient income (Funkhouser, 1995). However, income does not necessarily have a linear effect. As Cox, Eser and Jimenez (1997), demonstrate too, that income may have a different effect at different points of the income distribution.

In contrast to altruism, self-interest is also a motivation to remit. In this case a migrant sends remittances with the aspiration to inherit, to demonstrate laudable behavior as an investment for the future or with the intent to return home. If a migrant wants to invest at home, the household can be a trustworthy and well-informed agent. If a migrant intends to return home, he may already invest in housing, livestock etc. and will ask the family to be the agent. The migrant may also send remittances to invest in his reputation at home. Inheritance may be used as a blackmailing device by the household head to receive remittances. According to this theory, remittances increase with the household's assets and income, the probability of inheriting (dependent on the age of parents, number of siblings, etc.), the migrant's wealth and income, and decreases with risk aversion. Only in the case of the aspiration to inherit, can self-interest be distinguished from altruism in the migrant's behavior and a larger income and or wealth of the household should lead to more remittances.

10

Finally, in a three generation setting, remittances may be sent to parents to ensure that the remitter's own children also take care of him in old age (Cox & Stark, 1994), known as the demonstration effect. Care and transfers have to be visible to the grandchildren generation for maximum effect. A less extreme view of the motivations to remit is tempered altruism. In this case the migrant and the family at home mutually benefit from migration, through some kind of implicit contractual arrangement. Altruism and self-interest can nevertheless play a role here, by making the contracts self-enforcing. The contractual arrangements discussed here are coinsurance, loan repayment and exchange for services.

Another type of contractual agreement between the household and family is loan repayment, for example repaying human capital investment or the cost of migration. A household finances a potential migrant's education if the family implicit lending rate is higher than the market interest rate and the youth borrowing rate is higher than the family implicit lending rate (Poirine, 1997). During the next time period the migrant is able to find a better-paid job in the city or abroad due to the education acquired and will send remittances to repay the family for the initial investment. At this stage the migrant might also become a lender, by financing other migrant family members, which increases overall remittances.

The U-shaped time profile of remittances is shown below in figure 3. In this case, the family contract has the aim of increasing income instead of reducing uncertainty.

Figure 2.: Profile of remittances
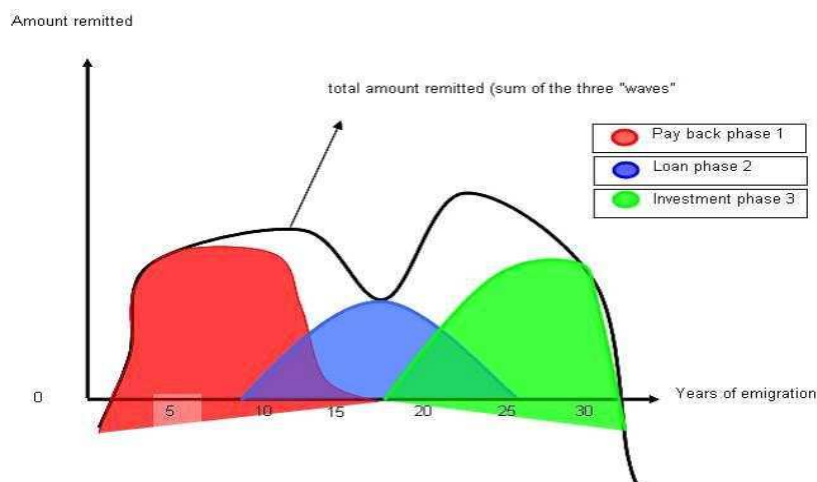
Source: Based on Poirine (1997)

In practice, only paying-back can be measured and there should be a positive link between the migrant's education level and remittances. This could also be interpreted as altruism however due to the close link between education and income. A final contractual arrangement is the exchange motive (Cox, 1987). Here transfers in the wider sense are paid to the household at home for services provided (e.g. child care). The theory can also be applied to remittances, whereby remittances buy various types of services, usually by temporary migrants (Rapaport and Docquier, 2005). If the migrant's income increases, remittances increase. If the household's income increases, thus making the services. more expensive, remittances can decrease or increase depending on the migrant's elasticity of demand. If the migrant's demand for the services is elastic, fewer services will be demanded and remittances decrease. If demand is inelastic, the same services will be bought, but at higher price, which leads to more remittances, despite the higher income of the household at home. Higher unemployment in the home country should mean fewer remittances since less money is then needed to make those at home perform their service (the opposite effect is found for altruism).

The strategic model, first explained by Stark (1995) and later by Stark and Wang (2002), stems from a strategic migration decision made because of wage differentials. Since high skilled migrants usually have a larger amount to gain by migrating, they are typically the first to go and then unskilled workers follow. As individual productivity is unobservable in the rich country, migrants are paid the average productivity of the group with which they are identified. For this reason, skilled workers may have an incentive to remit money home to keep unskilled workers in their home country, since migration of these workers may mean depressed wages for the skilled migrants (Docquier and Rapaport, 1998). The strategic behavior extension says that remittances increase with income and education of the migrant and with low income at home (Holst and Schrooten, 2006), thus again indistinguishable from altruism.
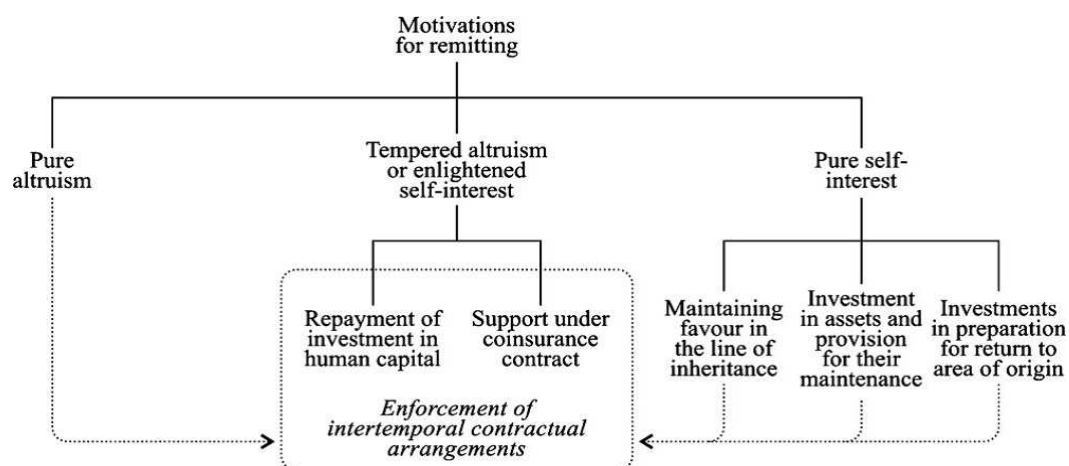
The level of migrants' remittance flows depends on both the migrants' ability, *i.e.* their income and the savings from income, and their motivation to remit savings back to the home country. Of course, the willingness to remit is also determined by the duration of migration (how long do migrants intend to stay abroad, temporarily or permanently?), the family situation of migrants (single, married, with or without children?), and network effects (do migrants move alone, with family members, and do they keep attachments to those left behind?) (For the growing importance of network effects see Munshi, 2003). One way of looking

at the determinants of remittance flows is by analyzing the motives that migrants have to remit money. The literature distinguishes between pure altruism, pure self-interest, informal agreements with family members left in the home country and portfolio management decisions. As Stark (1991) points out, no general theory of remittances exists. The studies that analyze this phenomenon provide useful descriptive evidence and results from empirical research, but they only explain it partly, and are characterized by certain geographical, socio-cultural and temporal limitations.

3.1. Motivations and Consequences of Remittances

In their analysis, Lucas and Stark (1985) focus on *motivations to remit* as an entry point to understanding the complexity of household arrangements involving migration. They outline a taxonomy of motives comprising pure altruism, various forms of pure self-interest, and intermediate motivations that represent contractual agreements between the migrant and the family at the origin. This taxonomy and the relationship between different motivations is illustrated. Migrants are seen to be altruistic if they derive utility from their family's utility, which in turn can be assumed to depend on the family's level of consumption. Lucas and Stark argue that if remittance-senders are motivated by altruism, remittances should be positively related to the migrants' own income, and negatively related to the family's non-remittance income.

At the other extreme, Lucas and Stark point to three forms of self-interested motivations. First, remittance-senders could be driven by the aspiration to inherit. If this is a significant motive, remittances should be positively associated with the assets of the receiving household (i.e. the potential inheritance). Second, migrants could remit money for acquiring assets in the home area and ensure their maintenance. Third, remittances could be sent in preparation for return to the community of origin. These transfers can help establish a base of physical capital, enhance social prestige, or strengthen relationships with relatives and friends. The last of these, write Lucas and Stark (1985:904) 'serves to illustrate how inextricable are motives of altruism and self-interest. In the end one cannot probe whether the true motive is one of caring or more selfishly wishing to enhance prestige by being perceived as caring'.

Figure 3.: Motivations for remitting


```
                          Motivations
                          for remitting
                               |
   ┌───────────────────────────┼───────────────────────────────┐
   │                           │                                │
 Pure              Tempered altruism                    Pure self-
 altruism          or enlightened                       interest
                   self-interest
                        ┌───────┴───────┐         ┌────────┬────────────┐
                   Repayment of    Support under  Maintaining  Investment   Investments
                   investment in   coinsurance    favour in    in assets and in preparation
                   human capital   contract       the line of  provision    for return to
                                                  inheritance  for their    area of origin
                                                               maintenance
                        Enforcement of
                        intertemporal contractual
                        arrangements
```

Source: *Lucas And Starc 1985, page 602*

The consequences of remittances are determined by the purpose of migration and remitting. Remittances increase the inflow of foreign exchange to home countries and thereby increase the demand on domestic currencies. When remittances later is used for consumption or investment it further bring impact on the home economy as either increase in consumption or as in increase in investment. The effects from the increased demand on the home currency are not clear. Some research shows that the real exchange rate (RER) can appreciate as an effect of remittances (See for example Bourdet and Falck, 2006) and give rise to the Dutch Disease. The theory of Dutch Disease sees capital inflow cause appreciation on RER, which makes the export sector less competitive and domestic consumption favor tradable imported goods and non tradable domestic products. This change has a negative effect on GDP if tradable sectors are more productive than non tradable sectors. Countries with high unemployment or underemployment are less likely to experience Dutch Disease (McKinley, 2005, p.2-4).

Whether remittances would be sent with the intentions of a portfolio investment or altruistic helpfulness the remittances will affect the economy different. Capitals that are used in portfolio investment increase the economic activity since investments are done with the intentions to generate profits and productivity, in the same matter as FDI does. Capitals sent in the mind of altruistic helpfulness do not bring any demand for profits and productivity. Whether the remittances are dominating altruistic there is more likely that the

inflow will have smaller effect on the economic activity. The effect could even become negative whether the capital makes the receiver less productive than the productivity the capital generates from being used. The idea that remittances work as compensation capital for poor economic performance was supported by Chami et al (2005, p.77) which found negative correlation between the size of remittances and the home country's GDP for the period 1970-1998.

The reason for the negative correlation between the size of remittances and GDP is due to the moral hazard and asymmetric information. The receivers are assumed to have the same income no matter if they work or not since remittances compensate for low income. The receivers then maximize the utility by spending more time for leisure since it becomes cheaper in an opportunity cost perspective. The remitter's utility is thereby assumed to be a function of their net consumption and the receivers' utility meanwhile the receivers' utility is a function of their work effort and the size of the remittances. The model also assumes the presence of asymmetric information; the remitter cannot observe the receivers' work effort, which induces the moral hazard problem in remittances. This model explains that there may be a problem with decreased productivity when remittances are present and it states that remittances may not be the best tool for development and economic growth.

On the other hand increased remittances per GDP reduce the aggregate output volatility in the home country (IMF, 2005, p.77). Increased remittances may have a multiplier effect on GDP which dampening economic crises and recession. The capital inflow has positive impact in countries with high unemployment, even when most of it goes to consumption (Maimbo and Ratha, 2005, p.5). Remittances give the receiver a higher disposable income, and higher disposable income has a spending effect, which has a positive multiplier effect on GDP. In the long run the higher capital inflow gives possibilities to accumulate capital through domestic saving and investment which have a positive effect on GDP (Bourdet and Falck, 2006, p.7).

Remittances can be used to finance basic consumption and to help improve health conditions. Remittances are also acknowledged to have a positive impact on education in the home country (Bugamelli and Peterno, 2005, p.5). These improvements may have positive effects on poverty in the home country.

3.2. The indirect economic effects of remittances on migrant-sending countries

Most migration impact studies have focused only on the *direct* social and economic effects of migration, that is, the impact on migrants and their households. However, remittances may also have significant impacts on non-migrant households, and hence may reshape sending communities as a whole (Taylor 1999:65). Such indirect effects are usually not captured by remittance-use studies. For instance, research has tended to negatively evaluate consumptive expenses as non-developmental. However, consumptive expenses, provided that they occur locally, can have positive impacts by providing non-migrants with labour and income. This is confirmed by empirical evidence that consumption by migrant households can lead, via multiplier effects, to higher incomes for non-migrant households (Adelman et al. 1988; Durand et al. 1996a).

The same holds true for so-called "non-productive" investments. For example, academics and policy makers have almost universally bemoaned the high amounts of money that migrants tend to spend on housing. This is partly because such "diatribes by academics and policy makers against migrants for their profligate and unproductive ways" (Taylor et al. 1996:411) reflect common elitist views on the irrational spending behavior of lower classes, which in any case have a weak or absent empirical basis. Various empirical studies have reported that construction activities can generate considerable employment and income for non-migrants (Taylor et al. 1996). This also applies to many other expenses such as feasts and funerals (Mazzucato et al. 2006). In this way, the benefits of remittances might accrue to households other than the ones that directly receive them (Taylor, 1999).

These expenses increase consumption levels that may—by easing capital and risk constraints on local production—in turn facilitate local investments by migrants and non migrants alike (Stark 1980; Stark and Bloom 1985). In this way, expenditure on housing and consumption may have significant multiplier effects in the wider economy.

4. Impact of Remittances and Economic Growth

Whereas views on the impact of international remittances on social and economic development in migrant-sending societies have recently inclined

toward the positive side, the impact of remittances on national economic growth and employment are rather unclear (World Bank 2006:xiii). The consequences of remittances on long-term economic development are not well understood (Kapur 2003:vii). There seems to be no conclusive evidence to sustain either neoclassical or dependency theory, because relevant studies have yielded contradictory findings. Other studies are less upbeat and mention the potentially adverse effects of remittances in that they create a strong disincentive for domestic savings and support private consumption of (imported) goods instead of financing investment, which can potentially hamper competitiveness and increase trade deficits (Kireyev 2006).

There are some indisputable welfare effects of migrant remittances.

First, remittances are an important source of income for many low and middle-income households in developing countries.

Second, remittances provide the hard currency needed for importing scarce inputs that are not available domestically and also additional savings for economic development (Ratha, 2003; Taylor, 1999; Quibria, 1997). But the magnitude of the development impact of remittances on the receiving countries was assumed by many scholars to depend on how this money was spent. Thus, a significant proportion of the literature studies the use of remittances for consumption, housing, purchasing of land, financial saving and productive investment. There is no doubt that spending on entrepreneurial investment has a positive direct effect on employment and growth. However, other scholars documented that even the disposition of remittances on consumption and real estate may produce various indirect growth effects on the economy.

Most of the theoretical researches considering the multiplier effects of remittances use models that capture both migration and remittances effects on welfare. They consider remittances as a possible offset to the decline in output suffered by developing countries, caused by the loss of trade opportunities as a result of emigration. The results show that if low-skilled migrants emigrate, the welfare of the source country rises in the case that remittances are in excess of the domestic income loss. If highly-skilled persons emigrate and/or if emigration is accompanied by capital, remittances have a welfare increasing effect for the non-migrants only when the capital/labor ratio of the source economy remains unchanged or rises. If the capital/labor ratio falls, the welfare effect is indeterminate or even negative (Quibria, 1997).

The effect of remittances in the economic growth depends on the way which they are spent. If remittances are spent for consumption, the welfare impact of

remittances depends on the relative factor intensities of traded and non-traded goods (Djajic, 1998). The empirical evidence indicates that multiplier effects can substantially increase gross national product. Thus for example every "migradollar" spend in Mexico induced a GNP increase of USD 2.69 for the remittances received by urban households and USD 3.1 for the remittances received by rural households (Ratha, 2003). In Greece, remittances generated at the beginning of the 1970s a multiplier of 1.77 in gross output, accounting for more than half of the GDP growth rate. Furthermore, high proportions of employment were supported by remittances: 10.3% in mining, 5.2% in manufacturing and 4.7% in construction. And the capital generated by remittances amounts to 8% of the installed capacity in manufacturing. Of particular interest is the finding that spending on consumption and investment produced similar multipliers of respectively 1.8 and 1.9. And contrary to common opinion, expenditure on housing was found to be very productive, with a multiplier of 2 (Glytsos, 1993). By carrying out an econometric test on data from 11 Central and Eastern European countries, Léon-Ledesma and Piracha (2001) found that remittances significantly contribute to the increase of the investment level of the source economies. Drinkwater *et al.* (2003) attained similar results through a study of 20 developing countries. Moreover, their results showed that remittances also diminished unemployment, but insignificantly.

Remittances do not only have positive effects on the source economy. If remittances generate demand greater than the economy's capacity to meet this demand, and this demand falls on non-tradable goods, remittances can have an inflationary effect. In Egypt, for example, the price for agricultural land rose between 1980 and 1986 by 600% due to remittances (Adams, 1991). Along with the positive effects remittances had on Jordan's economy, in the years 1985, 1989 and 1990, they seem to have intensified recession very strongly and generated negative growth rates of over 10%. Other potential negative welfare implications of remittances are the encouragement of continued migration of the working age population and the dependence among recipients accustomed to the availability of these funds. All these could perpetuate an economic dependency that undermines the prospects for development (Buch *et al.*, 2002).

Finally, because remittances take place under asymmetric information and economic uncertainty, it could be that there exists a significant moral hazard problem leading to a negative effect of remittances on economic growth. Given the income effect of remittances, people could afford to work less and to

diminish labor supply. Using panel methods on a large sample of countries Chami *et al.* (2003) found that remittances have a negative effect on economic growth (which according to the authors indicates that the moral hazard problem in remittances is severe).

The long-run motivation for attracting increased remittance inflows is to promote economic growth and development in recipient countries. In line with this ambition, understanding of the appropriate channels through which remittances influence economic performance is essential to formulating sound policies to maximize their overall impact on an economy. The major potential channels of the positive effects of remittance inflows on the growth and development prospects of developing economies include how these remittances impact on domestic investment, balance of payments, ease domestic credit constraints, exports, diversification of economic activities, levels of employment and wages, human capital development and technological progress. On the contrary, remittance inflows may also have adverse effects on the growth and development prospects of developing economies in a number of ways.

One of the critical negative effects of increased remittance inflows on a developing economy is the infection of the Dutch Disease through reduction in international competitiveness. A continuous and significant inflow of remittances can lead to increase in demand for the domestic currency. This increase in demand for non-tradable may further lead to the appreciation of the domestic currency, hence real appreciation of the exchange rate, which in turn reduces the international competitiveness of the country's exports whilst imports are made relatively cheaper. In effect, remittances may, through a number of mechanisms, exacerbate the balance of payments position in the long-run Ahlburg (1991), later Brown and Ahlburg (1999), have argued that remittances undermine productivity and growth in low-income countries because they are readily spent on consumption likely to be dominated by foreign goods than on productive investments.

Theoretically, the degree of impact of remittance inflows on external competitiveness of a receiving-developing country may vary depending upon some specific characteristics. For instance, because unemployment is high in many developing countries, there may not be any significant increase in the production costs of export commodities even in the face of an increased demand for non-tradable.

Can remittances alone trigger economic growth? Probably not. Although remittances play an increasingly vital role in securing and actually improving the

livelihoods of millions of people in the developing world, it would be naive to expect that remittances alone could solve more structural development obstacles such as an unstable political environment, misguided macroeconomic policies, lack of security, bureaucracy, corruption or deficient infrastructure.

Also, legal insecurity of property tends to have devastating effects on people's ability and willingness to invest (see De Soto 2000). However, if development in origin countries takes a positive turn, if countries stabilize politically and economic growth starts to take off, then migrants are likely to be among the first to join in and recognize such new opportunities, reinforcing these positive trends through investing, circulating and returning to their origin countries.

References

ADAMS, RICHARD H. (2006): "Remittances, poverty, and investment in Guatemala." In: Çaglar Özden and Maurice Schiff (eds.), op. cit.

ADAMS, RICHARD H. (2003): International Migration, Remittances, and the Brain Drain: A Study of 24 Labor-Exporting Countries. World Bank, Washington, DC.

ADAM, RICHARD H. (1998): Remittances, investment and rural assets accumulation in Pakistan. Economic Development and Cultural Change 47, 155-173

ADAMS, JR., RICHARD (1991): The Effects of International Remittances on Poverty, Inequality and Development in Rural Egypt. Research Report 86. International Food Policy Research Institute, Washington, DC.

ADELMAN, IRMA, J. EDWARD TAYLOR AND STEPHEN VOGEL (1988): "Life in a Mexican village: A Sam perspective." Journal of Development Studies, Vol. 25, No. 1, pp. 5–24.

ALPER, A.M. AND B. NEYAPTI (2006): "Determinants of workers' remittances: Turkish evidence from high frequency data." *Eastern European Economics*, Vol. 44, No. 5, pp. 91–100.

APPLEYARD, REGINALD (1989): "Migration and development: Myths and reality." International Migration Review, Vol. 23, No. 3, pp. 486–499.

ALMEIDA, CARLOS C. (1973) "Emigration, espace et sous-développement." International Migration, Vol. 11, No. 3, pp. 112–117.

Amjad, Rashid (1986): Impact of workers' remittances from Middle East on Pakistan Economy: some selected issues. The PDR 25:4, 757-782

ARIF (1999): vol 166 Remittances and investments at the household level in Pakistan

AHLBURG, DENNIS A. (1991): Remittances and Their Impact: A Study of Tonga and Western Samoa Canberra: National Centre for Development Studies.

AHLBURG, DENNIS A, AND RICHARD P C BROWN (1997): Are Migrants' Remittances Sensitive to Changes in Their Income? Discussion Paper 217. Brisbane: Department of Economics, University of Queensland.

BEIJER, G. (1970): "International and national migratory movements." International Migration, Vol. 8, No. 3, pp. 93–109.

Burki, Shahid. J. Migration from Pakistan to the Middle East. In D.G Papademetriou and. P.L. Martin (ed). The Unsettled Relationship: Labor Migration and Economic Development. London. Greenwood press

BUCH, C. M., A. KUCKULENZ AND M. LE MANCHEC (2002): Worker Remittances and Capital Flows. Kiel Working Paper No. 1130.

BORJAS, G. J. (1991): \Immigration And Self-Selection," in Abowd and Freeman 27-96.

BOURDET, Y., AND FALCK, H. (2006): "Emigrant's Remittances and Dutch Disease in Cape Verde."International Economic Journal 20, 267-284

BUGAMELLI, MATTEO AND FRANCESCO PATERNO(2005): "Do Workers Remittances Reduce the Probability of Current Account Reversals?". Mimeo.

Immigration, Trade and the Labor Market. Chicago: University of Chicago Press, pp. 27-96.

CALERO, CARLA (2008): Remittances, Liquidity Constraints and Human Capital InvestmentsInEcuador. World Development (www.elsevier/locate/worlddev)

CUKA E., H. PAPAPANAGOS, N. POLO AND P. SANFEY. (2003): "Labor Market Developments in Albania: An Analytical Overview," *Review of Development Economics* 7: 217-27.

CHAMI, R., C. FULLENKAMP AND S. JAHJAH. (2005): "Are immigrant remittance flows a source of capital for development?" *International Monetary Fund Staff Papers*, Vol. 52, No. 1, pp. 55–81.

CHAMI, R., FULLENKAMP, C., & JAHJAH, S. (2003): Are Immigrant Remittance Flows a Source of Capital for Development. IMF Working Papers 03/189.

CONWAY, DENNIS AND JEFFREY H. COHEN. (1998): "Consequences of migration and remittances for Mexican transnational communities." *Economic Geography*, Vol. 74, No. 1, pp. 26–44.

COX, D., ESER, Z., & JIMENEZ, E. (1998): Motives for private transfers over the life cycle: An analytical framework and evidence for Peru. Journal of Development Economics, 55(February), 57-80.

COX, D., & STARK, O. (1994): Intergenerational transfers and the demonstration effect. Boston College Working Papers in Economics 329.

DJAJIC, SLOBODAN. (1986): "International migration, remittances and welfare in a dependent economy." *Journal of Development Economics*, Vol. 21, No. 2

DE SOTO, HERNANDO. (2000): "The Mystery of Capital", Bantam Press, London.

DURAND, J., W. KANDEL, E. A. PARRADO, AND D. S. MASSEY (1996A): "International Migration and Development in Mexican Communities." Demography 33(2): 249-64

DJAJI'C, S. (1998): "Emigration and Welfare in an Economy with Foreign Capital," Journal of Development Economics, vol. 56: 433-445.

EBRD (1997): *Transition Report.* London: European Bank for Reconstruction and Development. Funkhouser E. 1995. "Remittances from International Migration: A Comparison of El Salvador and Nicaragua," *The Review of Economics and Statistics* 77(1): 137-146.

ELBADAWI, I., AND R. D. ROCHA (1992): "Determinants of Expatriate Workers' Remittances in North Africa and Europe," World Bank Policy Research Working Paper 1038 (Washington: World Bank)

GEDESHI I. (2002): "Role of Remittances from Albanian Emigrants and Their Influence in the Country's Economy, "*Eastern European Economics* 40(5): 49-72.

HADERI, S., H. PAPAPANAGOS, P. SANFEY AND M. TALKA. (1999): "Inflation and Stabilization in Albania," *Post-Communist Economics*, 11: 127-41.

HATTON, T. J. (1995): "A Model of U.K. Emigration, 1870-1913," *The Review of Economics and Statistics* 77(3): 407-415.

IMF. (1997): "Albania: Recent Economic Developments," Washington D.C.: International Monetary Fund. February 20, 1997.

IMF. (1998): "Albania: Staff Report for the 1998 Article IV Consultation and Request for Arrangements Under
IMF. (2005). World Economic Outlook 2005. Washington: International Monetary Fund

Ibqal and Sttar (2005): Impact of Remittances on Economic Munich Personal RePEc Archive R Growth and Poverty: Evidence from Pakistan

Enhanced Structural Adjustment Facility," Washington D.C.: International Monetary Fund. April 30, 1998.

HATTON, T. J. (1995): A model of UK migration, 1870-1913. Review of Economics and Statistics 77, 407{415.

HOLST, E., & SCHROOTEN, M. (2006): Migration and money- What determines remittances?
Evidence from Germany. Institute of Economic Research, Hitotsubashi University, Discussion Paper Series, 477.

HOLST, E., & SCHROOTEN, M. (2006): Migration and Money-What determines Remittances?

Evidence from Germany. Institute for Economic Research, Hitotsubashi University
Discussion Paper Series, No. 477.

KULE, D., A. MANÇELLARI, H. PAPAPANAGOS, S. QIRICI AND P. SANFEY. (2002): "The Causes and Consequences of Albanian Emigration during Transition: Evidence from Micro Data," *International Migration Review* 36(1): 229-39.

KOZEL, VALERIE, AND HAROLD ALDERMAN (1990): Factors determining work participation and labor supply decision in Pakistan. The PDR 29, 473-479

LUCAS, R. AND O. STARK (1985): Motivations to Remit: Evidence from Botswana, Journal of Political Economy 93(5): 901-18.

LUCAS, R.E.B. AND O. STARK (1985): Motivations to remit: Evidence from Botswana, Journal of Political Economy, 93, 5: 901-18

LIPTON, MICHAEL (1980): "Migration from the rural areas of poor countries: The impact on rural productivity and income distribution." World Development, Vol. 8, No. 1, pp. 1–24.

SIDDUDI AND KEMAL (2006): Remittances, Trade Liberalization, and Poverty in Pakistan: The Role of Excluded Variables in Poverty Change Analysis

McKINLEY. TERRY (2005): - Why is "The Dutch Disease" always a disease? The Macroeconomic consequences of scaling up ODA – International Poverty Centre, Working Paper 10, United Nations Development programme, Brazil

MAIMBO, SAMUEL MUNZELE AND DILIP RATHA (EDS.) (2005): Remittances: Development Impact and Future Prospects. World Bank, Washington, DC

MALIK S. J., AND N. SARWAR (1993): "Some Tests of Differences in Consumption Pattern: The Impact of Remittances Using Household Income Expenditure Survey of Pakistan 1987-88." The Pakistan Development Review 32(4):699–711.

MANÇELLARI, A., H. PAPAPANAGOS AND P. SANFEY (1996): "Job Creation and Temporary Emigration: The Albanian Experience," *Economics of Transition* 4(2): 471-90.

Naseem 2004 Economic Growth and Development in South Asia, with and without Regional Cooperation

MOLLE, W.T.M. AND VAN A. MOURIK. (1989): "A Static Explanatory Model of International Labour Migration to and in Western Europe," in Gordon and Thirlwall, eds.

MUNSHI, K. (2003): Networks in the modern economy: Mexican migrants in the US labor market, Quarterly Journal of Economics, 118, 2: 549-99.

NISHAT, M., AND N. BILGRAMI (1991): "The Impact of Migrant Worker's Remittances on Pakistan Economy." Pakistan Economic and Social Review 29:21–41.
———. 1993. "The Determinants of Worker's Remittances in Pakistan." The Pakistan Development Review 32(4, Winter):1235–245.

European factor mobility: Trends and consequences. New York and London: St. Martin's Press, pp. 30-52.

FUNKHOUSER, EDWARD (1995): Remittances from International Migration: A comparison of El Salvador and Nicaragua. The Review of Economics and Statistics 77, no. 1:137-146.

GLYTSOS, N. (1993): Measuring the income ects of migrant remittances: A methodological approach applied to Greece. Economic Development and Cultural Change 42, 131±168.

MAZZUCATO, V., M. KABKI AND L. SMITH (2006): "Locating a Ghanaian funeral: Remittances and practices in a transnational context." Development and Change, Vol. 37, No. 5.

MUENT, H., F. PISSARIDES, AND P. SANFEY. (2001): "Taxes, Competition and Finance for Albanian Enterprises: Evidence from a Field Study," *MOCT-MOST*, 11: 239-51.

MURPHY, R. (1998): "Nano exit raises hope for stable Albania," Reuters. September 29, 1998.

MOLLE W.T.M. AND VAN A. MOURIK (1989): "A Static Explanatory Model of International Labour Migration to and in Western Europe," in Gordon and Thirlwall, eds. European factor mobility: Trends and consequences. New York and London: St. Martin's Press, pp. 30-52.

KAPUR, DEVESH (2003): Remittances: The New Development Mantra? United Nations Conference on Trade and Development, Geneva.

KAPUR, DEVESH AND J. MCHALE (2003): "Migration's new payoff." Foreign Policy, Vol. 139, November/December, pp. 49–57.

KIREYEV, A. (2006): The Macroeconomics of Remittances: The Case of Tajikistan. International Monetary Fund, Washington, DC.

KINDLEBERGER, CHARLES P. (1965): Europe's Postwar Growth: The Role of Labor Supply. Oxford University Press, New York.

PAPADEMETRIOU, DEMETRIOS G. (1985): "Illusions and reality in international migration: Migration and development in post–World War II Greece." Interna ional Migration, Vol. 23, No. 2, pp. 211–223.

PAPADEMETRIOU, D.G., AND P.L. MARTIN (EDS.) (1991): The Unsettled Relationship: Labour Migration and Economic Development, New York, Greenwood Press.

PAPAPANAGOS, H. AND P. SANFEY. (2001): "Intention to Emigrate in Transition Countries: the Case of Albania," *Journal of Population Economics.* 14: 491-504.

POIRINE, B. (1997): A Theory of Remittances as an Informal Family Loan Agreement, World Development, 25, 221-241.

PENNIX, R. (1982): A Critical Review of Theory and Practice: The Case of Turkey.

QUIBRIA, M.G. (1997): «International Migration, Remittances, and Income Distribution in the Source Country: A Synthesis», Bulletin of Economic Research, 49,1:29-4International Migration Review 16, 781-818.

TAYLOR, J. EDWARD (1999): "The New Economics of Labor Migration and the role of remittances in the migration process." International Migration, Vol. 37, No. 1, pp. 63–8

TREYZ, G. I. et al. (1993): "The Dynamics of U.S. Internal Migration," *Review of Economics and Statistics* 75(2): 209-214. UN/ECE (United Nations Economic Commission for Europe). 1995.

ILAHI, N. AND S. JAFAREY (1999): Guestworker Migration, Remittances and the Extended Family: Evidence from Pakistan, Journal of Development Economics 58(2): 485-512.

OZDEN, C. AND M. SCHIFF (EDS.) (2006): International Migration, Remittances and the Brain Drain, Washington, DC: The World Bank.

SAMSON, MICHAEL (1996): "Re-evaluating South Africa's Fiscal Constraints on Transformation." *A Consultancy Report to NEDLAC Commissioned by the Economic Policy Research Institute.* Cape Town: EPRI.

STARK, O. (1995): Altruism and beyond: An economic analysis of transfers and exchanges within families and groups. Cambridge, MA: Cambridge University Press.

STARK, O. (1991): The migration of labor, Oxford and Cambridge, MA: Basil Blackwell.

STARK, ODED (1991): The Migration of Labor. Blackwell, Cambridge and Oxford.

STARK, ODED (1980): "On the role of urban-to-rural remittances in rural development." Journal of Development Studies, Vol. 16, No. 3, pp. 369–374.

STARK, ODED AND D.E. BLOOM (1985): "The New Economics of Labor Migration." American Economic Review, Vol. 75, No. 2, pp. 173–178

STARK, O., & WANG, Y. Q. (2002): Migration Dynamics. Economics Letters, 76, 159-164 Rapoport, H., & Docquier, F. (2005). The Economics of Migrants' Remittances. IZA Discussion Paper Series, No.1531.

RATHA, DILIP (2003): "Workers' remittances: An important and stable source of external development finance." Global Development Finance 2003. World Bank, Washington, DC.

RHOADES, R.E. (1979): "From caves to main street: Return migration and the transformations of a Spanish village." Papers in Anthropology, Vol. 20, No. 1, pp. 57–74.

RUSSELL (1992): "Migrant remittances and development." International Migration, Vol. 30, No. 3/4, pp. 267–288

RUBENSTEIN, H. (1992): "Migration, development and remittances in rural Mexico." International Migration, Vol. 30, No. 2, pp. 127–153.

WAHBA, S. (1991): "What Determines Workers Remittances?" Finance and development, Vol. 28, No. 4, December, pp. 41 - 44.

WORLD BANK (2006): Global Economic Prospects 2006: Economic Implications of Remittances and Migration. World Bank, Washington, DC.

Simple Taxation System – Fair Taxation

Attila Bánfi

Expectations vis-à-vis any system of taxation are contradictory. Taxes are a typical manifestation of burden sharing, that is, they are a primary source for covering community services, yet a substantial reallocation of incomes takes place among the actors of society/economy via taxes. It is a general expectation that the system of taxation should be *efficient* and *fair*, and that *tax administration* should be of the least possible *cost* (from the viewpoint of not only the tax administration staff, but also of the taxpayers).

According to the *efficiency* (neutrality) expectation, it is desirable that taxes should distort the behaviour and decisions of economic agents as little as possible. A fully inactive state in terms of its economic functions (which it has never been and will never be) intervenes in the economy through its mere existence, and thereby it unequally influences the distribution of social benefits and social disadvantages. In the case of the tax systems of the modern state capitalist regimes, which deliberately undertake economic functions and which are frequently regarded as attractive, particularly these days, efficiency aspects can hardly be found in their direct form and, in all likelihood, only with great difficulty even in their indirect form. Their impact, however, is more complicated than this as taxes are incorporated into the decision-making systems of social groups, that is, producers and users (consumers), prompting all actors to go for *tax optimisation* (adapting to changes in yields and prices in general due to changes in taxes by transforming the structure of production or consumption, if possible, and in recent decades preferring countries that offer more favourable conditions for taxation internationally when selecting the venue of taxation).

Fairness (just burden sharing) would prevail, if two taxpayers of the same financial position prior to taxation would remain in the same position after taxation as well (this could not be met on account of what has been written in relation to the efficiency criterion), or if taxes were less of a burden on lower income people and essential foodstuffs and consumer goods than on higher income people and non-essential articles. Excessive illusions cannot be entertained even in relation to meeting this expectation either. According to a frequently voiced practical argument there is a top limit on the progression of tax

rates on the part of income that can be centralised because it is necessary to maintain an interest in increasing incomes (and in their declaration) even among the richest. Based on this logic, the basic principle of modern taxation can be formulated, according to which the tax receipts needed by the state presuppose *a large number of taxpayers covering the widest range of society,* and cannot leave even the essentials of life intact (thus, not even the lowest income segment of society, which may be spared of the payment of income tax).

This principle of income and turnover taxation supplemented by taxes on additional possible tax bases (e.g. certain types of real estate, large inheritance, natural resources used or utilised) could be consistently enforced, if governments continued to aim at abiding by justness and fairness within the limits referred to, or at least they attempted to rectify obvious unfairness from time to time. Experience, however, does not always substantiates intention. Burden sharing is exceeding unequal and unjust among the members of society, partly on account of tax evasion, which is very frequently ordinary tax fraud[1] or tax optimisation[2], and partly on account of the tax benefits of economic groups massively capable of enforcing their interests.

Taxpayers have unequal access to the various techniques of tax avoidance. Hungarian family businesses (which are small and not rarely medium-sized undertakings) rather frequently employ their employees at the minimum wage, or at wages substantially below their actual earnings with a view to reducing taxes on wages and wage-type incomes, and particularly the contributions, which are of even higher rate than taxes. At the other end of the scale, the truly high corporate and individual incomes, and – if the state also levies an estate duty, large estates tend to land in a tax paradise (offshore companies). The taxation of this social group is the result of characteristically deliberate tax optimisation[3]; recently, the state has had an excessive tendency to take into account this fact

---

[1] More on tax avoidance and tax fraud: Melinda Slánicz (2010)

[2] Its objective could be a reduction in the amount of tax, or tax avoidance and it could also be the management of assets, which "bears in mind the minimisation of underlying taxes, for instance, capital gains tax, various profit taxes or the estate duty, which could reach as much as 40 percent in a few countries ... In America and in Western Europe, the latter, that is, passive asset management is the most frequent reason for using offshore companies." (Péli, 2010)

[3] The subject of tax optimisation determines consideration for his public consumption practically himself (which holds even if he does not consider this aspect when making his decision). Frequently, he substantially undervalues his personal public consumption (and that of his dependents, which should be mentioned at this point), and through this, he not only fails to exercise the expected fairness, but has other social groups to finance his own public consumption. The public consumption of the richest members of society is characterised not only by making use of usual everyday services. Government support of airports, the road network or cultural institutions (just to mention a few examples, which are overrepresented in the consumption of this group) is possible only from tax revenues paid (or generated in the future). Members of this group enjoy similar subsidies also when they are shareholders or members of the top management in sectors benefiting from tax preferences.

when amending the system of taxation (for instance, when abolishing the estate duty). Consequently, a hidden redistribution is affected to debit taxpayers unable to avoid taxation or to optimise their taxes, that is, to debit those living off wages and salaries, and through the not rarely excessive taxation on turnover and consumption at national economic level. There can be no doubt that tax revenues lost through tax avoidance, tax fraud and tax optimisation must be secured somehow from somewhere[4].

Whether there are endeavours to enforce fairness and equity in taxation, or whether these are waived, betrays a great deal on the moral status of a society and primarily, the political and economic elite in power. The collection of certain taxes, such as estate duty on large estates, or property tax on high value industrial property, or the obviously generated mining royalties in full points to the pursuance of criteria of social ethics, while the absence of their taxation means that these criteria are disregarded and their significance is not recognised. This is a particularly sensitive problem, if the given society is characterised by "a rapid increase in the generation of assets and an increasingly intensive inequality of chances". (Örkény, 2007)

The requirement of simplicity and clarity in the tax system, and hence of minimising the cost of administration, which also depend on factors other than these, orientates towards a tax system that is as simple as possible. The simplicity and consistency of the system of taxation and the budgetary coverage function presuppose one another. If the system of taxation is simple, then there are no tax benefits whether by rate or tax base, and there is no taxation by bands in it, which is one type of tax benefit. These "deficiencies" guarantee the levying of an equal tax burden on the various factors of production (e.g., employer and employee): this is where simplicity and consistency meet. If these are met, the magnitude of the tax rates depends only on the targeted size of the tax revenue.

Enter the reversal of the above argument. In practice, those imposing taxes introduce various benefits, allowances into the system of taxation for reasons of stabilisation and redistribution. Thus, the tax burden on an individual taxpayer becomes a function of his social and economic status (employer or employee; native or foreigner; has a family or not, etc.). To some extent the law provides an opportunity for choosing this position, so the tax payment obligation may be reduced. Owing to this, the system of taxation is not calculable, and the coverage function is impaired. The stabilisation and redistribution functions of the state

---

[4]In the case of more complicated systems of taxation (practically, the taxation systems of all modern states are like this) "taxpayers unable or unwilling to engage in tax fraud (with some euphemism, tax optimisation) experience over-taxation... In a society burdened with excessive hidden economy, hiding of incomes and tax fraud, it is not possible to drastically simplify the system of taxation overnight." (Kürthy, 2010)

can be taken out of the system of taxation because these functions can also be met on the expenditure side of public finances.

The simplicity of the system of taxation does not mean an end to the major types of taxes, only their harmonisation. In theory, a drastic streamlining (a single kind of tax, the simplest type tax system) is possible but, in practice, tax competition among countries under the circumstances of economic openness excludes this. *If, for instance, there was only the value-added tax at an appropriately high rate for covering the government's need of tax receipts existing in an economy, the taxpayers of that country would obtain the lower value-added tax rates of other countries. At the same time, the actors of other countries would, with the same finesse, obtain the low tax rates of the simplifying country. The end result is drastically destructive tax competition.*

1. Here and now: in 2009 and 2010

There is no eternal measure of social justice and fairness, it must always be considered within the given context, and under the given conditions that is always here and now; endeavours to their enforcement should always be relative to the former or current characteristics. International practice easily offers examples of and arguments for the application or discarding any type or any kind of taxes. These, however, cannot be sufficient for choice. The characteristics of the recent past and the present of a society must (should) determine the character of its system of taxation at all times.

The global crisis hit Hungarian society in the autumn of 2008 – finding a substantial segment of it in a rather weakened condition. Following the transition, members of society, who had earlier been deprived of many things, desired to obtain a larger home, a weekend house, travelling, modern cars all at the same time– generally speaking, the very fast imitative following of the Western (American) model of consumption. The income position of a not insignificant part of society did not enable this; hence they ran heavily into debt. The crisis was tough in taking its victims from the social strata which otherwise moderately drew in external (bank) funds to finance their consumption; those becoming unemployed and families with only a single wage-earner very frequently landed in a debt trap. The crisis threw light on a great many things, illusions vanished in a moment. Naturally, belief in ongoing growth and individual life strategies built on it became the first victims. It became evident that not event real estate is a safe harbour as it does not maintain its value, and

the savings of many years and decades invested in real property lost their value overnight – there is no demand for homes and houses even at low prices. As a self-reinforcing process, the crisis impaired the prospects even of families on which it had no direct impact. The curbing of consumption became the social norm as against the earlier consumption orientation. At an individual level, the thumb rule became to wait and see until economic growth is stabilised and the rate of employment improves in a lasting and perceptible manner. At a social level, however, only with a vigorously growing global economy or at least main export markets can economic growth be envisaged in the absence of internal, national growth of consumption, and in Europe it would do no harm if this was supported by growing domestic consumption. In the autumn of 2010, when taking stock of the magnitude of economic growth in Europe, it cannot always be decided whether the cross bar is really quivering, or it only seems to do so to us. There is no doubt that Europe drives Hungary's growth – to the extent that Europe is capable of, without a sustained rise in domestic consumption.

In the period between the transition and the global crisis, large estates were accumulated in Hungary. The large-scale privatisation cases were always results of *individual* negotiations even if they were put out to tender, and even more so, if the model was developed by the management of the company to be privatised. If somebody had access to a valuable piece of the common property in the course of privatisation, it was equivalent to obtaining social support for the launching of his individual enrichment, and subsequently, for the further accumulation of assets that could not be replaced by any other thing.[5]

It is an international experience that favoured sectors are generally able to enforce their interest in political and governmental decision-making. Hence it is not surprising that this remained characteristic of large Hungarian companies even after market mechanisms gained ground. (The best known Hungarian example is that of lex MOL.)

At the time of the outbreak of the crisis, it seemed for a while that a positive outcome was possible, a new order of values could be created if mankind or at least the societies providing life models are bright enough and draw the lessons of the earlier model based on over-consumption. According to the expectations

---

[5] If a person was sufficiently talented, and the company privatised by him had a valuable market profile, the company managed to produce its purchase price in the course of a few years of operation. If the state was the seller, the purchase price was always substantially below what it could have been. (This is indirectly substantiated by the transactions when the company was resold within a short period following privatisation. Almost in every case, the on-sale prices were higher, by orders of magnitude, than the price applied in privatisation. The opposite case, if any, indicated individual unsuitability for managing the assets and operating the company, etc., rather than tough and successful bargaining by the state.) The new owner took out credits always with the new acquisition as security. He could do so as he had obtained the coverage.

of many, the new model could have focused on the conservation of the natural environment. When enforcing the environmental aspects, the issue of justness and fairness is equivalent to enforcing the principle "let the polluter pay", i.e., damage caused to the environment should be restored by the polluter. Now, it seems that there is not going to be any change in models, and this cannot be comforting. It is no less so if the change in models does not take place because the crisis is not over yet, as we are only at its beginning, or at best in the middle of it.

A characteristic of the 20th century reaching well into the beginning of the next one has been the growing pace of tax centralisation and the growing ratio of taxes on income from work (including social security contributions (Stiglitz, 2000. p. 398). Ecologically oriented tax reforms link the tax receipts collected on account of polluting the environment to the reduction of the burden on labour – as if designating the preferred subjects of just and equitable taxation.

In the following, I am not going to address the general features of the Hungarian system of taxation, not even to the point of introducing the problem. I shall deal only with a few kinds of taxes, which, if applied by the government, the Hungarian system of taxation would be more just and more equitable for the majority of society carrying the tax volume.

2. Royalties due to the state, realised by companies

The state is entitled to land use and mining royalties on the utilisation of arable land for other (industrial, communal) purposes, and the raw materials and energy carriers brought up to the surface. The Hungarian government waives not only these natural royalties, but at times, as in the case of electricity generated by the Paks Nuclear Power Plant, generates a royalty of a financial nature (dotation) for the other power plants. At other times (Robin Hood tax, the special tax on the energy sector), taxes of an amount substantially below the amount of the natural or ordinary monopoly royalties undoubtedly due to the state are imposed for a transitory period of two or three years, provoking reactions from those concerned and the intellectuals moulding public opinion with a vehemence depending on party sympathies. Generally, as in the case of solutions that are not clean-cut, the trap here, too, is that the roles are reversed: payment of these transitory taxes appears as sacrifice on the part of the eternally favoured, while the regular transfer of the royalties generated by them in the absence of extensive individual bargains to the state, should be part of the ordinary course of business.

The privatisation of the energy sector *generally* lacked a professional concept by the government. The result is well-known: either the company developed its own privatisation model (MOL) or, as it happened in the case of power plants and electricity providers, competing groups of foreign trade investors, the "buyers of continental influence" (M. *Szabó*, 2006) divided the Hungarian electricity market among themselves with state assistance in the mid-1990s. The government (or to be more exact, ÁPV [the state privatisation agency] which regarded the replenishment of the empty state coffers as its task as soon as possible based on the mandate given by the prime minister)[6] not only made it easy for them to acquire markets, but even guaranteed an eight percent return on assets (subsequently raised to ten percent). In the interest of privatisation revenues totalling a billion dollars, the government ceded the mining royalties generated year-after-year constituting stable revenue for the community to them.

According to experts in environment protection, virtually every segment of the energy sector is characterised by the realisation of high extra profits and royalties; cross-financing is also typical. The royalty, the extra income are consequences of their monopoly positions. The members of the Levegő Munkacsoport [Clean Air Action Group], who developed a complex eco-social reform of public finances in 2008 (Lukács – Pavics - Kiss, 2008) quantified the royalties generated in the electricity, natural gas, crude oil and the Hungarian hydrocarbon sector in general, and retained within these sectors in other works (Pavics – Kiss, 2009). Below, I restrict myself to the presentation of those royalty items only in the case of which interrelations are self-evident requiring no additional explanation[7]:

- MOL buys crude oil at the lower Ural (Med) price, yet it applies the higher Mediterranean (Brend) prices as the price base in calculating the domestic prices of petroleum products; the amount of the extra profit realised by the company in this way amounts to HUF 20 billion a year;
- cheap electricity produced by the Paks Nuclear Power Plant which in the hands of the state creates extra profit – hidden subsidy – at the power plants; the amount thus distributed among the power plants totalled HUF 100 billion in 2007 (Lukács – Pavics – Kiss, 2008). The profits of the Paks

---

[6] "At the time of the Horn government 'the new government escaped from the monster of empty state coffers and the threat of the IMF (the World Bank would block the standby reserves of the country) by quickly selling the power stations and the electricity providers. ... the *revenue of one billion dollars* generated by this transaction was then regarded as extraordinarily high ... a genuine assessment of the condition of the assets has not taken place, similarly, nobody examined the damage to the environment caused till then in any depth. ... energy policy was essentially shaped from ÁPV Rt.' " (M. Szabó, 2006) (Highlighted by B.A.)

[7] I highlight only the energy sector and not even that in full from the proposal of the Levegő Munkacsoport (CAAG), and another study of the authors on the domestic subsidisation of fossil energy.

Nuclear Power Plant (to put it more accurately, the price of electricity generated by it) was reduced by the eight, and subsequently ten percent return on assets guaranteed for the power plants as early as in 2006, and also thereafter when the state's obligation to do so no longer obtained;

The major outflow of profits can be substantiated also by an EU comparison. In Hungary, the price of electricity sold to households was 7.7 percent higher than the average price in the European Union. Taking into account purchasing power parities, prices (more accurately purchasing units) were the highest here, exceeding the incidentally lowest French prices (purchasing units) more than twofold. "... the difference between the price for Hungarian households on 1 August 2009 (0.1605 euro/kWh) and the transfer price of 0.040 euro/kWh of Paks was about fourfold." (Pavics – Kiss, 2009, p. 10).

Buyers paid a total of one billion dollars for the Hungarian electricity market in the mid-1990s – this is hardly twice the amount of the 2007 royalties regrouped at the Paks Nuclear Power Plant and left with the companies instead of being collected by the state.

- the mining royalty on hydro-carbons, whose amount should be due to general government is the difference between the local cost of production and world market prices. In 2008, "MOL paid HUF 115 billion to the budget" on the production of crude oil and natural gas instead of the HUF 276 billion generated. In that same year, the amount of the mining royalty realised by coalmining as extra profit was HUF 10 billion.[8] (Pavics – Kiss, 2009, p. 12);

- The price of electricity purchased from the Ukraine is 33 percent lower than the world market price. In the absence of competition, a few electricity traders share the difference. At least half the foreign trade royalty generated is due to the state; in 2006 this amounted to HUF 23 billion. (Pavics – Kiss, 2009, p. 14)

The highlighted items quantified for one of the years between 2006 and 2008 alone together correspond to 35 percent of the budget deficit of HUF 870 billion planned for 2010, and 16.2 percent of the planned revenues from personal income tax (HUF 1,881 billion). (*Bill* on the 2010 Budget)

---

[8] The basis of the mining royalty is termed as the "generated value" of the mineral raw material. Its rate is 12 percent in the case of crude oil and natural gas, 5 percet in the case of non-metallic mineral raw materials and 2 percent in the case of other solid mineral raw materials and exploited geothermal energy. No mining royalty needs to be paid on crude oil produced using increased efficiency procedures. The rate of the royalty can be reduced for reasons of public interest. In the case of mining activities exercised on the basis of concession contracts, the minister establishes the rate of the mining royal individually.

3. Towards a modern local governance: taxation of industrial and commercial properties

It cannot be regarded as fortuitous that all Hungarian governments (finance ministers) of the past one-and-a-half decades – with the exception of the one in power for the past few months – attempted to make property tax mandatory and *value-based*.[9] Noting the social and political resistance (decrease in popularity), sooner or later the governments retreated with the exception of the Bajnai government.[10] There is no uniform and value-based property taxation in Hungary[11], yet I believe it could hardly be supposed that this state of affairs could be maintained in the medium-term. And that is so not only because property tax is suitable for channelling hidden incomes back into sharing the public burden, but also because it has somewhat of a wide range as a tax base, and property is a tax base that cannot be carried offshore – and other exemplary countries pursue uniform and value-based property taxation without exception. International experience proves that property tax is generally one of the least accepted kinds of tax,[12] but a fair consensus can be developed along principles that can be well delineated. These principles are the following:

---

[9] It is since 1991 (Act C of 1990) that Hungarian municipalities are authorised to decide independently whether or not to impose local taxes on the population living in their settlement, the local undertakings or the possible objects of taxation and if so, whether they wish to impose all of the local taxes without fail, or only in part, or not at all. Municipalities are authorised to deviate downwards from the centrally maximised tax rates at their discretion, but not upwards. The absence of the *mandatory nature* of local taxes counts as a curiosity even in international practice.

Typically, municipalities impose property tax on the basis of *area* (max. HUF 900 /m²/year) – in recent years, they have frequently imposed property tax on household property also. Earlier they tended to impose the tax on non-housing property.

In the case of a *value-based* tax, the valuation of real property and the maintenance of such values requires a large staff of well-qualified experts, as well as the accurate documentation and administration of the data of the individual properties (or at least of the districts or value zones, which can be regarded as having the same attributes). In Hungary and in the former socialist countries, area-based property tax is characteristic because data available from the land registry are sufficient for this.

[10] The 1996 proposal by Lajos Bokros, the finance minister of the Horn government, and subsequently the 2003 proposal by Csaba László (finance minister of the Medgyessy government) not even reached Parliament; the October 1998 submission by Fidesz was taken off the agenda of Parliament very quickly. In the autumn of 2007, at the time of the Gyurcsány government, Parliament adopted an amendment to the Act on Local Taxation, pointing towards a uniform property tax, but the members of Parliament withdrew this via a subsequent amendment.

[11] In 2009, the property tax proposed by the Bajnai government (and adopted by Parliament) was annulled by the Constitutional Court.

[12] Property tax is a *visible tax*, because (in contrast to turnover tax built into prices, or income tax, which is deducted), it is paid by the taxpayer and it can be both advantageous and disadvantageous. It is disadvantageous because knowledge of the amount of the tax may increase the resistance among taxpayers to the tax. For the same reason, it may be, or it could be, favourable because it could direct the attention of the taxpayer to the service provided in exchange for the tax, demanding that account be given for the use of the tax paid.

- practically the entire profession agrees that what is desirable is *a tax base as wide as possible and a tax rate as low as possible*, that is, it would be expedient to make the imposition of a *mandatory*, *uniform* and *value-based* property tax of around *0.2 percent* on any property within the confines of a municipality irrespective of the purpose of utilisation;[13]
- a property tax paid on housing property deductible from *personal income tax*, and property tax on commercial property *deductible from the local industry tax* is largely regarded as equitable;
- it is necessary to clearly define the range of exemptions, the circle of those exempt from the obligation to pay the tax (low income people, large families, pensioners);
- the database requirement of the method of valuation chosen cannot be excessive, and first and foremost, it must not violate the sensitivity and personal rights of the citizens;
- property tax is largely a *local* tax, or *shared* (central and local) tax – irrespective of whether it is collected by local governments or the central (regional) tax administration. Generally, the latter is more suitable for this task.

The point of departure of the tax introduced as of January 1, 2010, was populist (property tax = luxury tax). The so-called property tax (or asset tax) applied to real property and means of transportation (road vehicles, vessels and means of air transportation) above a certain value limit. This value-based property tax should not have been paid on all real property, but *housing property* in excess of a market value (or rather calculated value) of HUF 30 million, and its amount was not deductible from personal income tax. This tax would have affected only about 200,000-250,000 of the 4.2 million housing properties in Hungary.

In an international comparison, the value of privately held housing property in Hungary is substantially more significant than what would follow from the living standards and income relations of the households. The desire of the people for a home of their own, and their ties to it is not typical in an international comparison, this explains the frequently excessive burden undertaken to buy the first home, and then subsequently more valuable homes. A home with a market value of thirty or even forty million forints does not belong to the luxury category even in Budapest, and definitely not outside Budapest, and *their source is* characteristically *income from work* (that is taxed income), hence insisting on

---

[13] Property below the value of HUF 2 million is normally excluded from the range of taxable property for rational reasons as in their case, the cost of collecting the tax exceeds the amount of the tax collected.

the principle of "let the rich pay" against their owners is, in most cases, not only unjust but also untrue.

Prior to and even independent of the introduction of the uniform property tax, it would be expedient to impose a property tax precisely on the largest *industrial and commercial properties* with a mandatory nature in accordance with international practice at a rate higher than that usually applied to housing property, as a central tax or at most, *a tax shared with* municipalities. There is every reason to reduce the increasing differences between settlements associated precisely with the differences in the number and economic potential of companies operating at the individual settlements. The separate handling and target-oriented utilisation of the part of the property tax proposed on non-housing property to be absorbed by the central budget offers an opportunity for this.

Presumably, the *ecological footprint* of these large industrial and commercial estates is a multiple of that of a smaller settlement and is not restricted to the given settlement, but gives rise to changes in a substantially larger area. With regard to commercial property, this circle includes the shopping centres and the logistics centres of commercial networks. The implementation and operation of the major industrial and commercial properties required the construction of parking lots, roads leading there, that is, the covering of large areas. It is well-known that this has a detrimental impact on the water balance and flora and fauna of the soil, as well as the local climate. Produce get polluted by materials detrimental to health in a wide band along roads. For this reason, there is all the justification for breaking the pace of greenfield investments, instead the so-called rusty zones already interfered with should be preferred.

In 1950, the share of areas withdrawn from agricultural cultivation did not reach eight percent of the territory of the country, and it made up no more than 11.5 percent even in 1990, which is the year of the transition. Between 2000 and 2006, it seemed that this share was stabilising at a relatively high level of 17 percent, but the 16.4 percent of 2009 indicated some improvement. Close to one-fifth of the areas withdrawn from agricultural cultivation (19 percent) is re-categorisation as internal area, and the reason for this is almost exclusively the construction of industrial and commercial properties, shopping centres and, naturally, large housing estates. The appreciation of the Budapest agglomeration area is indicated by the fact that 24 percent of the nationwide figure for inclusion in internal areas is attributed to Pest county. (Withdrawal from agricultural production to the largest extent is in any case characteristic of this county, but this figure is only 14 percent of the nationwide.) (CSO)

A non-recurrent *land protection allowance* is payable in the event of utilising arable land for other purposes. Although the least possible size of area can and

should be used for purposes other than agriculture according to Act CXXIX of 2007, everyday practice (large parking lots and/or warehouses beside the facilities; shopping centres of at most two or three levels, etc.) does not indicate full compliance with the law. It seems, for potential investors weighing their decisions, the quality of the arable land to be used is not a priority aspect. In the case of arable land of a quality better than the average, the allowance payment obligation substantially higher than before still fails to exert a substantially restrictive impact.

A thorough analysis of expected environmental impact must (should) be a necessary part of the licensing process, preceding the establishment of new industrial and commercial properties. At the same time, the damage and other environmental impact caused by already operational properties should definitely be modelled prior to determining the tax rate.

The genuine value of industrial and commercial properties is obscure, while they should be individually determined for specifying the tax base at all times. These figures are not available in any registry, as it is not the properties themselves, but their owner companies that are the subjects of sale and purchase transactions.

A property tax imposed mandatorily on large industrial and commercial properties on a value basis would substantially reduce the possibilities of bargaining with the local governments, which is undesirable from the viewpoint of society as a whole. In the case of the local industry tax, so much favoured by local governments and companies and undertakings with a view to attracting a large company (relocation of its seat), local governments readily employ preferential rates (of even zero percent).

The total freedom guaranteed for Hungarian local governments with regard to the application of local taxation has a number of negative consequences, of which perhaps the most severe is that it restricts the coming into being of modern local governance, particularly at the settlements having bodies of representatives that are more patriotic than what could be expected.

4. Estate duty

As discussed above, a socially just and equitable property tax is one, where the tax paid on housing property originating from income from work is deductible from personal income tax. A similar expectation can be formulated in the case of estate duty and the duty on presents resulting in a rearrangement of assets

between generations: the tax exemption of inheritance stemming from income from work (the predominant share of which is housing property in Hungary today) can be regarded as socially just and equitable pointing towards equality of opportunity, while the uniform taxation of large estates manifested in various assets, which stem from incomes other than from work (business quotas, shares, bonds, real estate) can be considered similarly to the value-based taxation of large industrial and commercial property.

Current Hungarian practice, which abolished estate duty[14] in the summer of 2010 disregards the aspects of "generational equitableness"[15] (Örkény), although earlier eras had acknowledged it. Partly for the sake of an accurate definition, and partly to substantiate the statement, let us see a classical quotation: "Estate duty is the taxation of the *economic advantages concomitant with property* acquired by way of inheritance; it should therefore be categorised under the taxation of property, while from another point of view, as a tax collected on the occasion of the transfer of assets, it also has the nature of a turnover tax. Although this kind of tax had already been applied in Antiquity, its significance has increased primarily in our age in so far that many regard it necessary for the mitigation of inequalities in property and income." (Pallas Great Encyclopaedia) (Highlighted by B.A.)

The progressive estate duty or tax exemption up to a relatively substantial value of property/inheritance is a typical instrument of reducing social inequalities. An example of the former is that early in the 1940s "the rate of the estate duty [could reach] 80 percent in England and France", and at that time "the maximum tax rate was 60 percent in Germany, and 50 percent in Italy."[16] (Ungváry, p. 8)

In European countries, where there is an estate duty, it is a mix of legal solutions and tax allowances. Every *Anglo-Saxon* country guarantees exemption based on the magnitude of the assets. The following are international examples of partial or full exemption from estate duty:

- *United States:* the amount of the tax is specified in the annual budget, in 2009 it was applicable to estates not exceeding USD 3.5 million (roughly

---

[14] As of 1 January 2009, an estate was exempted from estate duty up to HUF 20 million/heir. Till then, the rate of the estate duty was generally between 11 and 40 percent, and it moved between 2.5 and 30 percent in the case of housing property. The actual rate depended on how close the kinship was (with the spouse belonging to the first group taxed at the lowest rate), and the amount of the inheritance with a large number of exceptions.

[15] "Whether we are talking about advantages or burdens, they should not be concentrated to a single generation. No generation should enforce its own interest to debit future generations – and the other way round: one must not sacrifice the life opportunities of the current generations and helping its members in a deprived position for the sake of advantages of the future generations. This requires a just sharing of the public burden in society embracing several generations." (Örkény)

[16] A radical increase in estate duty "has for long been urged by folkist writers and progressive middle-class democrats, such as the writer Sándor Márai." (Ungváry, p. 8)

HUF 700 million), irrespective of the number of heirs and the degree of kinship; 97 percent of Americans belong to this category;

- *England:* in 2009, GBP 312,000, the basis of determining the value is the price of an average house in London;
- *Germany:* EUR 307,000 for spouses and EUR 205,000 / capita in the case of children;
- *France:* spouses/life companions do not pay this duty at all, in the case of children it is EUR 156,000 / capita;
- *Italy:* EUR 1 million for every close relative;
- *the Netherlands:* in the case of spouses, in excess of half a million euro, for children it depends on age;
- *Czech Republic:* full exemption in the case of spouses and children. (Dr. Spielmann)

Currently, there is no estate duty in *Sweden, Slovakia* and (since 2008) *Austria, Egypt, Israel and Australia.* (Dr. Spielmann)

The abolition of the estate duty means that the parts of assets generated in the course of privatisation in Hungary exceeding HUF 20 million have been removed from bearing the public burden, that is, this gate to ensuring the equality of opportunity has also been shut. Historically, it would have been the time for exactly the opposite: while substantially raising the HUF 20 million limit for tax exemption, these assets should have been taxed at a more progressive rate than before. The wealthier part of society had it easy to avoid estate duty; its abolition may seem like just and equitable. In Hungary, there are practically no privatisable assets left, and in the future, the redistribution of assets and the network formation of large private companies can be expected. Large Hungarian companies entering the international scene continue to need the parent country because owing to domestic references, they cannot even dream of a global economic role. This would require the exemption of those living off income from work in Hungary, and an improvement in the chances of the next generations – that is, the coming into being of the richest circle of owners, who recognise mutual interdependence and elevate it to the norm.

* * *

A more proportionate and more equitable sharing of the public burden, the guaranteeing of more equal opportunities could be expected even if large private estates had not been accumulated over the past twenty years, if the majority of their owners had not acquired them in the form of presents from society. Similarly, it could be expected for the government to collect the royalties due to it and spend these amounts on financing public goods instead of the hidden

subsidisation of company groups that in any case enjoy monopoly/oligopoly positions. The types of tax and subsidy presented in my study prove that social redistribution will be focused on brushing aside any obstacles from the generation of large incomes and estates in the coming years as against making more intensive attempts at guaranteeing equality of opportunity in society. Yet, we can still hope for the more distant future.

References

CSUTORA, MÁRIA; KEREKES, SÁNDOR (2004): A környezetbarát vállalatirányítás eszközei. [Instruments of environment friendly corporate governance] KJK-KERSZÖV, Budapest

KEREKES, SÁNDOR; KISS, KÁROLY (2004): Környezetpolitikánk európai dimenziói. [The European dimensions of our environment policy] MTA Társadalomkutató Központ, Budapest

KEREKES, SÁNDOR (2007): A környezetgazdaságtan alapjai. [The foundations of environmental economics] Aula, Budapest

KEREKES, SÁNDOR; KINDLER, JÓZSEF (1997): Vállalati környezeti menedzsment. [Corporate environment management] Aula, Budapest

KÜRTHY, GÁBOR (2010): Az egyszerű adórendszer [The simple system of taxation], in Adózó munkaadók és adózó munkavállalók a korrupciómentes gazdaság, Budapest, pp. 49 - 70.

LUKÁCS, ANDRÁS - PAVICS, LÁZÁR - KÁROLY KISS.(2008): Az államháztartás ökoszociális reformja. [The eco-social reform of public finances] Levegő Munkacsoport, Budapest
http://www.google.hu/search?hl=hu&source=hp&q=Luk%C3%A1cs+Andr%C3%A1s%2C+Pavics+L%C3%A1z%C3%A1r%2C+Kiss+K%C3%A1roly.+Az+%C3%A1llamh%C3%A1ztart%C3%A1s+%C3%B6koszoci%C3%A1lis+reformja&meta=&btnG=Google+keres%C3%A9s

ÖRKÉNY, ANTAL (2007): Juss is, marad is (igazságos-e az örökösödési adó?) [Is the estate duty just?] Magyar Narancs, 2007.05.17.
http://www.narancs.hu/index.php?gcPage=/public/hirek/hir.php&id=1474

Pallas Great Encyclopaedia
http://mek.niif.hu/00000/00060/html/078/pc007859.html

PATAKI, GYÖRGY – TAKÁCS-SÁNTA, ANDRÁS SZERK. (2005): Természet és gazdaság, ökológiai közgazdaságtan szöveggyűjtemény [Nature and economy, acompendium of ecological economics], Typotex Kiadó, Budapest

Pavics, Lázár – KISS, KÁROLY (2009): A fosszilis energiák hazai támogatása, [Domestic subsidisation of fossil energy] Budapest
http://www.nfft.hu/dynamic/A_fosszilis_energiak_hazai_tamogatasa.pdf

PÉLI, ÉVA (2010): Offshore helyett nemzeti adózás [National taxation instead of offshore] (in: Adózó munkaadók és adózó munkavállalók a korrupciómentes gazdaság, Budapest, pp. 85 - 136)

SLÁNICZ, MELINDA (2010): Adókikerülés – adócsalás [Tax avoidance – tax fraud]. In: Adózó munkaadók és adózó munkavállalók a korr a korrupciómentes gazdaság, Budapest, pp. 71 – 85.

DR. SPIELMANN, ÉVA (2009): Nem halunk bele! Kedvezően változott az örökösödési illeték, Az üzlettárs [We won't die of it! Favourable change in the estate duty], 2009.02.01.
http://www.azuzlettars.hu/nem_halunk_bele

STIGLITZ, JOSEPH E. (2000): A kormányzati szektor gazdaságtana [Economics of the government sector], KJK – Kerszöv. Jogi és Üzleti Kiadó Kft. Budapest

ISTVÁN M. SZABÓ (2006): A megmaradás elve - Energiapolitika stratégia nélkül [The principle of permanence – Without a strategy of energy policy], Magyar Narancs, 2006. 05.11.
http://www.mancs.hu/index.php?gcPage=/public/hirek/hir.php&id=13106

SZALAI, ÁKOS (2001): Az értékalapú ingatlanadó Magyarországon. [Value-based property tax in Hungary] Magyar Közigazgatás 4., pp. 213-229.
http://www.unicorvinus.hu/index.php?id=24294&tx_efcointranet_pi1[fomenu]=publikaciok&tx_efcointranet_pi1[cusman]=sakos

Ungváry, Krisztián: „Őrségváltás" [Change of the Guard]
http://www.ketezer.hu/menu4/2006_01/ungvary.html

Integrated amending motion No. T/10554/1899. to Bill T/10554 on the 2010 Budget of the Republic of Hungary
http://www.parlament.hu/irom38/10554/10554-1899.pdfh

Main revenues of the central budget Chapter, 2010 budget, 42. pdf (application/pdf objektum
http://www.parlament.hu/irom38/10554/adatok/fejezetek/42.pdf

46

Act C of 1990 on Local Taxes
http://webcache.googleusercontent.com/search?q=cache:zPMftEzMeMgJ:net.jogtar.hu/jr/gen/hjegy_doc.cgi%3Fdocid%3D99000100.TV+1990+%C3%A9vi+C+t%C3%B6rv%C3%A9ny&cd=1&hl=hu&ct=clnk&gl=hu

AN EMPIRICAL STUDY OF THE PRICE DISPERSION OF THE STORES NEAR THE BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS[1]

ÁGNES HARAGH

Why do selling prices differ at various points of sale? Is there a connection between the location of shops and price dispersion? What impact does imperfect information between consumers and companies have on prices? Why does belief have it that the shops close to the Budapest University of Technology and Economics sell at higher prices? Which theories explain this phenomenon? My study aims to answer these and similar questions.

There is a rudimental belief in economics that homogenous products are sold at identical prices on the markets. However this theory of same prices does not cover the reality in several instances. We regularly recognize that a product is sold at different prices in several stores. There may be numerous reasons for price dispersion. I will examine the role of imperfect information and location in pricing.

The review of the theory helps the empirical research with the aim of finding a connection between the prices of the stores near the Budapest University of Technology and Economics (hereinafter referred to as the University) and between their measured distances from the University. In this paper my research will prove that we find expectedly and on average lower than average prices in stores away from the University.

1. Theoretical approach

Hopkins' (2006) simple definition of price dispersion means that different vendors offer the same product in the same market at different prices. This paper assumes that there is no quality difference between products and the stores do

not use the opportunity offered by price discrimination, but the existence of imperfect information and the consumption costs cause the dispersion of prices among the competitive and the reservation price (which is also identified by the monopoly price).

1.2. Location models

During the examination of the dispersion of prices of stores located near the Budapest University of Technology and Economics the variable is the distance which based on the experience, i.e. the prices of the university canteens and stores closer the University, are higher than the further's.

Considering this fact, the dispersion of the prices could be explained by the spatial locations of shops which generate different extent of costs concern to the consumers, which is based on the distance. From the economists who are concerned with spatial location, I would emphasize Hotelling's (1929) and Salop's (1979) study, who were among the firsts in creating a model where consumers are supposed to be uniformly distributed along a line or a circle, therefore the distance has a significant role in these models. In the market the stores define their prices based on the routes between them. They can attract with lower prices the further consumers, who make their purchasing decisions based on their variant costs due to the different distances and prices. From the point of the University, the cost borne by consumers based on distance is their free time, which can be viewed as an opportunity cost.

1.3. Imperfect information models

Salop and Stiglitz (1977) two-price equilibrium theory could explain the higher prices near the University. According to the theory, known also as the tourists and locals model, two groups may be distinguished: the informed group (the locals), who have low search costs, and the uninformed group (the tourists), who have higher search costs. Because the informed consumers, due to their low search costs, recognize all available prices on the market, they can buy at the cheapest place. The other group randomly purchases in stores as long as the price of the product does not exceed their reservation price. Therefore a two-price equilibrium experience forms in the market, where at the lower (competition) price the informed consumers purchase, and a proportion of their share of

uninformed consumers, while at the store with higher, monopoly prices (which is the same as the reservation price of consumers) only the random, uninformed group provides demand.

By this theory two consumer types can be distinguished around the University. On one hand there are the residents of Lágymányos[2] and the college, which could be considered as the group of informed consumers, as their everyday shopping leads them to gather information about prices, and they purchase at the store with the best prices. The students and the teachers constitute the other, uninformed group, because they have less information about the prices and they have limited time (the school breaks) during their stay at the University, and the obtaining of perfect information represents a larger cost to them than purchasing at a higher price. In this case the factor of the distance plays the role of the separation. The uninformed consumers concentrate in a substantial base (which is the University), hence the probability of random purchasing is higher in a buffet or a store near the University. Of course, the higher search costs also play a role in the dispersion of prices, as Stigler (1989) shows that consumers continue the search until the marginal costs of search are equal to the marginal benefit of it. The consumers can search lower-priced stores outwards of the University, but because they find high prices there also, it is not worthwhile for them to further increase their search costs, and they purchase the more expensive product. This also explains why the higher priced stores are located closer to the University.

1.4. Model of sales

Varian (1980) theory takes a special place among models of price dispersion, because it not only analyzes the "spatial" price dispersion, but the "temporal" dispersion is also taken in account. The "temporal" price dispersion means that the prices change over time because of periodic sales. In the Salop-Stiglitz model, the search costs may disappear over time because of the experiences and the learning process, if it is not a single purchase. The consumers choose a store with a lower price after buying in two of them (because they can compare prices of two or more randomly selected stores). The Varian-made theory, however, assumes that there is an uncertainty in the model due to the discounts. For this reason, only for consumers with low search costs is it worthwhile to gather information from prices, and not randomly purchase.

---

[2] This is the part of Budapest where the University is located.

2. Empirical review

Between the University students it is accepted that the buffets are the most expensive shops and if we go further the prices decrease. For example, at the Móricz Zsigmond circus, we can buy much cheaper drink, beer or sandwiches. To prove this I made a questionnaire where the question was about the prices of the stores (are the shops nearby the University more expensive than the others further away?). 72% of the respondents answered positively (13,7% negatively and 14,3% chose "I do not know"), which is almost three quarters of the respondents.
Based on the experience and the theories described yet, the price dispersion is caused by the distance and therefore the existence of imperfect information and the searching costs. I will prove this with an econometric analysis.

2.1. Data collection

I assume that the students and the teachers of the University are the subjects of the examination, so I chose products which are typically attached to consumption during a day spent at the University. Between the prices of these products there can be a significant difference because the students can find them in the buffets and also in the shops. So this examination deals with products which are characteristically consumed by the students: chewing gum, soft drinks, biscuits, chocolates etc. I chose products which appear in the buffets, because their only consumers are the students, so their range of products is tailored for the students also.
After selecting the product basket I have to establish a threshold within which the stores near the University are located. Since the students have only 15 to 30 minutes break between two lessons, the area is determined by the distance which a student can go over this time limit. The study area base is on much of the Lágymányos area (except the part that lies south from the Petőfi Bridge, Iríny józsef street and the Október 23 street), so I chose to examine the area bounded by the Irínyi József street-Karinthy Firgyes street-Bartók Béla road (Figure 1).

Figure 1.: The examined area and the stores nearby the Budapest University of Technology and Economics
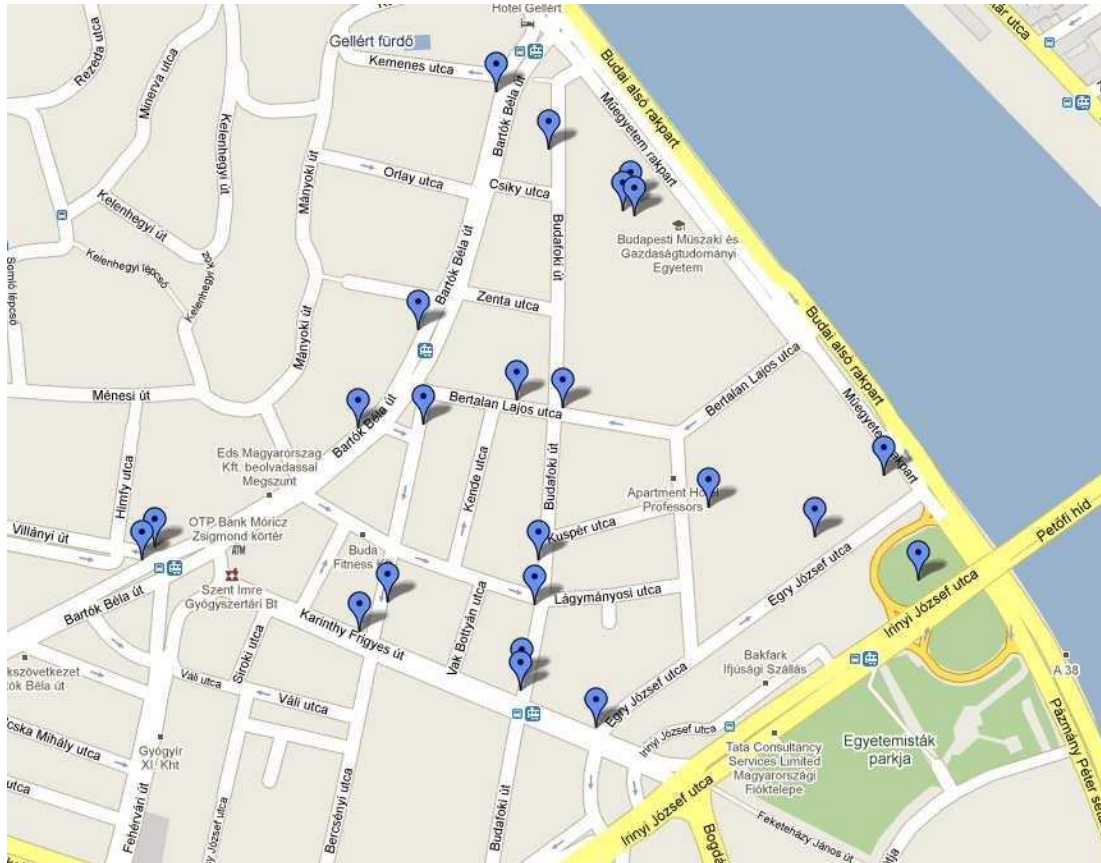
Lágymányos urban area has approximately 19,700 reported inhabitants, while 18,900 pupils study at the Budapest University of Technology and Economics. It is also seems that the consumers studying at the university have a significant "market power".

I went to all of the stores in this area and recorded each of the prices of the listed products. In the area 24 shops are located, but one of them did not allow the collection of data, so 23 stores play a part in my study (Figure 1). These include several university buffets, large stores, small bars and small shops who build directly on the needs of the students.

Unfortunately not all products could be found in every store, so instead of the originally planned 15 products, 10 were included in the analysis, and in one store there was only a small number of the products so it is not taken into account to avoid its distorting effects.

The determination of distances in not an easy task, since the University is large-scale and has no effective center from which the distance of the stores could be measured. I assumed that the canteens are located away from the University with zero distance, which is also not applicable in a case of a point. Therefore I measured the distance with the weighted average of the space between the stores and the exits of the University. I distinguished five exits, which are the following:

- At the building CH, which is an exit toward to the Gellért square
- The buildings K and F2E have a common exit between the two buildings
- The exit at the Bertalan Lajos street
- The exit near the buildings ST and J
- The exit at the Goldmann György square

The distances between the stores and each of the five exits were measured by using Google maps, and further I calculated the variable of the distance from that information. Each of the seven university buffets was given zero as a distance value, since it is located in the area of the University.

In addition to the distance, three other factors are included in the study, which could be relevant in the modeling of prices. These factors are incorporated into the econometric analysis as dummy variables.

For measuring the range of products, the size of the store is used. The product range is significant if the consumers want to buy different products which are not sold by all the stores. For this reason, a larger shop deals multiple products, while much less can be maintained in a buffet.

Prices are strongly influenced by late night opening, since it is known that the non-stop shops, which are also open at nighttime, have higher prices because of their longer opening hours and the additional costs caused by the night shift. Thus, this type of shop is usually expensive.

The third factor if the store is part of a retail chain (for example Coop or Reál). I think this is important for several reasons. If it is part of a retail chain, the chain determines the prices of the shops, or even if the franchise does not specifically determine the exact selling price, they can limit prices into intervals. A chain store with an advantageous location does not necessarily have the opportunity to raise prices. Another aspect of whether the store is part of a chain is that it will have a positive impact on some consumers purchasing decisions. For example, there are products that are found only in certain chains, and also specify the type of the store (cheap, but mostly Hungarian product or very high quality, etc.) with the marketing positioning.

I took these factors into consideration and used them to model the problem.

3. Econometric analysis

My object was to establish what the correlation is between the distance of the stores from the University and their prices. Therefore the initial hypothesis is that there is a negative correlation among the distance of shops (in meters) and their prices (HUF). I used the Gretl program package for the econometric analysis.
The symbols used in this study:

- APrice: average price
- ADist: average distance
- WDist2: weighted distance, where the rate of the weight is double
- WDist10: weighted distance, where the rate of the weight is tenfold
- WDistOwn: it is an own weighted distance, where the weights have different probabilities
- Dx_1: dummy variable for the size of the store
- Dx_2: dummy variable for the late night opening
- Dx_3: dummy variable for taking a part of a chain
- α: constant term
- β_i: theoretical parameters
- u: error term, also known as the error term
- R^2: coefficient of determination
- $\overline{R}^2$: adjusted coefficient of determination
- F: the value of the F-test
- Sw: value of the Schwarz's criterion of model selection
- HQ: value of the Hannan-Quin's criterion of model selection

I examined how the distance affects the prices. For this, however, I had to work with some kind of weighted average of the prices. Since the price level is constituted with a weighted average, I composed the average price variable (APrice) similarly. Since the goods consumed by students, buy them near to the University, it is justified to all the examined products include in the variable with the same weight. So I arrived at the average price dependent variable, which covers the average price of the different goods by stores.
It was more difficult with the distances, because the value of the distance is given, but there are tracks which the students use often. This can mean that from a given exit the students can reach several interchanges or colleges that could be the target of the pupils of the University. In many cases, if someone wants to go to building CH from the Goldmann György square, they do not have to leave the

area of the University. Similarly, it is hard to assume that a student would purchase in a store 900 meters away from the University, while there are lot of stores en route. Of course, for the informed consumers it is a perfectly logical step because of the purpose of buying at the cheapest price, but for the majority of the examined consumers, it is not.

The average distance is already included in the location of the stores, but with equal weight, even if it is 250 meters or 900 meters, although the shops near the University have more uninformed consumers. So, except the average distance (ADist) variable, I created even more weighted variables, where the stores located closer than 500 meters was given greater weight: on the one hand WDist2, where the nearest shops had a double weight, on the second hand WDist10, where ten-fold weight was assigned, and finally WDistOwn, where I assigned probabilities for each shop depending on the chance that a student who left the University would arrive at that shop.

In addition to these variables, the already mentioned size of the store nominated with Dx_1, the late night opening with Dx_2, and part of a retail chain with Dx_3.

In the case of the size of the store, the control group means the small shop (ie the dummy variable takes the value 0, and if the store is big, takes the value 1). The shops that are not open at night constitute the control group of the late night opening Dx_2 variable, and the ones that do not belong to a chain are the control group of the Dx_3.

At first a simple correlation between the distance variables and the average price is examined. Table 1 shows the results of this:

Table 1: The correlation between the average price and the different weighted distance variables (p-value in parentheses)

Matrix of the correlations

	APrice
ADist	-0,6659 (0,0007)
WDist2	-0,6725 (0,0006)
WDist10	-0,6813 (0,0005)
WDistOwn	-0,7061 (0,0002)

The negative correlation indicates that between the average price and each of the distance variables a reverse relationship exists. Say if the distance increases, the prices will fall and vice versa. In addition, it seems that the strongest connection is between the WDistOwn weighted average of the distance and the average price.

First, I present a simple model in which the dispersion of prices is only explained with the difference between the distances of the stores measured from the University. Then the APrice, average price dependent variable, is only explained by a constant (α) and by the distance. I examined the connection between the average price and each of the different weighted variables of the distance.

To observe the relationship of the average price and the average distance I used the following equation:

$$A\Pr ice = \alpha + \beta ADist + u \qquad (1)$$

I examined the model with the method of the Ordinaly Least Squares (OLS) where the regression coefficients are only meaningful if the change of a variable is observed when we assumed that the other members remain unchanged.

The OLS estimation result (Table 2) is that the dispersion of the average price attributes to the average distance with R^2=44,35%, where R^2 is the coefficient of determination, measures the goodness of fit.

Table 2.: The result of the OLS estimation, which shows the connection of the average price and the average distance (p-value in parentheses)

Model 1

constans	172,171 (0,0000)
ADist	-0,0279 (0,0007)
R^2	0,4435
$\overline{R}^2$	0,4157
F	15,9387
Sw	168,4575
HQ	166,7895

On the basis of the Table 2, the relationship between the average price (APrice) and the average distance (ADist) can be written as:

$$A\hat{\Pr}ice = 172{,}171 - 0{,}0279\,ADist\,. \tag{2}$$

This means that the increase of the distance variable by 100 meters is expected to cause an average HUF 2,79 decrease in the average price according to this model.

The result is the same by examining the other weighted distance variables, which is shown by Table 3 and Table 4.

Table 3.: The average price dependent variable is explained by the WDist2 variable in Model 2 and by WDist10 variable in Model 3 (p-value in parentheses)

<table>
<tr><td colspan="2" align="center">**Model 2**</td><td colspan="2" align="center">**Model 3**</td></tr>
<tr><td>constans</td><td>171,867
(0,0000)</td><td>constans</td><td>171,211
(0,0000)</td></tr>
<tr><td>WDist2</td><td>-0,0292
(0,0006)</td><td>WDist10</td><td>-0,0309
(0,0005)</td></tr>
<tr><td>R^2</td><td>0,4523</td><td>R^2</td><td>0,4642</td></tr>
<tr><td>$\overline{R}^2$</td><td>0,4249</td><td>$\overline{R}^2$</td><td>0,4375</td></tr>
<tr><td>F</td><td>16,5177</td><td>F</td><td>17,33095</td></tr>
<tr><td>Sw</td><td>168,1059</td><td>Sw</td><td>167,6213</td></tr>
<tr><td>HQ</td><td>166,4379</td><td>HQ</td><td>165,9533</td></tr>
</table>

According to Model 2, the dispersion of the average price variable is explained by the WDist2 weighted variable with R^2=45,23%. The increase of the distance by 100 meters is expected to cause an average HUF 2,92 decrease in the average price. In the case of the WDist10 variable, whose connection with the average price is described in Model 3, R^2=46,42%. This means that the distance variable explains the dispersion of prices by 46,42%, and the decrease of the distance by 100 meters is expected to cause an average HUF 3,09 increase in the average price.
With regards to the WDistOwn variable, Model 4 shows the results in Table 4.

Table 4.: The results of the OLS estimation of the average price and the WDistOwn variable (p-value in parentheses)

<table>
<tr><td colspan="2" align="center">**Model 4**</td></tr>
<tr><td>constans</td><td>171,742
(0,0000)</td></tr>
<tr><td>WDistOwn</td><td>-0,0314
(0,0002)</td></tr>
<tr><td>R^2</td><td>0,4986</td></tr>
<tr><td>$\overline{R}^2$</td><td>0,4735</td></tr>
<tr><td>F</td><td>19,89004</td></tr>
<tr><td>Sw</td><td>166,1627</td></tr>
<tr><td>HQ</td><td>164,4946</td></tr>
</table>

The WDistOwn weighted distance variable explains the distribution of the prices with R^2 =49,86%, and the growth of the distance by 100 meters is expected to cause an average HUF 3,14 decrease in the average price in Model 4.

Notice that after the weighting, the explanatory power of the model increases.

Then for further examination, the dummy variables can be taken into account. Examining the correlation between the dummy variables and the prices I experienced the following: the size of the stores and part of retail chain negatively correlate to the average price (the correlation is -50,6% with the previous, and -25,9% with the latter), while the late night opening shows a positive connection (35,6%). Table 5 includes these results.

Table 5.: The correlation between the dummy variables and the average price (p-value in parentheses)

Matrix of the correlation

	APrice	Dx_1	Dx_2	Dx_3
APrice	-	-	-	-
Dx_1	-0,5062 (0,0162)	-	-	-
Dx_2	0,3556 (0,1043)	0,2082 (0,3526)	-	-
Dx_3	-0,2598 (0,2430)	0,4623 (0,0303)	-0,1324 (0,5568)	-

This means that if the shops are larger or part of a retail chain, the offer price should be lower. However the stores with a late night opening should be set at higher prices.

The multicollinearity between the dummy variable can be excluded by the Table 5, since the value of each pair of correlation coefficients does not explain the existence of multicollinearity.

In the regression model the dependent variable is the average price (APrice), while the independent variables are the distances (ADist or WDist2 or WDist10 or WDistOwn), and the dummy variables as the dispersion of prices are explained by the size of the shop, the late night opening, part of a retail chain, and the distance of the store by the University. I assumed that there is a linear relationship between the dependent variable and the independent variables. The relationship between them can be written as the following equation:

$$A\Pr ice = \alpha + \beta_1 Dist + \beta_2 Dx\_1 + \beta_3 Dx\_2 + \beta_4 Dx\_3 + u \,. \qquad (3)$$

The Table 6 includes the results of the OLS estimation.

Table 6.: The results of the OLS estimation for the APrice dependent variable where the independent variable for the distance is the ADist variable (p-value in parentheses)

Model 5

	model A	model B	C
constans	172,321 (0,0000)	171,613 (0,0000)	171,586 (0,0000)
ADist	-0,0287 (0,0005)	-0,0279 (0,0003)	-0,0333 (0,0000)
Dx_1	-4,6053 (0,3548)	-5,6931 (0,1655)	-
Dx_2	20,5722 (0,0005)	21,0337 (0,0002)	20,4757 (0,0003)
Dx_3	-1,5475 (0,6897)	-	-
$\underline{R^2}$	0,7525	0,7501	0,7211
$\overline{R}^2$	0,6942	0,7084	0,6917
F	12,9201	18,0071	24,5574
Sw	159,9069	157,0284	156,3537
HQ	155,7368	153,6923	153,8516

Between the average price (APrice) variable and the average distance (ADist) variable, the size of the store (Dx_1) variable, the late night opening (Dx_2) variable and the part of retail a chain (Dx_2) variable the correlation is $R^2 = 75,25\%$ (i.e. in 75,25% they move together). However only the constant term (α), the average distance (ADist) and the late night opening (Dx_2) dummy variable have less then 10% p-value (which is the p-value of the null hypothesis' two-tailed test result which analyzes the assumption that the examined coefficient is zero (Ramanathan, 2003)), and even they are relevant at 1% significance level. If the variable with the highest p-value (which is the taking a part of a chain dummy variable, Dx_3) is omitted, both the Schwarz and the Hanann-Quinn model selection criteria's value decrease, i.e. the model improved. The increase in the value of the adjusted determination coefficient shows that the decrease of the determination coefficient (R^2) and the increase in the degrees of freedom lead to an improvement in the fit of the model together.

The Model 5 shows that the model B is expectedly and on average a better model to describe the average price than the model A.

Since the dummy variable of the size of the store (Dx_1) is not significant either at the 10% level in the model B, it is omitted from the model C to examine the changes. However, the model C (Table 6) shows that this omission does not necessarily leave the estimate undistorted. In between the model selection criteria the value of the Hannan-Quinn and the value of the adjusted determination coefficient is also deteriorated. Thereby, to avoid the distortion, the Dx_1 variable should remain in the estimation, and model B is relevant. This can be written as:

$$A\hat{\Pr}ice = 171{,}613 - 0{,}0279ADist - 5{,}6931Dx\_1 + 21{,}0337Dx\_2 \qquad (4)$$

The explanatory power of the model is $R^2 = 75\%$, which represents a considerable result of a cross-sectional analyzes. The value of the coefficients shows the effects of the changes of the variables to the average price, while the other variables are not modified. Thus, the decrease of the distance by 100 meters is expected to cause an average HUF 2.798 increase in the average price. If the dummy variable of the size of the store takes 1 for value (which means a big size), the average price of the shop expectedly and on average lower by HUF 5.69 than the average price of a smaller store. When a shop decides to open during late night hours, this is expected to cause an average HUF 21.03 increase in the average price of the store.

The OLS estimation is shown in Table 7, where, instead of the APrice, average variable, the WDistOwn weighted distance variable explains the dispersion of prices.

Table 7.: The results of the OLS estimation for the APrice dependent variable where the independent variable for the distance is the WDistOwn weighted variable (p-value in parentheses)

Model 6

	model A	model B	model C
constans	170,382 (0,0000)	170,270 (0,0000)	170,435 (0,0000)
WDistOwn	-0,0277 (0,0014)	-0,0278 (0,0008)	-0,0337 (0,0000)
Dx_1	-5,9958 (0,2474)	-5,9105 (0,1748)	-
Dx_2	17,9996 (0,0026)	17,9534 (0,0015)	16,7773 (0,0026)
Dx_3	0,1314 (0,9737)	-	-
R^2	0,7238	0,7238	0,6932
$\overline{R}^2$	0,6588	0,6778	0,6609
F	11,138	15,7228	21,4618
Sw	162,3177	159,2281	158,4502
HQ	158,1476	155,892	155,9481

Similarly to Model 5, in case of Model 6 the model B gives the expectedly and on average highest efficiency estimate (5).

$$A\hat{\Pr}ice = 170{,}270 - 0{,}0278 WDistOwn - 5{,}9105 Dx\_1 + 17{,}9534 Dx\_2 \qquad (5)$$

Using the WDistOwn weighted distance variable instead of the ADist average distance variable, the model explained the dispersion of price with a lesser extent (R^2 =72.38%), which still represents a great explanatory power. The increase of the WDistOwn weighted variable by 100 meter is expected to cause an average HUF 2.78 decrease in the average price. In case of a larger shop size, the stores offer lower prices expectedly and on average by HUF 5.91, while the late night opening attracts higher average prices by expectedly and on average HUF 17.95. After look at the Table 8, where the distance variable is replaced with the WDist2 weighted distance variable.

Table 8.: The results of the OLS estimation for the APrice dependent variable where the independent variable for the distance is the WDist2 weighted variable (p-value in parentheses)

Model 7

	model A	model B	model C
constans	171,367 (0,0000)	171,067 (0,0000)	170,978 (0,0000)
WDist2	-0,0284 (0,0009)	-0,0281 (0,0005)	-0,0338 (0,0000)
Dx_1	-5,3601 (0,2919)	-5,8306 (0,1679)	-
Dx_2	19,9641 (0,0009)	20,1850 (0,0004)	19,4661 (0,0007)
Dx_3	-0,6952 (0,8598)	-	-
R^2	0,7373	0,7369	0,7067
$\overline{R}^2$	0,6755	0,6929	0,6758
F	11,9314	16,8012	22,8866
Sw	161,2117	158,1622	157,4604
HQ	157,0416	154,8261	154,9583

In Model 7, the same can be seen as in the case of Model 5 and 6, so here also model B gives supposedly the closest modeling of reality (6).

$$A\hat{P}rice = 171,067 - 0,0281WDist2 - 5,8306Dx\_1 + 20,1850Dx\_2 \qquad (6)$$

The WDist2 weighted distance variable, the Dx_1 and the Dx_2 dummy variable explains the dispersion of price with R^2 =73.69%. The decrease of the WDist2 weighted variable by 100 meters is expected to cause an average HUF 2.8 increase in average price. The larger store size expectedly and on average cause a HUF 2.8 diminish in the average price of the shop and the late night opening stores determine higher prices by expectedly and on average HUF 20.18.

Finally the model with WDist10 weighted distance variable is presented in Table 9.

Table 9.: The results of the OLS estimation for the APrice dependent variable where the independent variable for the distance is the WDist10 weighted variable (p-value in parentheses)

Model 8

	model A	model B	model C
constans	169,851 (0,0000)	170,062 (0,0000)	169,903 (0,0000)
WDist10	-0,0278 (0,0022)	-0,028 (0,0013)	-0,0341 (0,0000)
Dx_1	-6,1746 (0,2475)	-5,8351 (0,1963)	-
Dx_2	18,9269 (0,0021)	18,7452 (0,0013)	17,7849 (0,0020)
Dx_3	0,5312 (0,8963)	-	-
R^2	0,7101	0,7097	0,6807
$\overline{R}^2$	0,6418	0,6614	0,6471
F	10,4076	14,6719	20,2537
Sw	163,3875	160,3191	159,3258
HQ	159,2174	156,983	156,8238

Model 8 differs from the above estimates in that the Dx_1 dummy variable (which depends on the size of the store) is omitted, because of the improvement of both of the two model selection criteria (the Schwarz and the Hannan-Quin). Despite the deterioration of the adjusted determination coefficient, in the case of Model 8, model C fits better expectedly and on average to the average price (7), thanks to the improvement of the two model selection criteria.

$$A\hat{\mathrm{Pr}}ice = 169,903 - 0,0341 WDist10 + 17,7849 Dx\_2 \qquad (7)$$

The WDist10 weighted distance variable and the Dx_2, late night opening dummy variable explains with $R^2 = 68.04\%$ the APrice average price dependent variable, which value is lower than the explanatory power of Model 5-7, because in this case there are less independent variables which constitute the model.

The increase of the WDist10 weight distance variable by 100 meters is expected to cause an average HUF 3.41 decrease in the average price. The late night opening means an average price rise by expectedly and on average HUF 17.78.

Regarding to the OLS estimations, the Table 10 shows the four types of distance variables, and the differences between the estimated models using them.

Table 10.: The results of the OLS estimation in case of the different distance variables (p-value in parentheses).

		The comprehension of the models				
		Model 5 (ADist)	Model 6 (WDistOwn)	Model 7 (WDist2)	Model 8 (WDist10)	
	constans	171,613 (0,0000)	170,270 (0,0000)	171,067 (0,0000)	169,903 (0,0000)	
	Coefficient of the distance	-0,0279 (0,0003)	-0,02777 (0,0008)	-0,0281 (0,0005)	-0,0341 (0,0000)	
	Dx_1	-5,6931 (0,1655)	-5,9105 (0,1748)	-5,8306 (0,1679)	-	
	Dx_2	21,0337 (0,0002)	17,9534 (0,0015)	20,1850 (0,0004)	17,7849 (0,0020)	
	R^2	0,7501	0,7238	0,7369	0,6807	
	$\overline{R}^2$	0,7084	0,6778	0,6929	0,6471	

Each of the models illustrates that the distance is a significant independent variable of the price dispersion. At three distance variable from four, the size of the store is considered to be a relevant variable, and only in the case of the WDist10 weighted distance variable left out of the model. Increasing the separate distance variables by 100 meters is expected to cause a decrease in the average price by HUF 2.8-3.4, which means that the distance variable is a robust independent variable.

4. Conclusion

The relationship between the theoretical models is that the costs borne by consumers in the location models could be the search costs due to the imperfect information. Both models contain the distance, such as the inducer of the search costs, of the costs borne by consumers and as well as of the price dispersion. Hence the empirical research is based on the relationship of the prices and the distance.

The econometric model consists of four independent variables. The average price dependent variable is explained by the size of the stores, the late night opening, the part of a retail chain, and in the four models, four types of the weighted

distance variable. After completion of the model selection of the four weighted distance models, three of them contain the expectedly and on average best-fit model in addition to the distance. Also, the dummy variables measure the size of the store and the late night opening. The increasing of the distance variable by 100 meters is expected to cause an average HUF 2.8 increase in the average price, in all three models, using different weights. In the fourth case only the weighted distance variable and the late night opening dummy variable remain the part of the final model, so here the effect of the distance is clearly different. However, even this difference is not significant, since the augment of the distance variable by 100 meters is expected to cause an average HUF 3.4 decline in the average price.

The empirical analysis suggests the existence of a negative correlation between the distance and the prices, in case of the stores nearby the Budapest University of Technology and Economics, and from the point of view of the price, the distance is a robust independent variable, so the study is proved the thesis outlined in the economic theories on the findings of the price dispersion.

66

References

CARLTON, DENNIS W. AND PERLOFF, JEFFREY M. (2003A): *Modern Piacelmélet* c. tankönyv Elhelyezkedési modellek c. fejezet 247-258. oldal, Panem kiadó 2003, Budapest.

CARLTON, DENNIS W. AND PERLOFF, JEFFREY M. (2003B): *Modern Piacelmélet* c. tankönyv Korlátozott információ az árról c. fejezet 470-490. oldal, Panem kiadó 2003, Budapest.

HOPKINS, ED (2006): Price Dispersion. (http://www.homepages.ed.ac.uk/hopkinse/), publications. Date of the download: 23 April, 2010.

 http://www.homepages.ed.ac.uk/hopkinse/price-d.pdf .

HOTELLING, HAROLD (1929): Stability in Competition. *The Economic Journal*, Vol. 39. pp. 41-57., 1929.

RAMANATHAN, RAMU (2003): *Bevezetés az ökonometriába alkalmazásokkal.* Panem kiadó, 2003, Budapest.

SALOP, STEVEN (1979): Monopolistic Competition with Outside Goods. *The Bell Journal of Economics*, Vol. 10. pp. 141-156., 1979.

SALOP, STEVEN AND STIGLITZ, JOSEPH (1977): Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion. *Review of Economic Studies*, Vol. 44. pp. 493-510., 1977.

STIGLER, GEORGE J. (1989): *Az Információszerzés Közgazdaságtana, Piac és állami szabályozás* c. kiadvány, 232-253.oldal, Közgazdasági és Jogi Könyvkiadó, 1989, Budapest.

VARIAN, HAL R. (1980): A Model of Sales. *American Economic Review*, Vol. 70. pp. 651-659., 1980.

REFORMING ADMINISTRATIVE SYSTEMS: ON THE ROAD TOWARDS A SUSTAINABLE EUROPE[1]

GYÖRGY ÁDÁM HORVÁTH

1. Introduction

Almost 54 years have passed since the foundation of the European Union, and in its history spanning over half-a-century it has changed from a free-trade association to a significant global power, and a potential future state. In contrast to its original six members, currently 27 states pertain to this Union. It can be expected that the number of states will increase once again in the near future.

In varietate concordia, or "United in Diversity", is the motto that was chosen to best represent the values of the European Union. Rightfully so, as successive expansions have made sure that there is no shortage in diversity. The Union is therefore continuously wedged between a pair of challenges: unification on the one side, and the preservation of diversity on the other. Europe derives its global potential from diversity – although paradoxically this is also the greatest threat to its efficient functioning. Efficient, in this case not referring to administrative efficiency, being *speedy* and *sufficient*, nor to economic efficiency, being a simple synonym of *cheap*, but rather efficient in serving sustainable development. One does not need to look further than the single largest democratically elected institution on the continent to realise the shortfalls.

2. Cumbersome institutions – (un)necessarily so?

The *European Parliament*, with its 736 members elected by nearly 500 million voters remains on a constant shuttling mission between its meeting places in Brussels and Strasbourg. This system makes the whole affair an unnecessarily costly one – in fact, moving the whole parliamentary apparatus back and forth ten times each year on the 856 kilometre return journey costs European taxpayers some 200 million euro annually. But besides all this, the environmental impact is not really accounted for, and for a good reason too: it's not possible. Such practices rightfully upset citizens when it comes to spending public money, and it's not surprising that more and more Europeans want to change this practice. Well over a million, in fact, have already signed a petition demanding Brussels to be the single seat of the European Parliament. This would be the most sensible solution of this particular problem, as the key European Union bodies are mostly located in Brussels, having purchased and rented millions of square metres of office space.

This issue shows us several important aspects of the European institutions, but of citizens also. Firstly, as the European Union remains a very diverse entity, its institutions must be capable to cope with this diversity, making them very big and complicated anyway. Additionally to this, some old compromises have remained valid even until today, which – at the time of their agreement – were a necessary part of a political game of tit-for-tat, ultimately establishing stability. However, as time has passed, so has the necessity for some of these burdensome constructs. More importantly, a simple, efficient and clear solution has been identified, but unfortunately, could not yet be agreed upon by all parties – member states, in this case, – because of fears that this move would be disadvantageous[2]. On the other hand, this particular situation shows something

---

[2] In fact, in this particular case, the only real loser of the solution would be the City of Strasbourg. During the periods of session, the local hospitality industry experiences a surge in demand, so prices must rise, as expansion in supply capacities lags behind to meet the demand. However, as demand remains nearly completely price inelastic, accommodation capacity must be increased. This, however, means that for much of the year – save ten short weeks – the hospitality industry is in a continuous downturn, which must be compensated by "getting the money where you can", i.e. from the *eurocrats*, who are willing to pay any amount – from taxpayers' money. Additionally, only a fraction of the costs totalling 200 million euro are actually spent *in Strasbourg*. Most of it is transportation cost of equipment and people, and true local spending is minimal. The other regrettable fact is that as the European Union has grown over the years, so has the demand for office space in the Strasbourg European Parliament building. The building may look unfinished from the outside, but from the inside it is very much complete. Recent experiences show that it is becoming so cramped, that it now truly hinders efficient working. If

completely different too. In Article 11 of the Treaty of Lisbon, the European Union has opened its gates for citizen participation by empowering EU citizens to petition the European Commission if they get one million signatures from across Europe for a given cause, putting it another step closer to a Citizens' Europe. The One Seat petition has also proved that citizens' initiatives can be functional, and there are causes for which one million European Citizens will come together. Regrettably, this very first "test-run" of citizens' initiatives has also proved that some rationalisation issues may be obstructed by political game-play and so-called national interests.

As of today, this problem has not been resolved. This also means that citizens are unnecessarily given an opportunity to get disappointed, whilst their attention has been drawn to the obvious waste of public funds. Subsequently, they are more likely to lose faith in their elected representatives, which, paradoxically, will be a crucial negative factor in aiming for a sustainable Europe.

3. Sustainable Development in European policymaking

During the last few decades, the concept of sustainable development has entered and has become a noteworthy feature in European policies. The inclusion of this concept in European policies has culminated in its inclusion in the European Union's successive Sustainable Development Strategies. This initiative is remarkable and unique in global terms too, because it establishes a single binding framework for an entity of over half-a-billion people. The Union's Sustainable Development Strategy relies heavily on the individual member states, and it requires each member state to determine, implement, review and improve its own strategy. Each member state, just as the European Union itself, is required to produce a biannual report on the progress achieved, and must produce updated versions of the Sustainable Development Strategy regularly.

Besides this, the Sustainable Development Strategy proscribes that the concept of sustainable development be included not only into the Strategy itself, but into *all* policies of the European Union – thus including those of its member states as well. As the Strategy puts it, the Principle of Policy Integration is to "[p]romote integration of economic, social and environmental considerations so that they are

---

capacity was to be further expanded, the problem would not get better, but worse, because in fact the European Union would be maintaining *an even larger building* that would still remain *not used* for 307 days a year.

coherent and mutually reinforce each other by making full use of instruments for better regulation, such as balanced impact assessment and stakeholder consultations." [EU-SDS 2006, p.5] This principle backs up a previous principle, that of Policy Coherence and Governance, which ought to "[p]romote coherence between all European Union policies and coherence between local, regional, national and global actions in order to enhance their contribution to sustainable development." [EU-SDS 2006, p.5] This integrative approach attempts to prevent a merely superficial inclusion of the concept of sustainable development – i.e. in the form of a mention thereof in a preamble, or something equally unimportant – but rather it establishes that it is the duty of lawmakers to make sure that any policy, which may be seemingly unrelated to sustainable development, if it cannot advance it, will at least not hinder it.[3]

However, if we reflect on the three dimensions of sustainable development – namely environment, society and the economy – we have to recognise that "seemingly unrelated" of the previous paragraph will become rather difficult to interpret. Indeed, as it has been said on countless previous occasions, sustainable development is a lot more than just environmental protection, so policy integration is truly a key element in attempting a successful implementation of measures.

The Sustainable Development Strategy sets the framework not only for member states, but for sub-national – local and regional – entities as well, so, ideally, they should be capable of operating in functional symbiosis with each other, enhancing the positive factors and cancelling out the negative ones. This reinforces the need that hierarchically structured entities have to cooperate with one another both horizontally and vertically, doing so in such a way that they do not contradict one another. To prevent such contradictions, a strict hierarchy exists in legal systems – i.e. an entity on a lower level may not contradict a regulation passed on a higher level – but on a lateral level such clear preventative measures seldom exist, which leave the gates gaping open for contradicting and potentially mutually disadvantageous effects.

However, diversity is not only an inherent feature of the European Union between its peoples and languages and cultures. The European Union is home to twenty-seven member states, all with their own particular administrative structures and practices. Even though all member states' legal systems fall into one of three types of legal systems, in effect, there are at least 27 distinct legal and administrative systems in existence side by side.

---

[3] This certainly puts the two-seat policy of the European Parliament in a peculiar situation.

4. European integration: a concerted dissonance?

It may be regarded as an unfortunate inheritance of the European Union that all member states could retain their original administrative structures without having to do any significant coordination. The reasons for leaving everything as they were are both historical and political, and they were even understandable at the time, as the odds were stacked high against a successful European integration. Of course, with every expansion of the Union, a different level of harmonisation had to be done on the side of the acceding member states, but altogether, on Union level not much was done. However, the repercussions of this missed opportunity of streamlining the European administrative system can be hard felt today, and there is mounting pressure on lawmakers to reform the basis of the European structure quickly to prevent the additional structures from coming crumbling down. The need to do so is echoed by nothing else more painfully, than the double rejection of the – admittedly ambitious – European Constitution some years ago, which has proved that any member state, no matter how big or small, has the power to prevent all other member states from achieving what would be otherwise mutually beneficial, for reasons completely unrelated to all other parties. This is truly the Achilles Heel of the European Union, and arguably one of the most pressing points one has to work around. The Treaty of Lisbon was the first working attempt on Union-level to make the European Union a more efficiently working body, but in comparison to the European Constitution, it is little more than a lame shadow. Now it is the individual member states' task to carry out an administrative reform in conformity with 21st Century European reality.

5. Administrative systems in Europe: a review

Although all European administrative systems are different from one another, naturally a number of similarities exist, for historical and political reasons, but also because states were required to harmonise their legal systems to conform to European standards. Nevertheless, states can be classified into seven categories according to their administrative structures and practices. These are, in no particular order:

- **Scandinavia** (Denmark, Finland, Sweden)
- **The Baltics** (Estonia, Latvia, Lithuania)
- **The Benelux** (Belgium, The Netherlands, Luxemburg)
- **The British Isles** (United Kingdom, Ireland)
- **Western Europe** (Austria, France, Germany)
- **The Mediterranean** (Cyprus, Italy, Greece, Malta, Portugal, Spain)
- **Central Europe** and **The Balkans** (Bulgaria, Czech Republic, Hungary, Romania, Slovakia, Slovenia)

As it is apparent from this classification, many of these states have a shared history, which is reflected in their administration as well. However, over the course of the last few decades, a number of them have taken – frequently bold – steps in reforming their internal structures, with some remarkable outcomes. These reforms were not always based on rationalisation – in fact, few have had true economic reasons behind them, and a number of them were used as means of declaring political independence, especially in the former sphere of influence of the Soviet Union. Only in Scandinavia, the Benelux and Western Europe do we find true examples of administrative reforms for economic reasons. In nearly all other cases, it is a very spectacular political message – which is particularly proved by the desire of some of these states to *revert back* to their previous, better-working, traditional structures, rather than to remain in a hastily planned construct that is both costly and unworkable over a long period of time. A handful of states seemed to follow a different path, with only carrying out symbolic reforms, essentially leaving outdated structures in practice, and disregarding those newer and already proven to be efficient models of public administration. Presently, some of these states are putting the emphasis in relying more heavily on older structures, reverting even further back in administrative history, which may be disadvantageous for their integration in the European Union.

A positive sign, though, is that a state previously most reluctant to carry out true reforms has done something that was both smart and surprising. Greece, the long-time EU-member without proper regional administrations has not only re-affirmed its regions by shifting the responsibilities from its prefectures to its regions, it has also organised the first direct regional elections, briskly shifting Greece from having one of the most outdated to having one of the most modern and EU-conform administrative systems.

6. Contemporary trends in administration and reform

If we take a look at the internal administrative divisions of the member states of the Union, we will see diversity, modernity, but at the same time, some rather outdated structures as well. Of course, it is understandable why the European Union was reluctant to intervene directly with the domestic administration of its member states – it also had no mandate to do so, as that would have breached the sovereignty of its members. However, the European Union *has* created a set of guidelines, which, although are merely for statistical purposes, they also form the backbone of Europe's regional policy, and they provide a very decent model for realigning internal administration. We are of course talking about *NUTS*, the Nomenclature of Units for Territorial Statistics. Put in effect by Regulation 1059 of 2003, this system was established for the allocation of regional development funds when Ireland joined the European Communities in the 1970s. This system has been in existence since that, but it was actually in practice because of a Handshake Agreement. It wasn't until 2003 that the scheme was ratified by all member states.

The NUTS-classification is based on existing administrative structures, and the legislation does not provide for the realignment of administrative units for the sake of NUTS comparability. However, should states choose to reform their internal administration, *they remain free to do so*. As an outline, units should be comparable in size in terms of population. Furthermore, in all cases, the statistical divisions are to be there to respect the "actual political, administrative and institutional situation." Additionally, "[n]on-administrative units must reflect economic, social, historical, cultural, geographical or environmental circumstances". [Official Journal, L-154/2003, p1]

A number of EU member states have recognised the potential in adapting to this scheme. Their incentive was simple: if they were going to get financial support under this system, why not make it work well? Two paths lay ahead: taking previously existing internal structures and virtually joining them together as regions (as in the case of Hungary with its Planning and Statistical Regions), or by doing the complete opposite and scrapping all previously existing structures and re-aligning them again (as in some cases in the Baltic, for example). Of course, the reasons for taking either path were seldom socio-economic only; most often than not, they were coupled with a substantial political load as well.

7. A shattered Europe

Let us now review the current state of administrative divisions across the Union. Taking NUTS classifications as a basis, these are the major figures describing the European administrative divisions.

- Geographic Area   4 324 782 km^2
- Population   501 064 211
- Member States   27
- Regions (NUTS 2)   271
- (NUTS 3)   1 303
- First Level Local Administrative Unit   8 732
- Second Level Local Administrative Unit   121 601

At first glance, these numbers may not seem startling, and we may not recognise much more than the occurrence of a near tenfold difference between each successive level. Even if we contrast the mean area per First or Second Level Local Administrative Unit, the numbers seem to be around an expected level, if such a thing exists at all. But even if we look at the absolute numbers at, say, Second Level Local Administrative Units, and realise that nearly a *third* of all units are *in France*. Indeed, it is the distribution of these units that present a threat when it comes to agility of responses. Any argument defending such an absurdly high number of units by area or by population must be false, since Germany has a larger population than France, yet far *fewer* administrative units. Being the largest member state by area, it is hard to tackle a hypothetical argument that it is the *area* why there are so many units, but even in this case, if we take the second or third largest state, Spain and Sweden respectively, we encounter significantly *fewer* units. In fact, the difference between Sweden and France is over a hundredfold.

Before one settles into the illusion that this problem affects France only, one must realise that it holds for the majority of states. So why, then, is Europe so fractured, so shattered administratively? The reasons are of course historical, but from some practical aspects, they have to do with the dispersion of information. Especially in the case of France, amidst the French Revolution of the 18th Century, previous administrative units were abolished, and new ones – fitting the revolutionary ideals – were created. The reforms did show some reduction in numbers of units, but at the time this was rather irrelevant. The key aspect of keeping a country together was the ability to relay information quickly. Indeed,

the principle behind how large a particular unit was to be was defined by how quickly information could be relayed within the unit. At the time when the *départements* of France were formed, they would have to be no larger than the distance information could travel in a day from the capital of the *département* to a given point on its periphery. In the late 18th Century, the limit to the speed of information relay was the speed of a horse. However, this revolutionary ideal soon became outdated, as even by the 19th Century, railways first and telegraphs later would enable larger distances to be crossed in much less time. By the 20th Century, information could travel at virtually the speed of light, but this has left administrative units largely unaffected. The rapid acceleration of the speed of information relay has changed our lives, but seemingly it hasn't even scathed our administrative structure. It is therefore high time something was done about it.

8. Principles of Sustainable Administrative Systems

The roots of the current administrative system lie deep in the history of the continent. State borders have changed a number of times over the centuries, each time delimiting the borders of the economy. Therefore, as pointed out by Hevér, it is hardly surprising that the earliest spatial economic models have all dealt with some sort of a border to the reach of the economic activity. [Hevér, 2009] Under 21st Century conditions, such delimitations may seem greatly outdated, but in fact, they still form the cornerstones of many hypotheses today. Let us now follow up on Hevér's argumentation about geographic space and the economy. In the following argumentation, one ought to think of an administrative area as a *quasi-isolated state*; one that is not completely isolated, but one that ideally ought to be self-sustaining.

As the concept of sustainable development itself has three key dimensions, this will transitionally hold for the key aspects of sustainable administrative systems. Looking at the properties and nature of administrative units, whilst bearing in mind the dimensions of the concept of sustainable development, it can be deduced that the capability of an administrative unit to become sustainable depends on the following five aspects:

- the geographic size of the administrative unit,
- its richness or poverty in resources (including environmental source and sink functions),
- the geographic location of the administrative unit,

- the human population the administrative unit has to sustain,
- the legally defined roles (rights, duties, authority to take decisions), i.e. the weight of the administrative unit.

To offer an explanation as to why these five aspects were found to be so essential, the aspects ought to be grouped together. The first three each have to do with establishing the potential biocapacity of the administrative unit, which, when compared to *von Thünen*'s models, is a similar approach comprising more dimensions simultaneously. The first three aspects will act as the "supply factors", which will be determined by how "big" the unit is, how "rich" it is (not financially, but in resources), and where it is. It is easy to see that these three factors fundamentally determine the potential in biocapacity of a unit. Clearly, a unit will not become sustainable just because it is big, because that will be influenced by its location on the globe, as well as what stock of resources (and flows derived from that) a unit will possess.

The human population the administrative unit will have to sustain will also play a crucial role, as this will determine the demand for biocapacity. This will be influenced by a multitude of factors, including current and desired standards of living, education and culture, expected life span and other sanitary conditions, as well as traditions and beliefs.

Finally, the fifth aspect takes into consideration the legally defined roles and the political weight – the ability to assert the interests – of the administrative unit. Clearly, a unit with a lot of power and little external control will be tempted to misuse its resources, which may deplete quickly and irreversibly. Conversely, a unit with little power and a lot of external control will not be able to do this, but they may lose out on some of their local potentials, which will reduce their standard of living. Optimally, the size of each local administrative unit should be determined by its capability of managing its ecological footprint, whilst allowing the principle of subsidiarity to be used to its greatest extent.

As it has been showed previously, the European Union has to confront an administration that is both diverse and fragmented. The ultimate danger with this is that there are a lot of very small units which have extensive powers, with effects that can reach beyond their limits quite easily. Indeed, on the one hand small units will generally lack the holism required by sustainable development. On the other hand, they will be plagued by their duties which they may find costly – as a direct effect of losing out on economies of scale – and they may be tempted to disobey some of them as a cost-saving effort. Would it not be rational, then, to expand or join adjacent administrative units, so they can both

benefit from holism and economies of scale? Advocates of the principle of subsidiarity may find this to be their worst nightmare, but in fact it is not contrary to the principle. Certainly, from a sustainability point-of-view, it is also a desirable move.

9. New trends gaining pace

Some countries – particularly in Scandinavia – have attempted bold and profound changes over the last few decades, and their results were all positive. The paths they have taken were not identical, but as a common theme they have identified and scrapped unnecessary levels of administration. Denmark, Sweden and Finland – but also Norway, currently outside the EU but very much a strong partner in Scandinavia – have all reviewed their administrative structures, and have taken to sometimes seemingly extreme steps to rationalise. The key points to their success were forming larger communities and empowering the regions.

10. A Europe of Regions

Regions have been in existence in some EU member states for nearly three decades now, yet most of them have still not been empowered by true autonomy. France and Spain were amongst the very fist states to create regions, in both cases establishing them to prevent the country from falling apart. This was relatively quickly complemented by the delegation of powers, drawing it partly from the state itself and partly from lower-level administrative units. This trend has been followed by a number of states – large enough to actually rely on regions, which effectively leaves the smallest states of Cyprus, Luxembourg and Malta excluded.

However, it appears that there may be some reluctance from some medium-sized states to follow suit. Indeed, even though statistical regions have been called to existence in most states – albeit for financial reasons, – in several cases autonomy was not granted and roles were not delegated, which could be regarded as a violation of the Principle of Subsidiarity. Abundant proof has been gathered about the usefulness the regions and most recently, even one of the states most reluctant to modernise its administrative system – Greece – has followed suit. As the distribution of power shifts within the European

administrative system, states reluctant to adopt changes may have to face the challenge of a quick adaptation, which may hold a risk of failure.

11. Forming communes – localities re-imagined

In most – in fact, almost all cases in the European Union – forming larger units, *communes* by grouping them together could be considered a desirable advance move. If localities are given clear guidelines and assistance, and larger units are allowed to be formed along natural delineations, cohesion within the unit is more likely to be stronger than otherwise. Essentially, as information can be disseminated at the speed of light, and as people are getting more used to this pace, and indeed, as many of us are already getting used to dealing with our day-to-day business over the internet, the necessity of having so many agencies of local administration has all but vanished.

This trend has not only been observed in Scandinavia, but also the Benelux and Germany has started to follow suit, and even in France – where the constitution preserves the status of the 36 000 localities – some changes have been taking place along these lines. In the case of France, since their localities cannot be scrapped, they bridged this problem by allowing adjacent localities to form corporations, which would essentially take over a number of duties from the localities, thus alleviating this problem, and thus preventing the localities from being scrapped altogether. As most citizens identify themselves with their localities, this was considered to be a favourable move. However, it can also be observed, that most corporations will be given *just enough* resources to carry out their duties, which may hold negative consequences for investing in the future.

In reality, fears of "losing a local identity" have seemed to be unfounded so far, as even within a larger *commune* the former locality may preserve its elements of identity, the symbols most people adhere to.

More recently, an even more peculiar formation has come to existence. Akin to the previously existing *euroregions*, which is a trans-border cooperation between some naturally existing regions – divided by a state border – *eurodistricts* have began to be formed. For now, these are mostly isolated cases of trans-border local cooperation, but perhaps in the near future this may be quite a common form of local administration. The potential that lies within this new structure is heralded by its first applicants: of all places, the first *eurodistricts* have been formed between France and Germany, between former foes, between two completely different jurisdictions. The underlying message is overwhelming:

state borders within the European Union become obscure and mostly irrelevant, and natural cohesions are allowed to emerge. To give an example of such a *eurodistrict*, we need not go further than the Strasbourg-Ortenau eurodistrict, which is already serving citizens on both side of the River Rhine, and which some are ready to bill "the Washington, D. C. of Europe"[4]. Irrespective of the fact that the capitals of the two countries are more than a thousand kilometres apart, citizens in this border area are still locals and close neighbours, and even across a state border it makes sense to harmonise services, coordinate tasks and share burdens, especially if it advances the efforts towards sustainable development. This is not irredentism, but the true Spirit of Europe in practice.

12. Conclusion

The European Union remains one of the most diverse entities on Earth, and it seems to have found a way to draw its strength from its colourful people. As the Union has committed itself to step on the path towards sustainable development, its administrative heritage must be reviewed for conformity. To better serve the needs of citizens, some trends can now be observed across the Union, streamlining bulky administration, whilst also empowering citizens and decision makers to make choices about their future more responsibly.

Whereas most states have carried out at least a bare minimum of administrative changes – and whereas others have taken outright bold moves towards streamlining their internal structures – a handful have been reluctant to do almost anything that would be both proactive and beneficial. As several states have already implemented measures, a smorgasbord of possible models is available, with doubtless clarity of their success. Indeed, the fact that some cooperative schemes have arisen under the most unlikely circumstances, should hint to those still remaining hesitant that in reality, their fears may be unfounded. For those refusing to adapt, this is likely to pose a threat of a costly and inefficient administration having to deal with the social, economic and environmental

---

[4] One need not wonder that such a formation would be called to existence in this particular case: as previously mentioned, the area relies heavily on the European Parliament and other European institutions for their income. With such a bold attempt to show the area as an attractive venue, united in true European spirit, it becomes somewhat more difficult to argue that the two-seat conundrum be solved quickly. However, it would be deeply cynical to say that this is the only reason such co-operations come to existence, as we can see from a number of other examples such partnerships are formed even where the European Parliament is not in session.

challenges of sustainable development, ultimately not living up to the expectations of citizens about the administration serving them.

References

GIRARDON, JEAN (2008): *L'intercommunalité*. Éditions Ellipses, Párizs. ISBN: 978-2-7298-3703-7.

HEVÉR, BOGLÁRKA (2009): Spatial Economics – Back to the Origin. Government and Private Sector in Transition Economies, pp. 49-51. *BERG Public Economic Series*, Vol. 18. ISBN: 978-3-931052-73-7.

Nahverkehr. Eurodistrikt Strasbourg-Ortenau. http://www.eurodistrict.eu/de/Nahverkehr-19.html. 11 January 2011

Regulation (EC) No 1059/2003 of the European Parliament and of the Council of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS). *Official Journal of the European Union*, L154/2003, pp 1-41.

Review of the EU Sustainable Development Strategy (EU SDS) − Renewed Strategy. Pp. 2-5. The Council of the European Union, Brussels, 26 June 2006. http://register.consilium.europa.eu/pdf/en/06/st10/st10917.en06.pdf

Strasbourg Region: A Bi-national Story of Success. COMET Competitive Metropolises – Information for Investors. http://www.oeaw.ac.at/isr/comet/documents/Final_Results/COMET_deliv _no11_leafletStrasbourg.pdf. 11 January 2011.

The Seat Question – Fact sheet on the European Parliament. OneSeat Campaign http://www.europafederalisterna.se/oneseat/?view=facts&lang=en. 13 January 2011.

CUMULATIVE LOGISTIC REGRESSION APPROACH TO DEVELOP SCORECARDS ON DATABASES WITHOUT DEFAULT EXPERIENCE

László Madar

1. Introduction and summary

For financial institutions that calculate their capital requirement for themselves under the Internal Ratings Based approach (abbreviated as IRB), it is a very important to assess the riskiness of all types of portfolio segments in an appropriate way. However, there are portfolios – and these ones can also be significant in size – that cannot be treated with ordinary statistical methods, because the target variable is inappropriate for modelling – there is no default, or only a very limited number of defaults. Commonly used statistical techniques, such as binary logistic regression may only be used at scorecard developments when there is explicit and numerous information about the defaults and losses happened historically in the portfolio.

These types of segments may be for example a portfolio of large financial institutions, internationally active corporations, local municipalities in some countries or project loans (or segments of project loans) when there is typically a limited number of population and from this population there is a default event number of zero, or in the worst case by considering recent recession periods can be only a handful per year. However, from these types of exposures can compose a significant amount of the financial institution's balance sheet amount. Because the regulation limits the exposure and risk-weighted amount (denoted as RWA) that may be kept unrated under the standardized approach, IRB institutions shall find techniques to rate all these clients.

Segments, where there is such a low number of clients or/and defaults, cannot be handled with industry standard methodologies, such as logistic regression or decision trees, because there is no statistically significant number of default events – the methodology breaks down, delivers unstable results. The logistic regression works only if there are a statistically acceptable number of events in the target variable, else the regression will deliver a very doubtful result that may

not be accepted at any confidence levels. That is, logistic regression or decision trees (or any other methodologies aiming for a default/non-default target variable) are working well only at high-default environment. At a low-default scenario, new ways of rating systems shall be found in order to arrive at an acceptable statistical result.

Institutions may overcome this problem by developing their rating systems on an expert basis or derive these rating systems from high-default environment; however, representativeness shall be still justified. Human errors may bias the final rating system, and may lead to false quantification of risk. In this study a statistical alternative is proposed that may be used as an IRB-conform development method for assessing the riskiness of clients. There are two ways proposed, the first, easier part is to develop a rating system that maps to an external rating scale, where the information of large external rating institutions is derived from their solicited ratings. However, there are portfolio segments where no external rating information is available, therefore internal knowledge shall be accumulated in a consistent way as the first step. After this, the final methodology aims to build a statistical model between the rank-ordered client population and explanatory variables. This enables financial institutions to create a scorecard, collecting all internal (and/or external) knowledge on these types of clients, and overcome the limitations of regular logistic regression methods.

The proposed logic defines a rank-ordered classification of clients based on their expected level of riskiness. This will be the target variable, which is not binary, rather a rank-ordered variable. Explanatory variables are collected in a usual way. A cumulative logistic regression is performed, and final results are interpreted accordingly.

2. Model setup

The used model is the cumulative logistic regression (also known as ordered or ordinal logistic regression). The model uses an ordered input variable and a set of explanatory variables to build a single model that can evaluate to which category an observation is likely to fall.

If there is a logical rank ordering of clients, then this model is able to set up a structure of dependency between some of the objective characteristics of the loan or client. Like other types of modelling tasks, this method enables financial institutions to create a rating/scoring system where traditional methods are likely to fail. This may be an appropriate tool for determining the factors in a logical

way, and setting the most appropriate weighting that describes complex risk perception of existing clients. This method of rating creation overcomes several problems that expert scorecards may contain. When building a scorecard for these types of portfolios in an expert way, there are two main sources of errors: first, the perceived risk of the non-defaulted portfolio may not describe the truth completely; second, the weighting of explanatory variables may be fortuitous. The second error source may be diminished completely by using the proposed approach detailed here, only the expected (perceived) riskiness of clients may be wrong. Due to the nature of the portfolio, this first type of error may never be diminished, only minimized by using external ratings if available (combination of internal and external information), or general logical methods from decision theory that enables the financial institution to combine expert opinions in a most effective way.

The setup of the model requires a rank ordering of clients along riskiness, that is, different levels of riskiness shall be set and assigned to clients. The number of categories to be set is important, normally they should be clearly differentiated, and therefore there should be not many categories where the loan types may be assigned to. This rank ordering may be external orderings of clients, such as solicited rating of rating agencies, where the rating agency performs a due diligence on the client and comes up with a final rating summarizing the overall level of risk of the client. In this case, the ratings may be used as the target variable, and the regression task becomes very easy.

However, not all portfolios have appropriate external ratings. The most common examples are local municipalities or project loans. For these types of portfolios internal ranking shall be created and summarized. Internal knowledge must be collected and evaluated in an appropriate way that minimizes perception errors and combines all expert knowledge the financial institution has. One very useful technique for this is the Analytical Hierarchy Process, abbreviated as AHP.

The Analytic Hierarchy Process (AHP) is a structured technique for dealing with complex decisions. Rather than prescribing a "correct" decision, the AHP helps decision makers find one that best suits their goal and their understanding of the problem. During the process of AHP decision makers (credit experts) first decompose their decision problem into a hierarchy of more easily comprehended sub-problems, so to speak they assess which mail elements shall be considered as most important to assessing the riskiness of a certain client or deal. The elements of the hierarchy can relate to any aspect of the client/loan, however here not the final vast number of explanatory variables (like financial ratios) are used, rather a few more general terms may be included (e.g. management quality, financial strength, etc.), altogether making the decision problem transparent. Once the

aspects are selected, the decision makers systematically evaluate its various elements by comparing them to one another one at a time, with respect to their impact on an element above them in the hierarchy. In making the comparisons, the decision makers may use exact data about the elements, but the aim is that they use their expert judgments about the elements' relative meaning and importance. It is the essence of the AHP that human judgments, and not just the underlying hard information, can be used in performing the evaluations. The AHP converts these evaluations to numerical values that can be processed and compared over the entire range of the problem. A numerical weight or priority is derived for each element of the hierarchy, allowing diverse and often incommensurable elements to be compared to one another in a rational and consistent way. Also, internal inconsistency of the final ordering may be measured and if this inconsistency is high, the process may be repeated and areas (or decision makers) from which inconsistency is coming may be corrected.

Pair-wise comparison is used in AHP because of the fact that human decision is much easier if two alternatives are to be compared, rather than assessing or ranking a vast number of alternatives. Still, if the alternative number is high (that is, there are hundreds of clients or loans in the portfolio), it is highly recommended to use a ranking AHP solution, rather than the baseline solution that would compare pair-wise all clients/loans.

The general form of a pair-wise comparison matrix is the following:

$$
S = \begin{pmatrix}
 & A_1 & A_2 & \cdots & A_n \\
A_1 & p_1/p_1 & p_1/p_2 & \cdots & p_1/p_n \\
A_2 & p_2/p_1 & p_2/p_2 & \cdots & p_2/p_n \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
A_n & p_n/p_1 & p_n/p_2 & \cdots & p_n/p_n
\end{pmatrix}
$$

where A_i denotes alternatives, p_i i=1..n are real numbers.

Such a matrix is a reciprocal matrix, that is $a_{ij} = 1/a_{ij}$, and it shall be best if it would be a consistent matrix also, that is $a_{ij}a_{jk}a_{ki}=1$, for all i, j, k. It is unlikely to happen (if only the alternatives are not measured by exact values), as the number of alternatives grow, the inconsistencies will be more abundant. Inconsistency of the matrix can be measured with several indices.

Decision is made with the help of a decision table. For a ranking AHP model, not all the alternatives need to be compared, only the decision factors must be weighted. However, for all alternatives (loans), the characteristics must be evaluated also, with the following formula:

$$A_i = \sum_{j=1}^{m} \frac{w_j}{w} \frac{1}{\hat{a}_i} a_{ij}$$

Where w_j is the weight of characteristic j, w is the sum of all weights, $\hat{a}_i$ is the maximum possible value of the given characteristic. This logic scores the alternatives based on previously set weights, and may be performed also for a larger number of clients/loans. The final result is a continuous scale that shall be quantized according to an expert basis to be used in the second step, during the rating creation process.

The cumulative regression model uses cumulative probabilities. The structure of the model can be described in the following way. Let the response variable $Y=1,2,..., J$ that is ordered by the riskiness of the loans/clients. The related probabilities are $\{\pi_1, \pi_2, ..., \pi_J\}$, and a cumulative probability of a response less than equal to j is:

$P(Y \leq j) = \pi_1 + ... + \pi_j$

Using this notation, a cumulative logit can be defined as the following:

$$\log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \log\left(\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right) = \log\left(\frac{\pi_1 + ... + \pi_j)}{\pi_{j+1} + ... + \pi_J}\right)$$

This describes the log-odds of two cumulative probabilities; that is how likely is the response to be a category of equal to or below j versus a response that's higher than j.

This model is a common way to approach the modelling of a J-category response variable through the use of cumulative logits, given that the response may be characterized by a set of explanatory variable vector $X=x_1, ..., x_n$:

$$P(Y \leq j \mid X) = \frac{1}{1 + e^{-\Theta_j + \beta_1 x_1 + ... + \beta_n x_n}}$$

It may be noticed that for all response the same coefficients ($b_1, ..., b_n$) are regressed and reported as a result. This means that the only difference between the estimated equations is the $-\Theta_j$ parameter which is the constant for category j. In this way, parallel logistic curves are estimated for each category j.

88

Finally the single probabilities of belonging to a pre-set category j may be derived as:

$$P(Y = 1 \mid X) = P(Y \leq 1 \mid X)$$
$$P(Y = j \mid X) = P(Y \leq j \mid X) - P(Y \leq j - 1 \mid X)$$
$$P(Y = J \mid X) = 1 - P(Y \leq J - 1 \mid X)$$

For the first and last category the calculation is special, because single probabilities are derived as the difference of the cumulative probabilities. The given case is assigned to that category, where its non-cumulative probability is the highest. Under normal circumstances, the estimation will result in a normal-curve shaped estimation for a single client regarding the possible range of categories, that is, estimated probability can be assigned to the final category precisely and with ease.

This model, called the proportional-odds cumulative logit model, has $(J - 1)$ intercepts plus n slopes, for a total of $n + J - 1$ free parameters. Notice that intercepts can differ, but that slope for each variable stays the same across different equations. There is an extended model, where coefficients may differ from each other, giving a much more flexibility, but requiring a very extensive dataset.

3. Data for modelling

In this chapter, the proposed cumulative regression model is tested against ordinary (binary) logistic regression. As there is no portfolio containing internal information available that may be used to set up a rating scale, publicly available data with external ratings will be used for testing purposes. A 100 case portfolio of bonds issued by large companies are selected, where there were no defaults during the observation period. The portfolio is ordered along their main external rating by S&P. A representative database were selected that were stratified to main external rating, since there are much more companies on the better part of the scale as from the worst category in this segment.

Also there is publicly available information about these clients, client-level information and issue-level information as well. From these public information a small set of explanatory variable could be created.

The following explanatory variables were used:

- operative margin of issuer
- capita, ratio of issuer
- gearing ratio of issuer
- leverage
- cash ratio
- acidity ratio
- volume/turnover
- assets/debt ratio

The cumulative logistic regression was run on a sample of 100 clients, rated from AAA to C (altogether 7 categories) by Standard and Poor's. The logistic regression was run by using the software of SAS Learning Edition. Due to the low number of cases, significance level was set to a large 15% to retain more and better performing variables even if they not fit into the usual 5% significance limit.
The final cumulative logistic regression result is the following:

Table 1.: Model result of the cumulative logistic regression procedure

Analysis of Maximum Likelihood Estimates					
Parameter		Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	9.185	4.0463	5.1528	0.0232
Intercept	2	10.546	4.0526	6.7716	0.0093
Intercept	3	11.474	4.067	7.9599	0.0048
Intercept	4	12.459	4.0921	9.2705	0.0023
Intercept	5	13.798	4.1449	11.082	0.0009
Intercept	6	15.288	4.2134	13.165	0.0003
Operative margin		1.8548	0.475	15.25	<.0001
Gearing ratio		5.9151	2.4078	6.0349	0.014
Capital ratio		-32.58	8.8585	13.522	0.0002
Volume/turnover		1.7584	0.4675	14.15	0.0002
Assets/debts		2.3146	1.5964	2.1021	0.1471

Intercepts are set for each category boundary, and five explanatory variables were selected to the final model.
Also a benchmark model was created in order to test the validity of the cumulative logistic regression. This was created using a trick to assign the lowest rated category to "event" (set rating class of "C" to default category, even if they

were non-defaulted), and with the help of this, a usual binary logistic regression was run. Here also confidence level was set to 15% in order to have the same conditions for binary as for cumulative logistic regression.

The result of the benchmark model was the following:

Table 2.: Model result of the binary logistic regression procedure

Analysis of Maximum Likelihood Estimates				
Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	9.51	2.6108	13.268	0.0003
Operative margin	2.6697	0.8453	9.9756	0.0016
Capital ration	-20.5375	4.5605	20.2801	<.0001
Volume/turnover	2.6439	0.9253	8.1644	0.0043

The benchmark model uses only three relevant variables. All model variables are very significant, that forecast a robust model.

4. Model testing and model choice

The final model was benchmarked with standard statistical tests, the ROC-Curve (AUROC index) and KS measure.

The ROC-curve is built up the following way. The x-axis contains the proportion of good customers to all good customers in a cumulative way. As we go through all possible cut-off points, a proportion of the good clients, and a proportion of bad clients (related to the total count of bads) are excluded. This excluded population proportion is shown in the chart, the curve point is determined by the relative measure of excluded goods and bads. If the rating system is better, it can identify bad customers sooner, resulting in a steeper ROC-curve (proportionally more bads are excluded than goods using the given cut-off). This curve has also an index figure, the AUROC or AUC index, that can be exactly mapped to the standardized Gini-index. The AUC=(B+C)/(A+B+C). The Gini=2*AUC-1, that rates the power between 0 % to 100 % from worst to perfect ordering.

The Kolmogorov-Smirnov statistic compares two distribution functions and calculates the maximum absolute deviation from each other:

$$T = \max \left| G(x) - F(x) \right|,$$

where G is one and F is the other distribution function (distribution of goods and bads).

The representations of these two tests are shown in the following graph.

Graph 1.: On the left hand side the theoretical representation of the ROC curve can be seen, on the right hand side the composition of the KS measure is shown
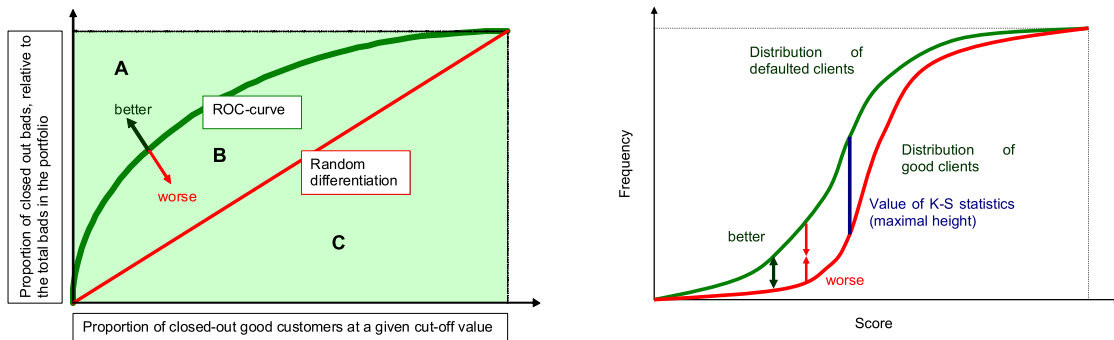
The two systems were tested against each other. For the charts to be drawn a good-bad split must have been made for plotting the ROC-curve. This may bias performance measurement to Logistic regression side. Bads were the non-investment grade bonds (With S&P rating of "BA", "B" and "C") while goods were the investment-grade issues from the population. The following graphs and test statistics were acquired:

Graph 2.: ROC-curve of the cumulative logistic (left) and binary logistic regression
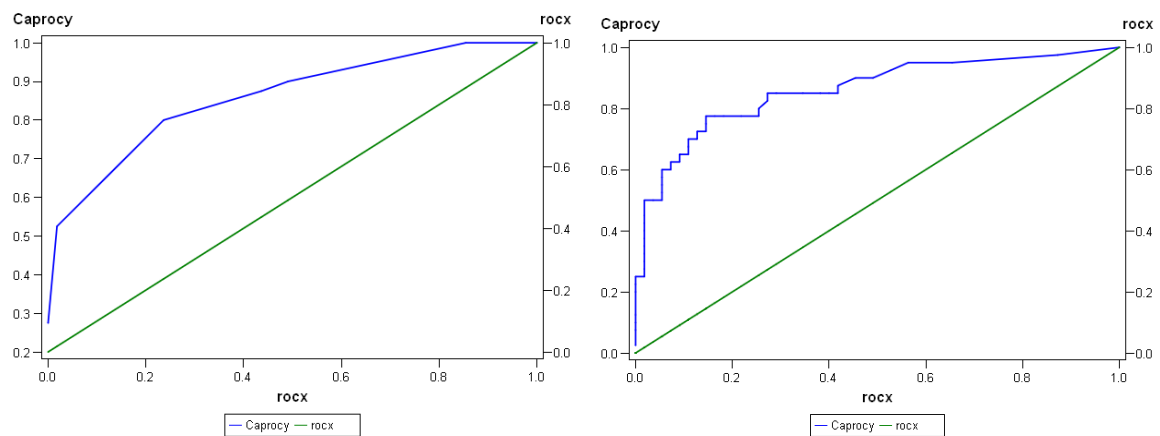
Graph 3.: KS-curve of the cumulative logistic (left) and binary logistic regression
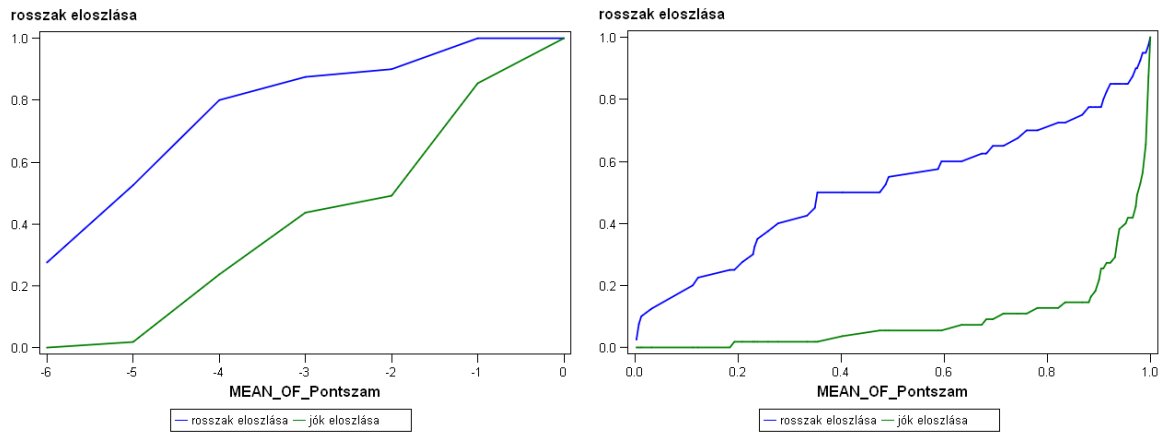
The power measures for the models are the following:

Table 3.: Test statistics of the cumulative logistic regression and the benchmark model

	AUC	GINI	K-S
Cumulative	0.859	71.7%	56.4%
Binary	0.860	72.0%	63.0%

Models are powerful; they show almost the same power measure. However, according to ordinary logistic scorecard tests, the binary logistic model performs a little bit better, but this performance improvement is well inside the significance borders of the estimation. However, the model fit to all types of categories is better at the cumulative logistic regression. This may be seen in the following tables, where actual categories are mapped against the estimated scores. Here, binary logistic regression result shows a large jump in the ordering the clients into categories.

Graph 4.: Model fitting plot of the cumulative logistic regression (left) and logistic regression (right)
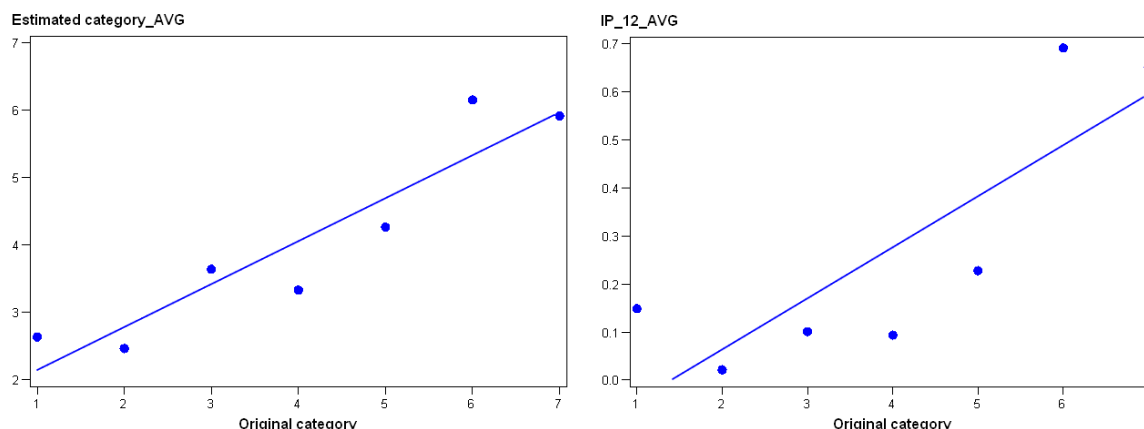
From the graphs above it may be seen that for the cumulative logistic regression the final fit is much better. The points are distributed evenly and seemingly in a random way along the estimation line, while binary logistic regression shows some autocorrelation, weighting to the worst category and not assigning good risk measures to medium categories.

Model choice should be for the cumulative logistic regression, since it can deliver the same discriminatory power as the binary logistic regression, but it can also fit to all categories precisely. The result of the regression not only says that the given client will be good or bad, it can also assign a level of expert riskiness to the given client or loan. This helps to build such a model that can be used in low-default or zero-default environment, and can result a well-working risk forecast system.

5. Conclusion

From this short test it seems that there are good ways to use statistical measures to contract diverse information in a logical way, and establish a well-working system for portfolios that cannot be handled with common statistical tools. Where the industry-standard best practice solutions fail, there has to be new ways developed. When there is no default experience then it is suggested to set up a logical internal ordering of clients and develop a scorecard based on objective measures to obtain a logical and well-performing statistical tool for assessing the single entities of the portfolio. Beside yearly review, a developed

94

scoring system is good for monitor and assess portfolio elements in a comparable and efficient way, if basic information, explanatory variables are provided. Cumulative logistic regression fits to rank ordering better than binary logistic regression, and delivers a simple scorecard that assesses all clients the same way. Use of this model is recommended where there are low default portfolios that need to be kept under the IRB approach.

References

HOSMER, DAVID W.; STANLEY LEMESHOW (2000): *Applied Logistic Regression*, 2nd ed.. New York; Chichester, Wiley. ISBN 0-471-35632-8.

SAATY, THOMAS L.; PENIWATI, KIRTI (2008): *Group Decision Making: Drawing out and Reconciling Differences.* Pittsburgh, Pennsylvania: RWS Publications. ISBN 978-1-888603-08-8.

SAATY, THOMAS L. (2008-06): "Relative Measurement and its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors - The Analytic Hierarchy/Network Process". *RACSAM (Review of the Royal Spanish Academy of Sciences, Series A, Mathematics)* 102 (2): 251–318. http://www.rac.es/ficheros/doc/00576.PDF.

SAATY, THOMAS L.; ERNEST H. FORMAN (1992): *The Hierarchon: A Dictionary of Hierarchies*. Pittsburgh, Pennsylvania: RWS Publications.
ISBN 0-9620317-5-5.

SAATY, THOMAS L. (2001): *Fundamentals of Decision Making and Priority Theory*. Pittsburgh, Pennsylvania: RWS Publications.
ISBN 0-9620317-6-3.

MODELLING OPTIONALITIES IN RETAIL BANKING PORTFOLIOS
– IMPACT OF INTEREST RATE CHANGES ON THE NET PRESENT VALUE OF BANKS –

PETRA KALFMANN

1. Summary

Interest rate risk in the banking book is in the focus of risk management practices since the risk factor was nominated in the Basel 2 framework. The interest rate risk as part of market risk has been measured and managed accordingly by banks in the trading book since 1996, but since banking book assets are mainly not traded it is not possible to use the same methodology for banking book risk measurement purposes. The interest rate risk of the banking book can be derived from the following risk elements: repricing risk, yield curve risk, basis risk and embedded options. The first three elements are related to market interest rate movements, and therefore their impact on the banking book can be measured with accordingly defined at-risk models (Earnings-at-risk, Economic Value of Capital).

The current study puts emphasis on a special part of interest rate risk of the banking book: namely the embedded options. This risk element is specific mainly in large retail portfolios, since the widening of banking products offered for retail customers have introduced many kind of optionalities into the portfolios. In case of loans the early prepayment possibility offered for clients is the main source of embedded options, while in case of deposits the contracts without predefined maturity (sight deposits) holds option like elements. The evaluation of embedded options and their pricing into the products is essential from the bank's point of view, since these possibilities affect significantly the asset structure, the profitability and the liquidity position of the bank.

The evaluation techniques used for asset side and liability side options are different, therefore in the present study I deal with asset side optionality: the early prepayment option.

2. Interest rate risk in banking book

The Basel 2 regulatory regime has introduced a new element into the capital regulatory framework: the so called Pillar 2, which is currently known and defined as the Internal Capital Adequacy Assessment Process (ICAAP). Pillar 1 covers the minimum regulatory capital requirement, while Pillar 3 puts emphasis on market discipline through transparent disclosure requirements.

In the framework of ICAAP the institution is responsible for identifying and evaluating the relevant risk factors the institution faces, and for each relevant risk factor sophisticated measurement methodology has to developed and introduced. In practice ICAAP methodology is introduced through the framework of economic capital models. Based on the guidelines set by the Basel Committee on Banking Supervision (BCBS) interest rate risk in the banking book has to be identified and managed in the framework of ICAAP.

The first guidelines on the measurement of interest rate risk in the banking book were published by BIS in July 2004[1], based on which the Committee of European Banking Supervisors (CEBS) published several guidelines discussed with the financial sector, the latest in 2006[2].

Interest rate risk in the banking book by definition is the risk to both earnings and capital in the banking book of institutions arising from adverse movements in interest rates. Credit institutions as financial intermediators face essential interest rate risk; because of maturity transformation they perform the institutions run the risk arising from the difference between the returns related to different maturities.

The interest rate risk in the banking book can be measured in a different way than the risk arising from trading book. The assets allocated to the trading book are evaluated on a daily basis with mark-to-market evaluation, and this way the loss/gain resulting from the movements of interest rates can be measured every day. In contrary the assets in the banking book are usually not traded and therefore mark-to-market evaluation cannot be used. Because of this reason there is no one commonly used and accepted best practice methodology for the measurement of interest rate risk in the banking book, like Value-at-Risk for market risk.

The effect interest rate risk in the banking book based on international literature, guidelines of supervisors and practice of international banking groups can be

---

[1] Principles for the Management and Supervision of Interest Rate Risk, BIS, July 2004
[2] Technical aspects of the management of interest rate risk arising from non-trading activities under the supervisory review process, CEBS, October, 2006

measured in two ways. The income based methodology is aiming at measuring the impact of interest rate movements primarily on the net interest income (NIM), since its scope is to calculate the short term impact of interest rate movements on profitability. The economic capital methodology is aiming at measuring the critical change in the "market" value of the bank capital with re-evaluating the cash-flow of the bank, since its scope is to evaluate the long term risk implications, which can be caught through the value change in the bank capital.

The sources of the interest rate risk in the banking book are the following based on CEBS guideline:

- **repricing risk**: risks related to the timing mismatch in the maturity and repricing of assets and liabilities and off balance sheet short and long term positions; e.g. if a bank funds its long term fixed rate loans from short deposits, in case of interest rate increase the profitability of the bank will decrease since its repriced funds will be more and more expensive.
- **yield curve risk**: risk arising from changes in the slope and the shape of the yield curve;
- **basis risk**: risks arising from hedging exposure to one interest rate with exposure to a rate which reprices under slightly different conditions;
- **embedded options**: risks arising from options, including embedded options, e.g consumers redeeming fixed rate products when market rates change.

3. Embedded options in retail banking book

As it was referred to in Chapter 2 optionality in retail banking books can be a sufficient risk factor. In this chapter I will give the definition and sources of the embedded options in the banking book of large retail portfolios.

3.1. Asset side optionality

The asset side optionality covers those risk factors which are embedded in the characteristics of the banking assets. The asset side optionality mainly covers those characteristics, that borrowers have the possibility to repay the loan (partially or totally) before maturity. This is called the **prepayment optionality**.

Prepayment optionality guarantees the borrower to repay its debt before maturity, so actually it can be defined as a **long call option** on the debt from the borrower's point of view, while it is a short call option from the bank's perspective. Since the borrower owns the right it is rational to price this risk into the cost of the loan. Prepayment optionality is a sufficient risk factor, since the early prepayment of a loan affects the maturity structure and the profitability of the bank.

3.2. Liability side optionality

The liability side optionality covers those risk factors which are linked to the retail deposits of the bank. This risk factor means that the depositors have the right to withdraw their deposits from the bank at any point in time. From this perspective it makes sense to distinguish sight deposits and term deposits: sight deposits don't have a contracted maturity, while term deposits have, so different modelling possibilities should be used to measure the value of the options linked to these products.

The withdrawal optionality guarantees the depositor a **long put option** on the deposit, while it is a short put position from the bank's point of view. Since the debtor owns the right it is rational to price this risk into the price of the deposits. The withdrawal optionality is a sufficient risk factor, since it affects the liquidity position of the bank.

4. Evaluation of prepayment optionality

The prepayment optionality can be derived from the following events:
- **refinancing**: this covers those events when the borrower repays the loan totally from another bank loan. From the borrower's point of view it is rational if the condition of the new loan is more advantageous than the existing one.
- **early repayment**: this covers those repayment events when the borrower repays the loan partially or totally from his accumulated savings, and not from a new loan.
- **sale of real estate**: this is a special event related to early repayment which happens when the borrower repays its loan before maturity because the

financed real estate (considered also as collateral) is sold and the financing is not required anymore.

- **default**: this is a special case of the three above mentioned events. In case of the default of the borrower it might happen that the borrower repays its loan by:
  - o refinancing its debt from another bank loan;
  - o its accumulated savings or other external financial resources;
  - o forced sale of the collateral.

In the following we focus on the prepayment optionality linked to retail mortgage loans, since the evaluation of prepayment possibility makes sense in case of long-term loans.

4.1. General framework

The prepayment optionality evaluation on the net present value of the banking book can be considered in the following framework.

- First a predictive model has to be developed for forecasting in what circumstances prepayment can happen.
- Second a business model has to be developed for measuring the impact of early repayment on
  - o profitability (short term view) and/or
  - o net present value of the bank (long term view).

4.2. Modelling prepayment optionality

The models developed for measuring prepayment optionality of retail mortgage portfolios have a well-based literature in US. Historically the market of mortgage backed securities (hereinafter referred to as MBS) was developed in the US; the pricing of these assets assume the modelling of prepayment optionality, which must be considered in the cash-flows (and as a consequence in the price) of MBS. The pricing of MBS is closely related to capital market prices, because MBS have had well-developed secondary market.

This is not the case in EU markets, just like in Hungary, where mortgage loans don't have a secondary market. Because of the difference in the market structure

of US and EU mortgage markets and financing, the US based models and relevant literature cannot be implemented without modifications on EU markets. The models concerning can be divided into two types:

- **models for optimal repayment**: these types of models are used for modelling prepayment optionality and its implications on NPV of banking books in those events when it is financially rational for borrowers to repay before maturity. Typically these events cover *refinancing events*.
- **models for non-optimal repayment**: these types of models concern those events when repayment happens not necessarily because it is financially rational, but because other factors affect the repayment, e.g. early repayment because of accumulated savings, sale of collateral. These factors can be considered as *behavioural factors*.

The financial rationality in case of **models for optimal repayment** implies those events when the borrower receives a more advantageous offer from another bank. The most important element of a new offer is the interest rate, and therefore the required inputs of these types of models are consisted of:

- market data, e.g. macro economic drivers,
- internal bank data: pattern of prepayment because of refinancing based on historical data, e.g. loan specific drivers.

The following chart presents the logical framework of models for optimal repayment and its effect on NPV of banking book:

**Chart 1.:** Logical framework of models for optimal repayment
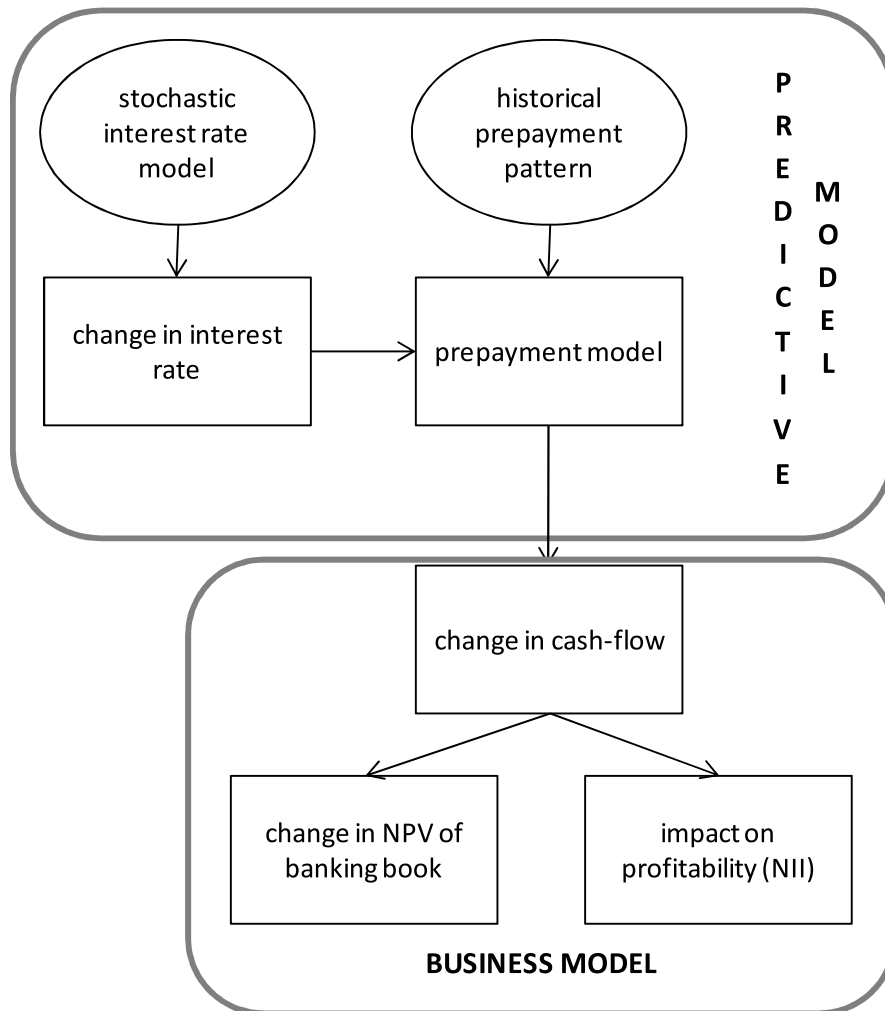
The behaviour of borrowers in case of **models for non-optimal repayment** implies those events when the borrower repays earlier because of non-financial reasons. This type of repayment can be modelled in two ways:

- modelling the behaviour of banking clients on prepayment excluding refinancing: this type of modelling assumes to develop a behavioural scorecard for clients who repayed their debt earlier (borrower specific drivers);
- measuring historical prepayment: if scorecard cannot be developed or in very poor quality the pattern of prepayment (excl. refinancing) should be measured.

- Since this model type is trying the capture the effect of behaviour of banking clients, it is essentially based on internal banking data on which a scorecard can be developed.

The following chart presents the logical framework of models for non-optimal repayment:

Chart 2.: Logical framework of models for non-optimal repayment
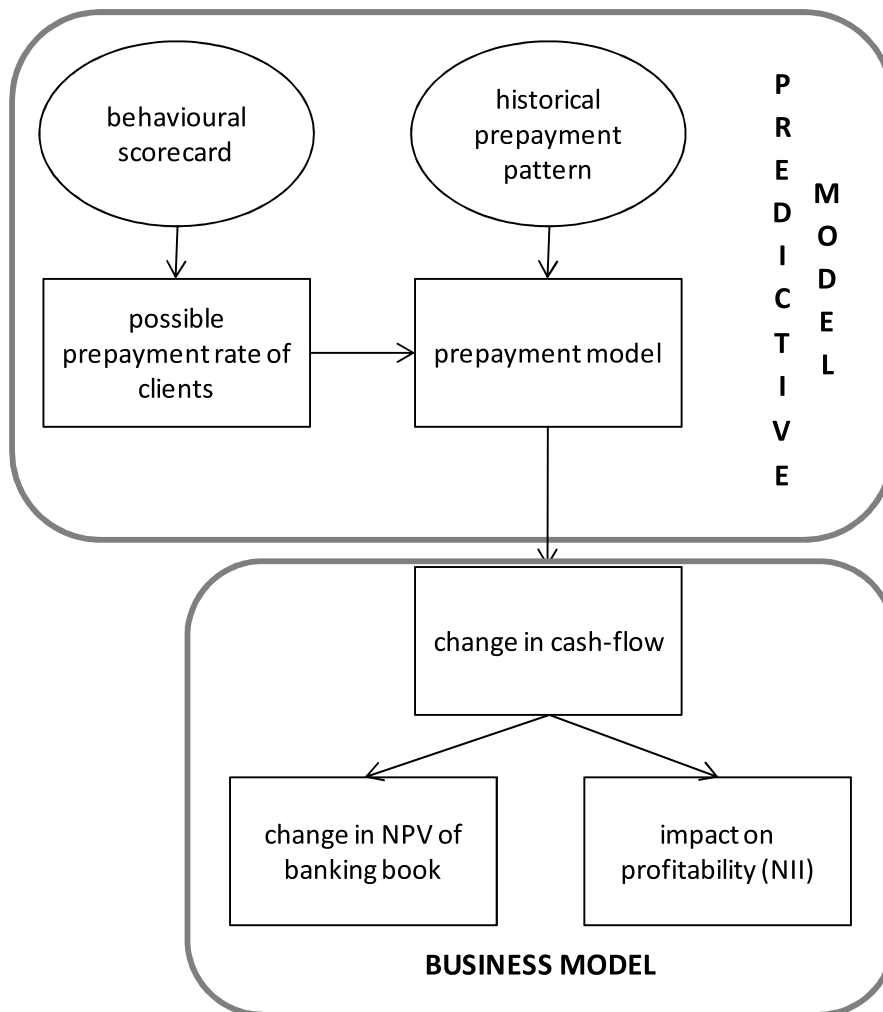
The result of both models is the estimation of change in NPV of banking book, which is ultimately the VaR of the theoretical value of the banking book. This is the possible loss of the banking book value in case of prepayment happens, so this the theoretical value of prepayment option.

There is scarce literature available on the outcome of above mentioned models, since finally banks use them for internal evaluation and pricing of prepayment options which is strictly confidential. In November 2001 a study was published about the mortgage prepayment risk based on UK data. Based on this study we can summarise the possible risk drivers for optimal and non-optimal models as following:

Macro economic drivers
1   interest rate changes
2   house price inflation
3   GDP growth
4   employment level

Loan specific drivers
5   age of loan
6   length of interest rate period
7   loan to value ratio
8   structure of prepayment charges
9   sales channel

Borrower specific drivers
10 age
11 marital status
12 region of living
13 occupation
14 borrower type (first time buyer/exisiting customer/re-mortgage)
15 ratio of loan to salary
16 other products held with the lender

The study was prepared with the involvement of market participants representing about 65% of UK market at the end of 2000. Based on their data the following risk drivers were found the most significant:

- *age of loan*: most borrowers don't want to remortgage or move house again shortly after having gone through the application process of a mortgage, so prepayment rather happens in the second half of the period;
- *house price inflation (HPI)*: when house price inflation was high the number of home moves increased and it resulted in higher activity in housing market with higher prepayment rate;
- *interest rate changes and interest differential*: the interest differential is actually the difference between the current rate of the mortgage and the new rate offered by another lender; it was observed that the higher the interest differential the higher the prepayment activity;
- *prepayment charge*: it was observed that the cost of prepayment over a certain level discouraged prepayment significantly.

For illustrating with example the following chart presents the house price inflation index of FHB Mortgage Bank[3] and the CHF interest rates[4]:

---

[3] FHB Mortgage Bank is publishing house price inflation index since 2009 for the Hungarian market.
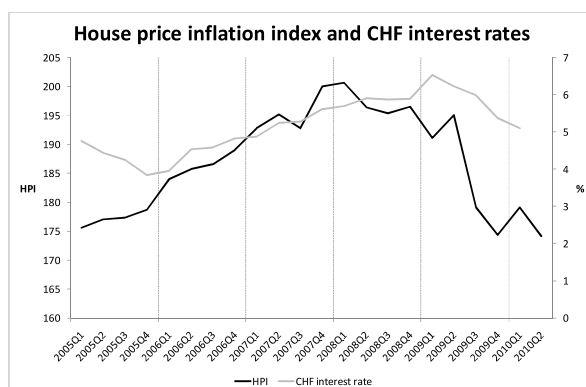
Source: FHB Mortgage Bank, NHB

Based on the chart I assume that prepayment rate was high in 2005-2006, when house prices were increasing and interest rates were slightly decreasing, probably there was a natural early prepayment ratio also in 2007-2008, when HPI and interest rates were increasing parallel, then in 2009-2010, which period was affected by the crisis.

4.3. Illustration with theoretical example

The following theoretical example presents the possible evaluation of prepayment possibility on the NPV of banking book. Let's assume a retail mortgage portfolio with the following characteristics:

	Average	Volatility
Loan volume (M HUF)	6,5	2,5
Interest rate	6,5%	2%
Maturity (yr)	15	5

The portfolio is consisted of 100 annuity loans, out of which 5 loans totally prepay before maturity.
The sample portfolio can be illustrated with the following characteristics considering volume, interest rate and maturity:

---

[4] The specialty of the Hungarian mortgage market is that 90% of loans are CHF-based, as a result the CHF mortgage interest rates influenced the market movements until 2010Q1. Since then FX-based loans are not allowed on the Hungarian retail mortgage markets.

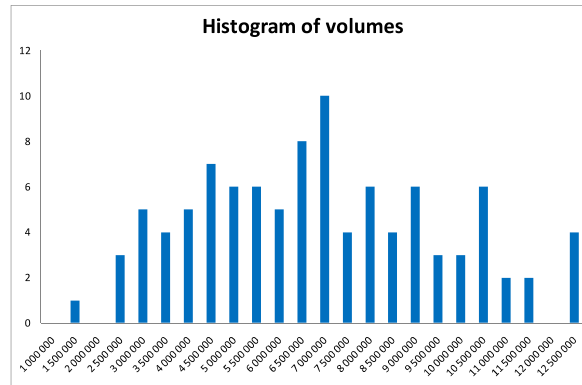**Chart 3.:** Histogram of loan volumes in the theoretical portfolio

Histogram of volumes

Chart 4.: Histogram of loan interest rates in the theoretical portfolio

Histogram of interest rates
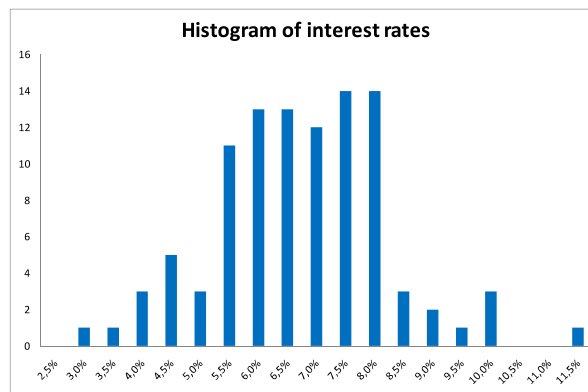
Chart 5.: Histogram of loan maturities in the theoretical portfolio

Histogram of maturities
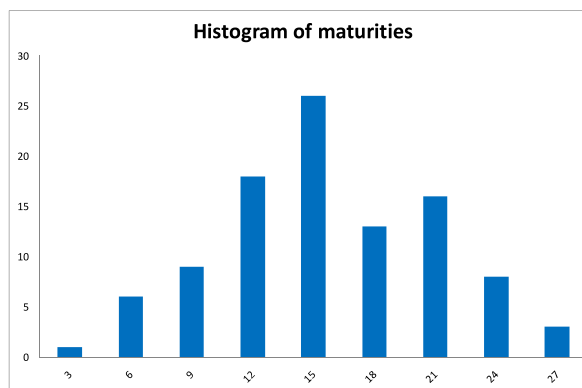
The original cash-flow of the portfolio is calculated as the aggregation of the cash-flow of each loan, and the present value of the portfolio is calculated as

aggregation of the PV of each loan (which currently equals the book value of loans[5]):

| Nr	1	(...)	95	96	97	98	99	100	Total
Loan	6 360 989		2 687 535	9 193 059	6 999 492	4 102 686	6 139 499	3 576 964	
Interest rate	7,67%		5,13%	5,58%	11,17%	6,90%	6,76%	6,05%	
Maturity (yr)	22		18	10	12	6	21	13	
Installment (monthly)	49 954		19 088	100 155	88 467	69 748	45 661	33 172	
PV	6 360 989	(...)	2 687 535	9 193 059	6 999 492	4 102 686	6 139 499	3 576 964	661 882 148
1	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
2	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
3	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
4	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
5	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
119	49 954		19 088	100 155	88 467	0	45 661	33 172	4 813 922
120	49 954		19 088	100 155	88 467	0	45 661	33 172	4 813 922
(...)									

Considering that 5 loans are totally repaid before maturity the above cash-flow table is modified accordingly (e.g. loan nr 96 is repaid totally):

Nr	1	(...)	95	96	97	98	99	100	Total
Loan	6 360 989		2 687 535	9 193 059	6 999 492	4 102 686	6 139 499	3 576 964	
Interest rate	7,67%		5,13%	5,58%	11,17%	6,90%	6,76%	6,05%	
Maturity (yr)	22		18	10	12	6	21	13	
Installment (monthly)	49 954		19 088	100 155	88 467	69 748	45 661	33 172	
PV	6 360 989	(...)	2 687 535	6 043 689	6 999 492	4 102 686	6 139 499	3 576 964	649 133 905
1	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
2	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
3	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
4	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
5	49 954		19 088	100 155	88 467	69 748	45 661	33 172	6 481 833
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
71	49 954		19 088	100 155	88 467	69 748	45 661	33 172	5 847 287
72	49 954		19 088	0	88 467	69 748	45 661	33 172	5 747 132
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
119	49 954		19 088	0	88 467	0	45 661	33 172	4 552 244
120	49 954		19 088	0	88 467	0	45 661	33 172	4 552 244
(...)									

For the calculations we consider that there was no change in the interest rates of the loans. The impact of monthly cash-flow changes on the aggregated cash-flow of the portfolio is presented on Chart 6:

---

[5] I assume that loan pricing is fair, i.e. at disbursement the NPV of a loan is zero.

Chart 6.: Impact of 5 early repaying loan on the monthly cash-flow of the theoretical portfolio
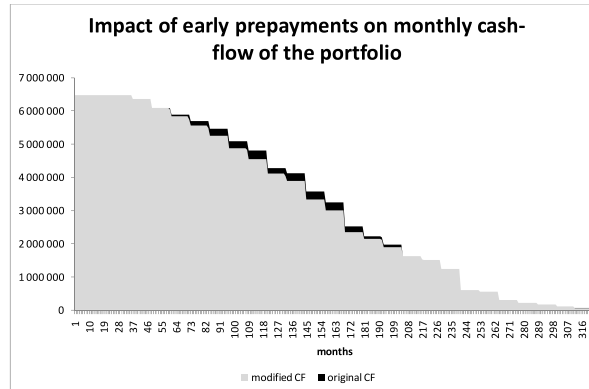
As a result of the early prepayments the bank suffers loss on profitability and NPV side also. Considering profitability the bank loses the interest which is not paid because of the prepayment. If we don't consider time value effect the total loss on the portfolio because of the early repayments is 1.2%, while in terms of PV the loss is 0.9%.

Interest (HUF)	Cash	PV
Original	387 179 282	264 938 640
Modified	382 438 075	262 559 800
Change	-4 741 207	-2 378 840
Change %	**-1,2%**	**-0,9%**

Considering the NPV of the banking book the total PV of the mortgage portfolio cash-flow has decreased by 1.9%:

Total (HUF)	Cash	PV
Original	1 046 997 221	661 907 124
Modified	1 021 787 322	649 180 965
Change	-25 209 900	-12 726 159
Change %	**-2,4%**	**-1,9%**

The decrease in the total portfolio value can be considered as the theoretical price of the prepayment optionality in the mortgage portfolio, and therefore this should be the fee charged on clients for early prepayment for covering the loss of the portfolio value.

References

BIS (2004): Principles for the Management and Supervision of Interest Rate Risk. July.

CEBS (2006): Technical aspects of the management of interest rate risk arising from non-trading activities under the supervisory review process. October.

PERRY S., ROBINSON S., ROWLAND J (2001): *A Study of Mortgage Prepayment Risk*. The Actuarial Profession.

KALFMANN P. (2008): Módszertani lehetőségek a banki könyvi kamatkockázat mérésére [Methodologies for managing interest rate risk in the banking book]. *Hitelintézeti Szemle* pp.20-40.

PARADOX OF PLENTY[1]

LÁSZLÓ SZALAI

1. Curse of the Resources

1.1. Myths about Reality

As on 12th October, 1492 at 2 a.m. a sailor, Rodrigo de Triana, from the crow's-nest of a sailboat called Pinta shouted „¡Tierra! ¡Tierra!", neither his captain Martin Alonso Pinzón, nor his superior, a certain Christopher Columbus, the „Great Admiral of the Ocean" have thought that this moment will irreversibly change the history. Unlike their employer, Isabella I. of Castile and her prominent advisors, who exactly knew that they had been financing a business enterprise a few years ago in order to build a far more powerful and wealthier Iberian Empire than ever seen before. All the circumstances were given; America's unexplored territories were hiding an enormous amount of natural resources, especially precious metals.

During the next century the European armies were defeating the thriving cultures of the New World one after the other. Divers of them were prospering well, but the strongest and richest was the mystic and exotic Aztec Empire, located at the beautiful lands in the middle of the Yucatan-peninsula. In 1519 a Spanish conquistador, Hernán Cortés started heavy offensive to overcome the Aztec troop led by their last Emperor, Montezuma II. Exploiting the local conflicts of interest the invasionary forces had vanquished the natives by sieging and taking over their Capital, Tenochtitlan at Lake Texcoco in 1521. A few years later Cortés returned to Europe as a very wealthy nobleman with a whale of loot robbed from the Aztec treasure. But as the pirate legendary tells, the gold was cursed.

Just before his death Montezuma had sworn revenge, cursing the enemy through his fabulous wealth. Of course those myths are about insatiable hunger and thirst

or everlasting pain and suffering, however it is not steady at all that they are just in the air. Whether the gold and silver inflowing by the dozen from America will bring the age of prosperity to the Spanish Empire?

1.2. Montezuma's Revenge

Until its ultimate defeat in 1588 the Invincible Armada was dominating the oceans of the world while securing the seaways for the merchant fleet, which transported a massive stock of natural resources to the homeland. This hundred years lasting vantage could forecast Spain's leading economic position so far, yet according to estimates the difference disappeared only within a century, moreover it turned to a considerable drawback after passing another hundred years. At the dawn of the XIX. century the Spanish per capita economic output which was just a slightly larger than produced in the Hungarian Kingdom and counted only the half of the English domestic product[2]. Maybe the Gold of Cortés and the (resource) curse stand in the background?

As Mauricio Drelichman and Hans-Joachim Voth show in their article the failure was not caused by the high amount of resources itself, but the sub-optimal utilization through the institutional environment. The explanation of the paradox in this context is that instead of following the path of the capitalizing Low Countries and England the feudal system was conserved in Spain by the enormous stock of precious metals. *„Rent seeking was successfully contained through repeated bargaining between the Crown and the centres of economic activity, to the detriment of the traditional nobility and the clerical establishment."* [Drelichman et al (2008), page 121.] This phenomenon distracted funds from those two groups who gave the axis of the transition in the industrializing countries. Eventually, the institutional system that evolved in Spain was unable to restrict the arbitrary power of the Crown, so the political and economic leverage concentrated alike in the hand of the royal family and the bureaucrats who were unconditional faithful to them. Already during the reign of Isabella the building of a legal system had begun which was able to supersede the progressive nobility from the jurisdiction[3].

Recognizing the possibilities in the trading of silver the Crown founded an empire-wide merchant monopoly, from which the revenues of treasury began to

---

[2] The GDP estimated to 1800 in 2005 international Dollars on purchasing power parity: England: $2717 per capita, Spain: $1443 per capita. Source: Mattias Lindgren, Gapminder, http://www.gapminder.com

[3] Many of these enactments inspired Machiavelli's The Prince

rise in a large measure. Following the laissez-faire principle they allowed indeed the private actors to participate, but precious metal transactions were taxed in 20% rate[4] and could be carried out only in Seville. The structure of the budget was transforming fast also inducing changes in the whole economic framework. The persistent growth in silver reserves eventuated an unbalanced expansion, by 1590 in the receipts of the Crown the share of silver-trade and the taxation connected to it rose up to 29%. However, the economic effects hung over the borders so they indirectly determined the development trend for Europe and the whole known World.

Due to the oversupply of silver the price had dropped in the continent, meanwhile on the other parts of the world – especially in the Far-East well-disposed with condiments and luxurious goods – it held its former value. The European merchants soon recognized the possibility of arbitrage and began to import the exotics paying with cheap Spanish silver for them. With the increasing stock of money in the most of capitalizing countries the general price-level rose, which antedated the symptoms of the further on detailed classical Dutch disease in Spain. The output of the sector based on the primary resource depreciated in such terms, that the revenues from it could not cover the ascendant prices of other tradable goods. Moreover, the restricted trade of silver invigorated the demand for the domestic currency, which led to a serious rise in the exchange rates. The other exportable products of Spain got relative expensive abroad which trough the decaying terms of trade undermined the competitiveness of the rest of its economy. Yet this is not the only mechanism which trough Montezuma's revenge evolves.

As mentioned before, in order to firm its power the Castilian Crown made considerable structural reforms, which were financed obviously from the precious metal revenues. Politics have not changed since then in the point that the path to despotism leads through delusion, machination and corruption. As follows the first institutional manifestation of the resource curse was the utilization of the rents for buying votes and bribing the representatives of different economic and political interests. However, just like the formerly discussed effects of the silver-trade the royal ambitions pay neither regards to the borders.

In 1516 due to the calculating marital politics of the Habsburgs the latter Holy Roman Emperor[5] Charles I. became the king of Spain. During his reign the territories had grown in such terms that the contemporary saying described it as „the Empire on which the sun never sets." At the same time, the controlling and

---

[4] The „Royal fifth"

[5] Under the name Charles V. His other main titles: King of Italy, German King, King of Naples.

defence of the land together with its further expansion required massive financial funds. In order to overcome his money worries Charles claimed for short term loans without collateral security (*asientos*) from the era's two influential banker families, the Fuggers and the Welsers. The Golden Bull, the constitutional document of the Empire however limited him to use the imperial revenues for military purposes and forbade spending it outside the borders. Hence, Charles utilized the formally independent Spanish silver-incomes to redeem the loans and finance his geopolitical designs of great dimensions.

The successor on the throne, Philip II., reported bankruptcy due to the enormous costs of the war fought with the French, so the quondam finance routine broke off with ten years of overdue debt left behind. However the silver flowed in henceforward to Spain made the opportunity for the King to refinance the treasury from the developing stock markets. With the hand of the banker families the Crown released perpetual bonds (*juro*), wherewith thereafter the banks could trade freely. The revenues from the subscription were used to angel even more grandiose military operations or rather to finance the King's current expenses, meanwhile the precious metal import had risen continuously in order to repay the former loans. At the temporary stabilized position the tax incomes were able to cover the interests of the bonds (*situado*), but the series of newly borrowed credits needed among others for the establishment of the Invincible Armada made it necessary again to involve the silver-incomes in the debt service. By 1598 the total value of the traded bonds was seven times larger than the annual revenues of the treasury[6].

---

[6] Drelichman et al (2008)

Figure 1.: Precious metal imports and short term liabilities[7]
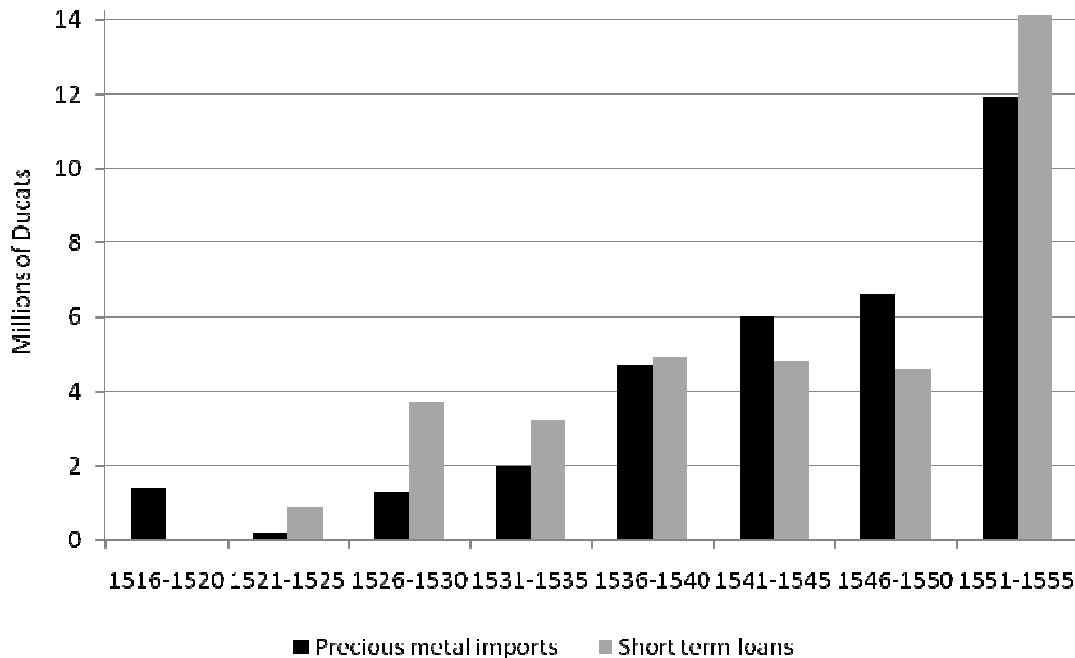
With the leveraged financial mechanism described above the King was able to ignore the *Cortes Generales*, an assembly of Spanish nobleman which functioned as a quasi-parliament and still kept its control over the tax rates. But after a while the permanent wartime expenditures of the Empire could be only covered by increased duties. At the same time the feasible bankruptcy threat with the collapse of the armies and denoted a significant political hazard, so the members of the *Cortes* voted the necessary impositions one after the other. Of course the fact that through the secondary markets of the bonds many of them possessed *juros* in large quantities helped them to get engaged in the support of the Crown's solvency. Hence, the increasing taxes worked as a post-financing mechanism beyond the bonds released by the Crown in its own scope of authority, so exploiting the weak institutions the King could spend above his means being cert aware of posterior approval.

The debt spiral evolved by the irresponsible mismanagement was swelled further through the Thirty Years' War, in which the Crown reported bankruptcy six times to the bondholders. The conflict settled with the Peace of Westphalia eventuated in the ultimate defeat of the Habsburgs together with Spain's

---

[7] The data are in millions of Ducats at 1580 level. Source: Drelichman et al (2008)

abdication of its imperial geopolitical designs. The new international balance of forces even along with the general concentration of political power[8] re-enabled the early capitalistic disintegration of some economic institutions, however the accumulated debt put back the development of the traditional Spanish sectors and the industrial output yet for a long time.

The century-lasting Spanish „golden age" based on the precious metals carried the seeds of its own downfall already at the discovery of America. Along with the economic consequences of the unbalanced expansion, the insalubrious development of the institutions leaded to such a sub-optimal allocation of the resources which condemned the superpower of the XVI. century to be one of the file-closers 200 hundred years later. Cortes' Gold brought curse on the vanquishers of the Aztecs indeed, even if Montezuma's revenge evolved through the paradox of plenty instead of the everlasting pain and suffering as known from the pirate legendary.

1.3. The modern Curst

The World has changed since the Age of Discovery, yet according to numerous empiric observations the Spanish example can still be adopted to draw relevant conclusions in point of the development of countries well-disposed with natural resources. The scope of the utilizable raw materials has flared, although the economic logic related to them remained almost the same. During the XX. century copper, nickel, aluminium and other precious- or rare metals[9] had been falling into lane with gold and silver. On the grounds of motorisation oil and natural gas, in due to the advance of extraction technologies diamond and other gemstones became more and more relevant. The abreast growing markets of the global economy together with its enormous need for energy created a massive demand for resources, so – just like in case of Isabella about 500 hundred years ago – all the circumstances were given for countries endowed well with natural wealth to develop by leaps and bounds. But as always, history repeats itself this time again.

Even in the modern era several wars[10] exploded in order to have possession over the natural resources, moreover the outcomes of these conflicts were also

---

[8] Beginning of the Age of Absolutism

[9] Among others: lithium (mainly used in batteries), uranium (nuclear plants), titanium (armaments industry and medicine)

[10] Some examples: The Nazi occupation of Norway in World War II for the heavy water stocks, the battle for crude oil and natural gas in the Caucasian Mountains. Further on the Gulf War and the American attack on Iraq.

dependent on the distribution of raw materials between the participants. Considering this the natural wealth is the source of economic and political power, a factor which is decisive in the view of a given country's all-time development potential. None the less, the empirical studies regularly verify that in general the countries which are disposed well with resources realize below average economic growth.

In his study[11] Mansoob Murshed rates 42 developing countries of the World into two main categories according to the seriousness of growth failures they have born. Following him the failure is pronounced „*severe*" if the given economy had already exceeded its 1998 per capita income at least once sometime between 1970 and 1990, or „*catastrophic*" if the setback is traceable compared to the performance of the 1960's or earlier periods. He shows that out of the analysed sample only six countries do not possess point-source natural resources in high measures. If we also consider the data which only available from 1965 then further three countries – Angola, Iraq and Ethiopia – could be featured with growth failures. According to Murshed all of these are resource-based economies.

---

[11] Murshed (2004), pages 1-2.

Figure 2.: The relation between primary exports and economic development[12]
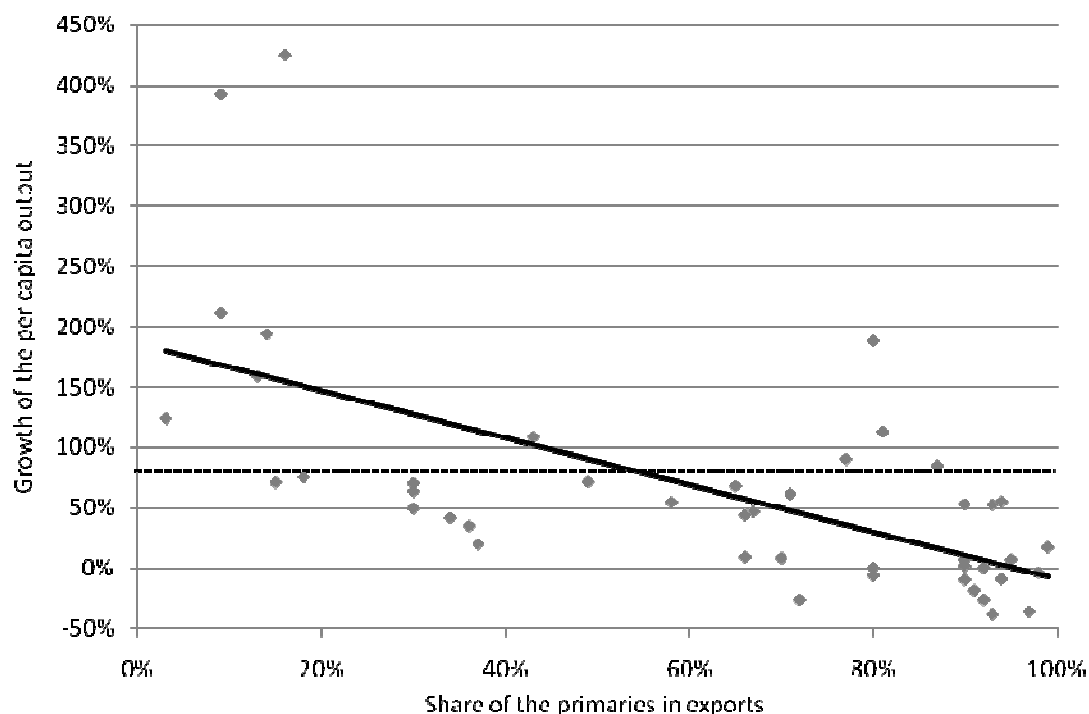
Notwithstanding the significant difference which appears among these countries in the stage of general development and productivity they are common in their 25 years lasting stagnant or decreasing economic output despite of the large natural reserves. If we agree that these resources are sustainable per se, then certain mechanisms have to exist in the background, which are responsible for the failure in case of the curst ones. However considering the numerous sharp exceptions the conditions of these are presumably embedded to a more widespread context.

Botswana, which is well-endowed with diamonds and realized a 188% real growth rate in the above mentioned period, is used to be presented as to most telling argument against the theory of resource curse. „The exception proves the rule" – we could say if we ignore Norway with its high reserves of crude oil and natural gas, or Chile which extracts iron ore and copper in great quantities. Australia is also passing on the same way by its world-leading position in coal exports with a total share of 29% on the global market. Although the relation

---

[12] The data are on purchasing power parity for the period between 1980 and 2005 in 2005 International Dollars. The Figure shows 45 random selected records out of the sample counting 112 countries in which cases the relevant data were available. The horizontal broken line denotes the average of the World (62.01%). Source: Mattias Lindgren, http://www.gapminder.com, NationMaster, http://www.nationmaster.com

between natural wealth and economic development is remarkable[13], on the other hand the cases which differ from the good-fitting examples are significant enough to necessitate further analysis. Important to notice that under exceptions I mean the countries which were able to set long term growth forth even though the high rate of primary exports, not the ones which do not prosper despite of their lower ratios. As Murshed refers on it as well, by the latters it seems less likely that the failure arises essentially from the lack of the resources.

In the following parts of my essay I present the most established theories construing the paradox of plenty, I touch upon the possible causes of exceptions and endeavour to confirm them with empirical evidences. For the analysis I utilize the statistics of 27 resource-based economies, in which case the share of the primaries counts at least 60 % of total exports and the relevant dataset is accessible.

2. Theories about the Paradox

2.1. Dutch Disease

In 1959 a high reserve of hydrocarbons was discovered under the Dutch shelf of the North Sea, which however could not be exploited efficiently through the underdeveloped extraction technologies and unfavourable economic atmosphere until the 1970's. Then the oil crisis in 1973 and the latter energy crisis[14], or rather the international and general boom in the commodity prices together with the advance in mining technologies had created ideal conditions to start the production. Yet, despite of the successful investments and selling at the global market the Netherlands did not show the unambiguous signs of prosperity.

Related to this issue the phrase „*Dutch disease*" appeared first in The Economist published November 26th 1977, wherein Nagasaka defined it as a „*...term used in correlation between natural resources and economic growth, that explains negative influence to a certain sector or the economy itself due to the failure in economic policy after having acquired the windfall profit from increased price of resources.*" [Masaki (2007), page 2.] According to Corden and Neary in the

---

[13] In the sample the share of primary exports explains almost 40% of the variance in growth rate per se.

[14] The first oil crisis was caused by the Yom Kippur War (a.k.a. October War or Ramadan War), wherein the OPEC-states imposed oil embargo to assist Syria and Egypt against Israel and the supporting United States. The second (a.k.a. energy crisis) eventuated in 1979, as Ayatollah Khomeini exploded the Iranian Revolution which generated a widespread panic at the stock markets and pushed the oil prices into heights never seen before.

followings I bring out the main mechanisms of the disease in the framework of small open economies.

Consider a national economy which produces two tradable goods and inherently non-tradable services. Designate the formers as natural resources (R) and manufactured goods (M) and presuppose that their prices are fixed through exogenous factors, while by the latter (S) the price is flexible and orientates to the domestic demand and supply. Take that all the three types of goods are made for final consumption[15] and the real wages are perfectly flexible, so the full employment is guaranteed at any equilibrium. In this simple model I ignore the monetary effects; all of the variables should be measured in real values relative to the international price of the tradable goods. I also presume that all sectors utilize a single specific input while the labour as a mobile factor always moves toward to equalize its returns to all sectors, thus its marginal product value is always equal to the real wage. Under these circumstances I show the effects of a windfall technological boom at the natural resource sector.

On the figure below L_S denotes the demand for labour in the service sector, L_M in the industrial production and L_T in all of the tradable sectors[16]. In the first base allocation of factors among the sectors evolves at the real wage level w_0, whereby M_0 workers will be employed at the industry. At the same time L_S describes the demand in terms of the relative price of the services which gets determined on the commodity market, hence it is necessary to analyse the overall and general state of the equilibrium. On the commodity market this equilibrium comes forward in the point of osculation of the production possibilities curve and highest reachable indifference curve (A), where the gradient of the common tangent assigns the price ratios. Since the price of the tradable goods is exogenous, this ratio simultaneously pegs the real exchange rate of the domestic currency as well.

---

[15] In the case of resources the assumption is equivalent with the production made directly for export. The latter abroad utilization as an intermediate input is irrelevant from domestic aspects.

[16] In this way $L_T = L_M + L_R$ where L_R stands for the demand in the resource sector.

Figure 3.: Effects of the resource boom in the labour market[17]
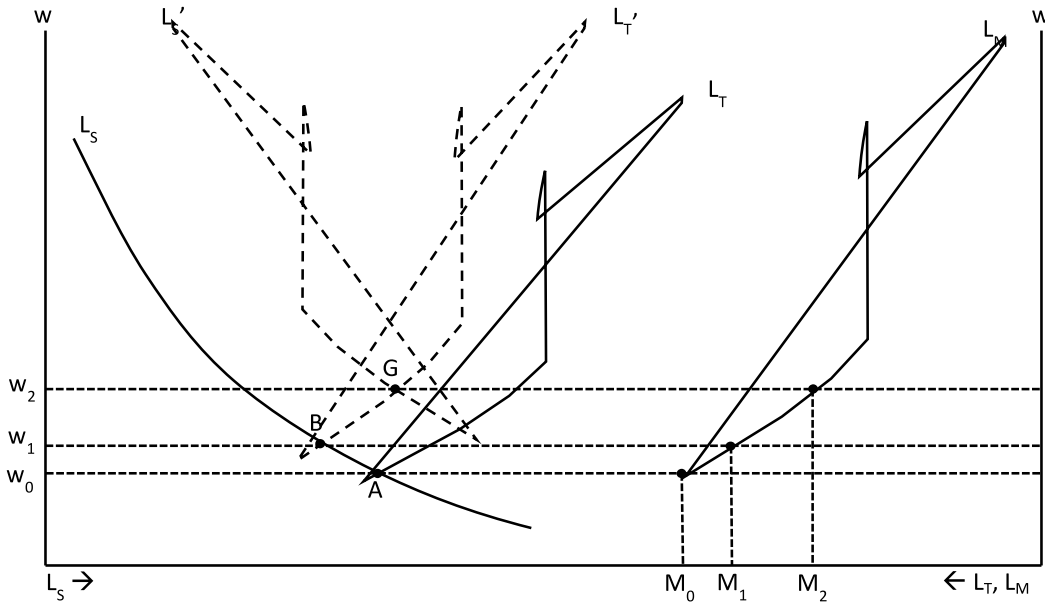
The effects of the positive shock develop in two stages. Foremost let's suppose constant price ratios, thus the L_S curve at the labour market remains unchanged but L_T drifts upward due to the increase in the marginal product value of labour in the tradable sector. In the hope of the higher real wage (w_1) workers start to flow into the resource sector which cuts down the quantity of labour applied in the manufacturing industry. The *resource movement effect*[18] which evolves through this mechanism causes *direct de-industrialisation*, to wit a decline in the manufactured output. On the commodity market this induces unbalanced expansion, the production possibilities curve gets extruded toward the tradable goods while the maximum producible amount of services does not vary. The structure of the output gets altered from the combination matches with point *A* into point *B*, whereby more tradables and fewer services will be produced than in the preceding equilibrium.

---

[17] Source: Own compilation on the grounds of Corden et al (1982), page 7.

[18] Please note that in this theorem „resource" refers on the manpower as a production factor instead of natural resources.

Figure 4.: Effects of the resource boom in the commodity market[19]
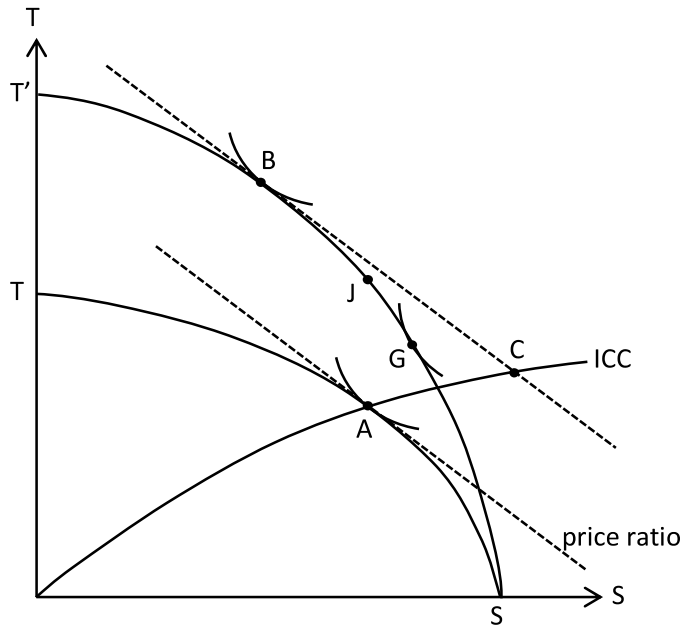
However through the unchanged price ratio the structure of the consumption also shifts toward the ICC curve into the point C which is located outside of the possible production set. But at the same time due to the increasing overall earnings and decreasing output an excess demand evolves in the service sector which sets the second stage of the shock afloat. Through the internal market mechanisms the flexible price of services heightens, therefore the price ratio gets modified perforce too. The marginal production value rises, which causes this time the upshift of the L_S curve at the labour market. The new equilibrium (G) will come up beside the further increase in real wages (w_2) inducing additional fall in the manufacturing employment (M_2) and output. The *spending effect* which arises from the ascendant incomes causes *indirect de-industrialisation*, although its sequels on the output of the service sector are uncertain.

In order to demonstrate the latter statement, consider the extreme cases. Firstly, if the income elasticity of services is zero, then the increase in real wages does not generate additional demand in the sector, thus the spending effect will not show up. Accordingly the ICC curve is represented by a vertical line, the equilibrium will eventuate somewhere between the B and J points and the quantity of services will relapse. Secondly, if by the extraction of resources only the specific factor is utilized without manpower, then the position of L_T curve

---

[19] Source: Own compilation on the grounds of Corden et al (1982), page 7.

remains unchanged, thus the resource movement effect will not develop. In this case – assuming positive income elasticity – more services will be realized due to the flare of the production possibilities. Consequently the parameters of the new equilibrium depend on the labour-intensity of the resource sector together with the income elasticity of services. Notwithstanding that considering the manufactured goods both of the above mentioned effects act in the same direction which induces the obligate direct and indirect fall in the industrial output.

The other main consequence of the model is the ascent in the relative price of the non-tradables due to both the resource movement and the spending effect. If we agree with the law of one price[20] in case of the tradables then the general increase in the price-level ceteris paribus causes the appreciation of the domestic currency in real terms. At the same time this mechanism „*raises the relative price of exports and reduces the relative price of imports.*" [Krugman (2003), page 381.] Accordingly the domestic manufactured goods lose from their competitiveness on the global market which implies a further downfall in the production. On long term the country gets defenceless against the volatility of natural resource prices, thus after the windfall profits the economy will face recession by the first serious decrease in international commodity prices.

As shown above, through the base model of Dutch disease the growth failures of well-disposed countries could be interpreted. However it requires further analysis to reveal whether the short term extra profit in the private sector or the governmental revenues originated mainly from the rents connected to the possession of resources are able to ease the harmful effects. Numerous empirical studies[21] confirm that the answer depends on the given country's institutional framework and on its quality.

2.2. The institutional approach

According to the generally most accepted institutional exposition of the paradox; the incomes from the extraction of natural resources are getting utilized in an inefficient or moreover contra-productive way due to the weak institutions. A typical issue about this phenomenon is *rent seeking* which evolves through the manipulation of the economic and political environment causing sub-optimal

---

[20] On an efficient competitive market which is free from transporting costs and trade restrictions the prices of same goods sold in different countries must be equal in terms of a given currency.

[21] Among others: Boschini et al (2003), Murshed (2004) and Ross (1999)

resource allocation, loss in efficiency, increasing differences in incomes, social conflicts, and finally as a self-inducing process it leads to the further retrogression of institutions. Rent seekers usually endeavour to establish such leverage wherewith they will be able to gain economic profit paying no regard to the market conditions. Thus they utilize their sources unproductively to affect the framework in an artificial or violent way instead of executing investments in order to raise the value added. Due to this behaviour a permanent battle is being fought for the disposition over the natural resources which induces distortion in the political, legal, and economic environment together with rendering more difficult to the „business as usual."

On the other hand, natural wealth itself also favours to the development of *rentier states*, wherein the main source of the governmental revenues is the rent originated from possessing the resources. In spite of the budgetary structures based on the taxes borne by several actors, in these systems the government has to face much less civil control as its receipts are mainly or completely independent from the domestic enterprises and households. The lack of feedback yields the ground to autocracy, thus – just like in the early modern Spain – both the political and economic power are going to get monopolised through the bureaucratic and dictatorial regulation mechanisms. Corruption sprawls among the close elite which induces further efficiency loss on both the economic and institutional fields. Through the weak civil sphere the rentier states usually stagnate at a stage of political development where the intrinsic institutional system of democracy has not evolved yet. At the same time the revenues from the resources create the possibility for the government to sustain the minimal required advocacy with some hypocritical social arrangements.

Robinson and his co-authors presented a mathematical model of these states which shows that politicians „*engage in inefficient redistribution by employing people in the public sector in order to influence the outcomes of elections.*" [Robinson et al (2006), page 466.] Furthermore they found that the over-extraction of resources is prevalent, especially during temporary crises as decision-makers „*discount the future* (value of natural wealth) *by the probability they remain in power.*" [Robinson et al (2006), page 466.] This unproductive squanderer of resources defines such unsustainable growth scenarios which end by necessity – at latest after running out of reserves – in a deep recession. In the early XXI. century the characteristics of rentier states are most typical for the OPEC-members, which had borne an average annual downfall of 1.3% in the per

capita output between 1965 and 1998[22]. In case of these states the above mentioned mechanism is more or less observable.

Thus, supplementing the classic theory of Dutch disease the institutional approach explains the paradox of plenty in wider political and social context. The economic monopolisation and political centralisation evolving through the rent seeking behaviour establish a combination of bureaucracy, despotism, and corruption which enhancing the pure economic consequences leads to a severe or catastrophic growth failure. However if we return to the assumption that natural wealth is valuable per se, then other mechanisms have to exist through which the resources could be extracted in a long term efficient way[23]. In the framework of institutional economics a stable and coherent structure of institutions which developed enough and leans on a strong social approval could guarantee the conditions needed for sustainable growth.

In these systems the structure of authority, the private ownership, the contractual relations, and the legal institutions compose a calculable environment wherein the economic actors are able to optimize for long term. In this context the natural resource endowment is rather a „blessing" than a curse as the windfall profits from the extraction could later materialize in investments increasing the long term value added. It becomes possible to utilize the reserves as intermediate input and turn the service sector or the high-tech manufacturing industry to the major factor of development. Through the protection of intellectual products the temporary surplus also favours to innovation, thus the new technologies and production methods afford a vantage in international competitiveness due to the improvement in efficiency and decrease in costs.

Consequently, according to the institutional analyses the dissolution of the paradox depends on the quality and stability of economic and legal institutions. In countries with weak institutions but severe reserves of natural wealth the resource curse evolves, as by others featuring strong structures a new set of development scenarios become available, through which the harmful effects of the Dutch disease could also be neutralized. Thus by the theory the exceptions mentioned before could be explained as well.

---

[22] Source: Robinson et al (2006), page 448.
[23] If such mechanisms do not exist then we should consider natural resources simply as bads, which contradicts the empirical observations.

2.3. Supplements and further expositions

The theoretic model of Dutch disease by Corden and Neary had several supporters and divers of them have remoulded, modified, or complemented the original version. Following Matsuyama Sachs and Warner[24] assume that the boom in the resource sector induces positive externalities in the form of *learning by doing* in the industry, which however do not prevail by the production of services. The larger the manufacturing sector is relative to all the tradables the stronger this spillover effect is. If the sector is relatively small than the externalities do not cause serious advance in efficiency, so the classic consequences of Dutch disease predominate. In default whereof the spillover effect is able to counteract the loss in competitiveness evolving through the decay in terms of trade and could hold up the de-industrialisation process. Thus the national economies which notwithstanding their natural reserves develop on the base of manufacturing could prevent the growth failures. Although Boschini and her co-authors revealed that the train of thoughts above supposes the monotonic negative disposition of resources, in other words *„the more natural wealth, the worse the outcome."* [Boschini et al (2003), page 6.] According to their standpoint this consequence seems to be inconsistent with the empirical observations.

As analysing the fluctuations in several growth datasets Ploeg and Poelhekke observed that the per capita products of well-disposed nations depend significantly on the volatility of commodity prices[25]. They also revealed that this relation is even stronger if the variables are measured rather in terms of rents connected to resources than on the base of export revenues. Likewise the volatility of the real exchange rate and the terms of trade are also crucial considering the development, since the strong fluctuations induce incalculable circumstances which are ruinous to the economy. This hangs together with the former consequence by Corden and Neary which says that the resource-based economies get defenceless against international mechanisms independent from them.

Beyond the mainstream approach the supplementing institutional theories have in common that they categorise the relevant natural resources among several aspects. Coincidently with others[26] the previously referred Murshed enhances the difference between *diffuse* and *point-source* resources. He considers fossil

---

[24] See in Boschini et al (2003), page 6.
[25] The authors refer to the phenomenon as „volatility curse".
[26] See Auty (1997) or Woolcock et al (2001)

energy sources together with precious metals and stones as point-source because they could be extracted at relative small, exactly delimited localities. Unlike the diffuse types such as soil, woodlands, fish, animals or even sunshine, which range more or less in the whole territory of the given country. Former researches claim that the advent of the resource curse is more likely in case of point-source reserves as they could be acquired, possessed, controlled, and extracted easier. In this way they imply stronger motives for the rent seeking behaviour, which as shown above is one of the most significant source of economic and political conflicts.

Boschini, Petterson and Roine group the resources a bit different as they suggest that what matters is that they call „*the appropriability of the resource,* (which is about) *how easy it is to realize large economic gains within a relatively short period of time from having control over it.*" [Boschini et al (2003), page 2.] The two aspects of the division are the *technical appropriability* which arises from the physical and economical properties of the resource and the *institutional appropriability* which depends on the political and legal conditions of their extraction. In this term a resource is the more appropriable technically as:

    a) the more valuable its given standard unit is

    b) the smaller this unit is considering its volume and weight[27]

    c) the less value added is needed up to final consumption

    d) the easier is to store and transport, and

    e) the less transaction cost is rendered to its realization.

The order could be set up by right of the above seems very similar to the categories range from point-source to diffuse. Thus the most technically appropriable resources are the precious metals and stones followed by rare metals, crude oil, natural gas, copper, iron, bauxite and coal. The least appropriable are coffee, cocoa, sugar, lumber, other agricultural products, fish, animals, and finally soil.

The institutional appropriability of a resource is determined through the role played by its extraction in the development of harmful economic effects under the given social, political, and legal circumstances. Boschini and her co-authors show that in case of less technically appropriable resources the institutional framework seems also less decisive, thus the relation between both the factors and the economic growth is weak. However as the degree of technical appropriability increases the institutions get more relevant as well. By the most technically appropriable resources the institutional appropriability becomes the

---

[27] a) and b) together give the „specific" value of the resource. For example a diamond is more technically appropriable than iron ore as its same unit of volume or weight is far more valuable.

most important factor in the view of long term sustainable economic development. The authors also made empirical research in order to prove their theory and summarized the results in hypotheses[28].

1. Institutional appropriability hypothesis: „*Natural resource abundance is negative for economic development only under poor institutions.*"
2. Technical appropriability hypothesis: „*The impact of institutional quality and abundant natural resources is more pronounced the more technically appropriable are the country's natural resources.*"

Accordingly, beside the „classic" macroeconomic and institutional expositions there are several other theories construing the paradox of plenty. In the following part of my essay I attempt to empirically prove some of them and afterwards on the grounds of the results I identify the major factors which are decisive considering the long term economic development.

3. Empirics

3.1. The quality of institutions

Considering the difficulties related to the quantification of institutional quality factors, their adaptation in statistical models decreases the reliability already by its own. However the theoretical researches assume their decisive role, thus the engagement of these parameters is necessary when constructing the models. In the followings I utilize the dataset assembled by the World Bank (WGI) to describe the institutional frameworks of national economies[29]. The data are aggregated into six clusters[30] which are:

a) Voice and Accountability (*VA*),
b) Political Stability and the Absence of Violence (*PV*),
c) Government Effectiveness (*GE*),
d) Regulatory Quality (*RQ*),
e) Rule of Law (*RL*),
f) Control of Corruption (*CC*).

---

[28] Boschini et al (2003), page 9.
[29] World Governance Indicators (2009). Source: Daniel Kaufmann (Brookings Institution), Aart Kraay (World Bank Development Research Group), Massimo Mastruzzi (World Bank Institute), http://www.govindicators.org
[30] For further information see Kaufman et al (2010) or Paradox of Plenty in Hungarian.

Although the researchers of the World Bank consider the above as the quality indicators of political coordination or governance, but the definition given by them reveals their strong relation with institutional quality. According to this governance is *„the traditions and institutions by which authority in a country is exercised."* [Kaufman et al (2010), page 4.] All the six clusters range between -2.5 and 2.5 where the higher value denotes the better institutional framework.

3.2. The variables and parameters

The following econometric analyses are about to verify the growth failures through the enquiry about resource-based development. Hence by the models I adopt the growth of the GPD per capita index on a period of 25 years as an independent variable[31]. My sample includes the data of 27 countries wherein the share of the primaries within exports totals up to at least 60%. The average growth rate takes 27.42%, the median counts 8.28%, and the standard deviation is 51.87%. The table below summarizes the notations, definitions, and ranges of all the variables utilized by the construction of the regressions.[32]

---

[31] The data are on purchasing power parity in 2005 international Dollars on the base year of 1980.

[32] Please note that the present version of this essay brings only the simplest calculations forward. For further analysis see Paradox of Plenty in Hungarian.

Table 1.: The variables of regressions[33]

Notation	Definition	Range
TechD1	The technical appropriability of the primary resource	Dummy
TechD2	The technical appropriabilities of the main resources	Dummy
TA1	The degree of the primary resource's tech. appropriability	2 − 5
TA2	The degree of the secondary resource's tech. appropriability	3 − 5
TA3	The degree of the third resource's tech. appropriability	1 − 5
AvgTA[34]	Weighted average of technical appropriabilities	2,5 − 4,6
PriES	Share of the primaries within the exports	0,65 − 0,99
GDPpC	GDP per capita on PPP (2005, International Dollars)	538 − 68696
Growth	The growth of GDPpC between 1980 and 2005	-0,39 − 1,88
DAvg	Deflection of Growth from the world average	-1,00 − 1,27
VA	Voice and Accountability	-1,61 − 1,66
PV	Political Stability and the Absence of Violence	-1,60 − 1,23
GE	Government Effectiveness	-1,31 − 1,81
RQ	Regulatory Quality	-2,40 − 1,64
RL	Rule of Law	-1,74 − 1,85
CC	Control of Corruption	-1,35 − 2,10
AvgI	Average of the institutional indexes	-1,67 − 1,70
ToT	The change in the terms of trade between 1980 and 1999	0,20 − 1,19

By building up the linear models I always apply the ordinary least squares method, although I do not necessitate the Gauss-Markov theorem, thus the regressions will not present the best linear unbiased estimations (BLUE) in all cases. I consider the test results significant on 10% level, thus by adopting the p-value approach of type I. error I reject the null hypotheses if p < 0.1. I quantify the reliability of the forecasts provided by the models through the root mean squared error (*RMSE*) together with the adjusted coefficient of determination ($\underline{R}^2$).

---

[33] Data sources: Gapminder, http://www.gapminder.com, NationMaster, http://www.nationmaster.com, and WGI, http://www.govindicators.org

[34] In case of *TA1*, *TA2*, *TA3*, and *AvgTA* variables the higher value denotes the more appropriable resource.

3.3. The Relevance of Institutions

As mentioned before in this section I only present the simplest methods through which the relation between institutional quality and economic growth could be confirmed. Consider first the development related to the composite variable of average institutional quality. The equation of the OLS-regression is the below:

$$Growth = 0{,}336192 + 0{,}345351 * AvgI$$

The estimation is consistent and unbiased but not BLUE. The adjusted coefficient of determination counts 34.46% and the root mean squared error is 40.4%. In this rendition the quality of the institutional framework explains more than a tierce from the deviation in growth among the resource-based economies per se, which could be interpreted as a decisive factor. Consider now the statement by Boschini and her co-authors which says, that as the technical appropriability increases the relevance of institutions gets also more and more significant. In order to show this effect I define a new independent variable (*RI*) as a product of average institutional quality (*AvgI*) and average technical appropriability (*AvgTA*). Through this definition the new variable will have the property that in case of better technical parameters it renders lower values to the poor institutional conditions. The regression gets modified as below:

$$Growth = 0{,}356590 + 0{,}107719 * RI$$

The value of adjusted R-squared has increased up to 38.66%, while the *RMSE* has reduced to 39.08%. Both the significant indexes have changed in favourable direction, thus this latter regression gives definitely better estimation than the former one. At first sight this fact seems to be able to confirm the appropriabilty hypotheses.

Figure 5.: A regression made by the adoption of the technical appropriability hypothesis
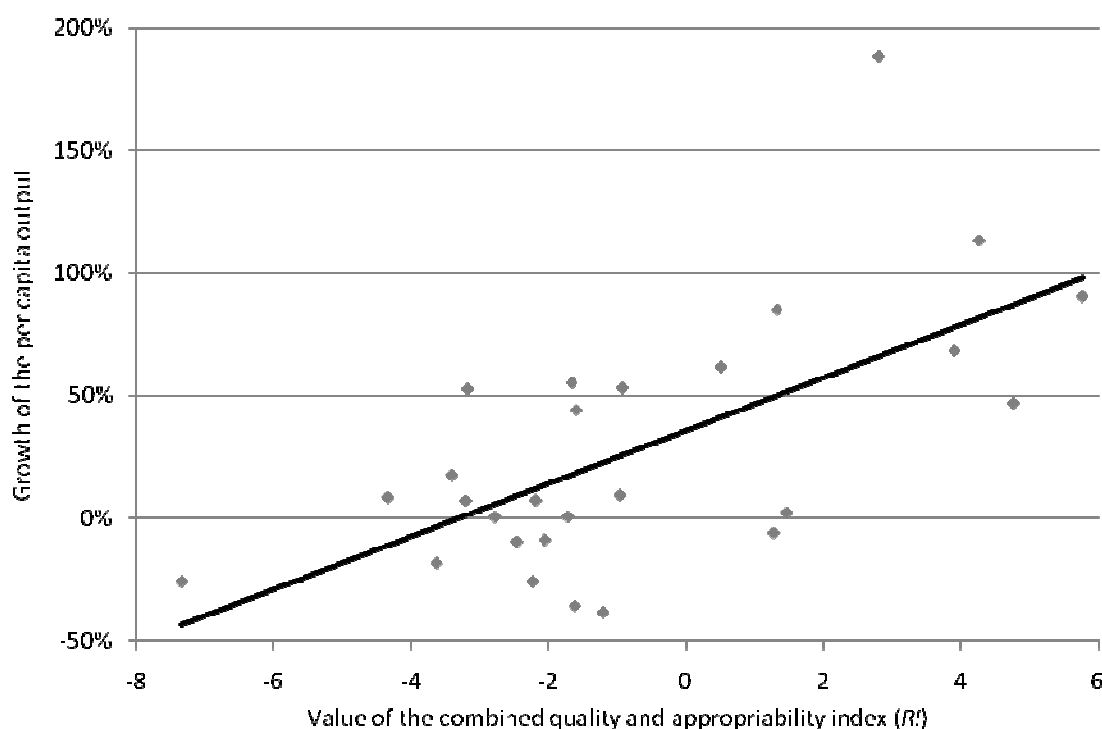
Continuing the above train of thoughts let's examine whether the institutional quality gets more important as the share of primaries grows within the exports. Define the RIE[35] variable as the product of RI and $PriES$, whereby – similarly to the former case – it renders lower estimated values to weak institutions as the appropriability and the share of primaries increase. The newly modified equation is as follows:

$$Growth = 0{,}379248 + 0{,}139757 * RIE$$

This estimation is also consistent and unbiased with an adjusted $\underline{R}^2$ value of 40.46% and an $RMSE$ rate of 38.51%. Both indicators have improved compared to the former regressions, thus based on the results the technical appropriability hypothesis could be not just confirmed but even extended as the follows: *,,The impact of institutional quality and abundant natural resources is more*

---

[35] The abbreviation refers to Resources-Institutions-Exports.

pronounced the more technically appropriable are the country's natural resources" …and the higher the share of primaries is within the exports.

Furthermore there is another way to demonstrate the relation assumed by the original hypothesis. At first divide the sample into three sections equal in their range according to the average institutional quality variable. Thus by the first tierce the institutions are weak, by the second they are average, and the third includes the countries featuring strong institutional background. Then consider the correlation coefficients within each of the groups as shown below:

Table 2.: Correlation among growth and technical appropriability under different institutional frameworks

	Full sample	Weak Institutions	Average Institutions	Strong Institutions
AvgTA	-0,25605373	-0,51402548	-0,1751999	0,75731455
TA1	-0,07695224	-0,38535335	-0,1377283	0,31888641

In this compilation the correlation coefficients of both the appropriability indicators increase monotonic in parallel with the institutional quality. Accordingly the appropriable resources induce worse effect on development under poor political and legal conditions. At the same time it is worthy of note that beside strong institutions the higher degree of appropriability does expressly favour to economic growth[36], although at the average level the relation is rather negative.

3.4. Terms of Trade

In this final section of my empirical analysis I return briefly to the consequences of the Dutch disease. As shown previously the Balassa-Samuelson effect develops due to the unbalanced expansion of the production possibilities after the windfall technological advance, thus the domestic currency gets appreciated in real terms. The process concludes into a decay in terms of trade as domestic exports get more expensive abroad while foreign products turn relative cheap inland. In other words the purchasing power of the domestic currency rises in terms of import goods and at the same time the purchasing power of the foreign currency falls in terms of export goods. If the Balassa-Samuelson effect manifests forsooth, than the terms of trade have to decline as the share of

---

[36] Considering the positive coefficients in the last column.

primaries grows within the exports[37]. In order to test this relation, consider the following regression:

$$ToT = 1,96972 - 1,60993 * PriES$$

The adjusted coefficient of determination is 36.99% beside a mean absolute percentage error (*MAPE*) of 30.14% and a *MRSE* of 20.58%, which together with the negative coefficient of the independent variable seem enough to confirm the theoretical consequence. Although a more robust result can be achieved if the *TechD1* dummy variable gets added to the regression:

$$ToT = 2,12573 + 0,252924 * TechD1 - 1,98757 * PriES$$

In this case the value of adjusted R-squared gives 54.96%, *MRSE* counts about 17%, and the *MAPE* is 23.27%. Considering these results the former conclusions are acceptable.

---

[37] In theory the larger this share is the stronger are the symptoms of the Dutch disease.

Figure 6.: The Balassa-Samuelson effect evolving through the boom in resource sector
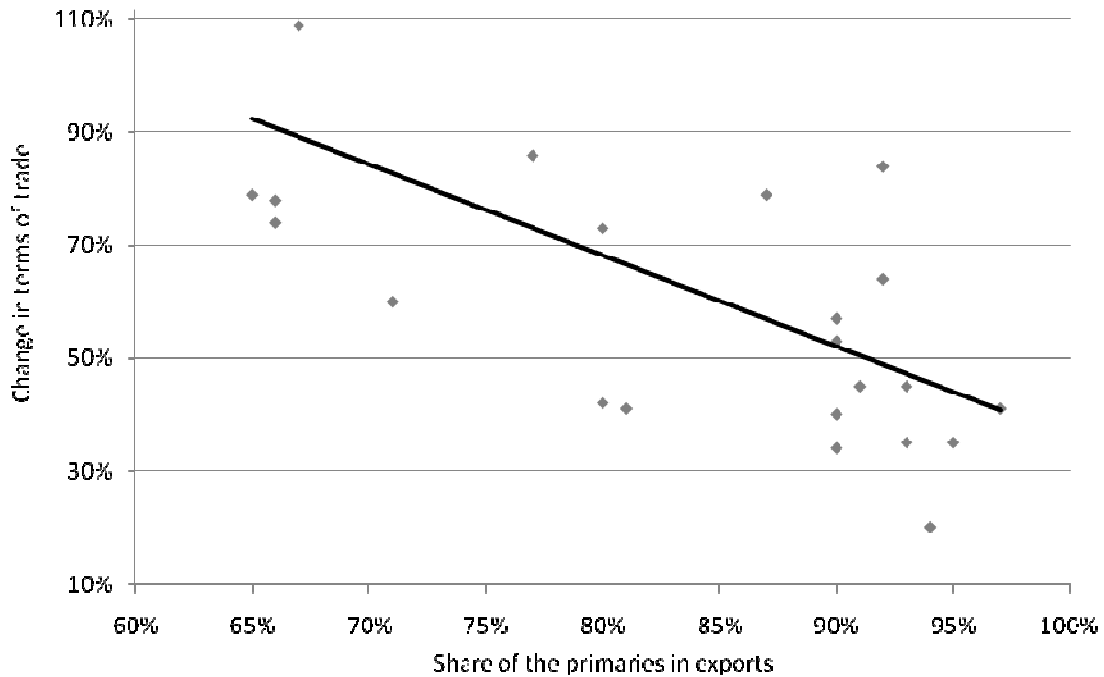
Notwithstanding these unambiguous results the relation between the decay in terms of trade and economic growth seems far less obvious in virtue of the empirical observations. However the correlation among them is significant, if the growth rate is estimated with a constant and the terms of trade index (*ToT*) then the regression explains only 5% from the deviation of the dependent variable, thus the final conclusion of the Dutch disease has to be rejected. Furthermore it is also impossible to assemble a non-restricted model from the available dataset whereby it would not be necessary to omit the *ToT* variable already at very high significance levels during the selections. Accordingly the harmful effects of the real appreciation cannot be proved.

4. Curse and Dissolution

In the present essay I brought forward that the paradox of plenty is a historically supportable and empirically verifiable phenomenon. On the grounds of the pertinent literature I delineated the major macroeconomic and institutional theories in relation with the growth failures of resource-based economies. Afterwards I presented the outlines of my empirical research wherein the main conclusions of the former explanations were tested, especially focused on the appropriability hypotheses.

At my point of view the major issues about the paradox can be summarized as follows:

1. The curse of the resources is not about the fulfilment of the mystic pirate legend but a mechanism which evolves through causality and can be interpreted with economic rationality.
2. The system of causes is discoverable and after the identification of the factors the curse can be lifted as well.
3. In case of a given country the seriousness of the growth failure is not definitely related to the quantity of resource reserves.

The statistical models verified that the statements of institutional economics are correct in the most cases. In the first place it is worthy of note that out of the several possible independent variables the indicators of institutional quality were always significant. The relevance of the physical properties of the natural resources was also confirmed together with their strong connection to the institutional framework. I also brought forward that the technically appropriable resources have dual disposition as they embarrass the development under poor political and legal conditions but favour to it in case of strong institutions.

Through the extension of the fairly simple relation between economic growth and primary export share the former exceptions can also be expounded, although Botswana which represented the far largest residuum in all the estimations still remains a unique case. Maybe the „Diamond of Africa" is the favourite of Montezuma as well…

References

BOSCHINI, A. D., PETTERSON, J., ROINE, J. (2003): Resource Curse or not: A question of appropiability. *Scandinavian Working Papers in Economics*, Vol. 534, Economic Research Institute, Stockholm School of Economics, Sweden.

CORDEN, W. M., NEARY, J. P. (1982): *Booming Sector and De-Industrialisation in a Small Open Economy*. International Institute for Applied Systems Analysis, Austria.

COTRELL, A., LUCCHETTI, R. (2009): *Gretl User's Guide: GNU Regression, Econometrics and Time-series*. Department of Economics, Wake Forest Univerity, USA.

DRELICHMAN, M., VOTH, H. J. (2008): *Institutions and the Resource Curse in Early Modern Spain*. Institutions and Economic Performance (Helpman E.), pages 120-147, Harvard College, USA.

MURSHED, M. S. (2004): When does natural resource abundance lead to resource curse? *Environmental Economics Programme,* Discussion Paper 04-01, International Institute for Environment and Development, United Kingdom.

MASAKI, Y. (2007): *Economic and Econometric Analysis of the Dutch Disease*. Keio Research Center for Foreign Language Education, Japan.

MEHLUM, H., MOENE, K., TORVIK, R. (2006): Institutions and the Resource Curse. *The Economic Journal*, Vol. 106, Royal Economic Society, United Kingdom

KAUFMANN D., KRAAY, A., MASTRUZZI, M. (2010): The Worldwide Governance Indicators, Methodology and Analytical Issues. *Policy Research,* Working Paper, Macroeconomics and Growth Team, Development Research Group, The World Bank, USA

KRUGMAN, P., OBSTFELD, M. (2003): *Nemzetközi gazdaságtan*. Panem, Hungary.

PLOEG, F., POELHEKKE, S. (2008): The Volatility Curse: Revisiting the Paradox of Plenty. *DNB Working Paper*, Vol. 206, De Nederlandsche Bank, The Netherlands.

ROSS, M. L. (1999): The Political Economy of the Resource Curse. *World Politics,* 51, Vol. 2, Institute for International Studies, Princeton University, USA.

RAMANATHAN, R. (2003): *Bevezetés az ökonometriába.* Panem, Hungary.

ROBINSON, J. A., TORVIK, R., VERDIER, T. (2006): Political foundations of the resource curse. *Journal of Development Economics*, Vol. 79, pages 447-468, Elsevier, The Netherlands.

USEEM, J. (2003): The Devil's Excrement. *Fortune Magazine*, Time Inc., USA
http://money.cnn.com/magazines/fortune/fortune_archive/2003/02/03/33 6434/
6:15 pm, 25th October 2010.