Charles University in Prague
Faculty of Social Sciences
Institute of Economic Studies

Three essays on banking and pensions

Dissertation Thesis

Juraj Kopecsni

2010

Contents

List of abbreviations

AIC	Akaike Information Criteria
BASEL II	The New Basel Capital Accord
BCC	Banker, Charnes & Cooper (1984) model
BIS	Bank for International Settlements
CCR	Charnes, Cooper & Rhodes (1978) model
CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
DMU	Decision Making Units
EAD	Exposure At Default
FTE	Full Time Employee
GLM	Generalized Linear Models
IRB	Internal Rating Based approach
KPI	Key Performance Indicator
LGD	Loss Given Default
LR	Loss Rate
MAD	Mean Absolute Deviation
MBO	Manager Business Objective
MSE	Mean Square Error
NII	Net Interest Income
PAYG	Pay-As-You-Go
PD	Probability of Default
PV	Present Value
RAROC	Risk Adjusted Return On Capital
RORAC	Return On Risk Adjusted Capital
QML	Quasi-Maximum Likelihood estimator
SIC	Schwarz Information Criteria
SME	Small and Medium sized Enterprises
SQI	Service Quality Index
SSW	Social Security Wealth
UCO	Universal Client Officer
VRS	Variable Returns to Scale

Foreword

The recent financial crisis has impacted several financial areas. Particularly, it influences (1) savings on pension accounts that are invested on financial markets and are faced with financial risk and risky changes driven by political decisions; (2) already regulated banking sectors through Basel II, where it is essential to identify in advance the key drivers of loss given default of firm sector and to identify the problematic cases from the credit perspective that directly impacts the real economy; and (3) profit-based strategies of the bank branch network by optimizing resource allocation of branch networks and by improving quality of customer services in order to garner the loyalty of existing customers and to fully utilize their possibilities.

The thesis examines the abovementioned areas and consists of three empirical on pensions and banking. It is a collection of essays dealing with different aspects of risk, and it contributes to the recent debates about politically embedded risk in the pension system, credit risk faced by private institutions, and the performance assessment of banks' branches, focusing on quality dimension. The thesis covers partially the area of public finance, credit risk and operational research as well. In all three essays, the most recent modeling techniques are applied: intergenerational and intragenerational impact analyses on a micro level in pensions, generalized linear models with fractional responses and ordinal regressions in loss given default modeling for firm clients in banking, and finally, non-parametric quality adjusted data envelopment analysis in banking, which draws inspiration from the operational research theory.

The first essay, "Policy Risk in Action: Pension Reforms and Social Security Wealth in Hungary, Czech Republic, and Slovakia," provides evidence on the policy risk of social security in Hungary, the Czech Republic, and Slovakia by computing the changes in social security wealth induced by the pension reforms undertaken since the 1990s.

The choice between the pay-as-you-go (PAYG) and fully funded pension system is sometimes expressed in terms of a trade-off between return and risk. The funded system should provide a higher expected return on workers' contributions at the cost of exposing workers to investment risk. Since contributions are invested in stocks and bonds, which yield uncertain returns, workers face uncertainty about the level of their pension when they retire. However, the PAYG systems are not free of risk either. The rules of the pension system may be changed any time, as governments respond to demographic, economic, or political shocks. As a consequence of this so-called policy risk, the contributions actually paid and benefits actually received by a worker may differ substantially from what she was promised by the pension legislation at various moments in her lifetime. Appropriate comparisons between the PAYG system and privately funded systems should therefore involve a comparison of two risky systems.

Analyzing the impact of reforms on workers of different genders, ages, and education levels allows researchers to document the aggregate, intergenerational, and intragenerational aspects of policy risk. Although policy risk has various sources, it always materializes through pension reforms, when past promises are replaced by new ones. The reforms usually involve numerous adjustments to contribution and benefit formulas, which are complicated, not very transparent, and contain a large number of parameters. Such adjustments may affect people of different ages and earnings histories differently, often in ways that may not have been recognized or anticipated by legislators. Pension reforms reduce social security wealth by amounts that sometimes exceed several years' worth of earnings and have large redistributive effects across and within generations.

These findings imply that uncertainties about the redistributive impacts, timing, and political dynamics of reforms contribute significantly to the policy, risk in addition to the inevitable demographic and economic risks. The policy risk of the PAYG system, as documented here, provides a new rationale for a pension system that combines the PAYG and funded pillars. A mixed system in effect follows old investors' recommendation: "Don't put all your eggs in one basket." Finding the optimal balance between the two pillars requires an appropriate quantitative comparison of the risks, one that would characterize the policy risk in a similar way to how stock market risk has traditionally been characterized. Making such a comparison presents a challenge, as the data-generating process driving the changes in SSW induced by pension reforms is fundamentally different from the data-generating process driving stock market fluctuations.

The second essay, "Modeling Bank Loan LGD of Corporate and SME Segments: A Case Study," addresses the estimation of loss given default. Loss given default (LGD) is one of the key parameters used to estimate credit risk in an internal rating based approach considered in the New Basel Capital Accord.

The New Basel Capital Accord was created with the objective of better aligning regulatory capital with the underlying risk in the bank's credit portfolio. It allows banks to compute their regulatory capital as a revised standardized approach based on regulatory ratings for risk weighting assets or using an internal rating based (IRB) approach, where banks are permitted to develop and to employ their own internal risk ratings.

The IRB approach is based on three key parameters used to estimate credit risk: PD – a probability of default of a borrower over a one-year horizon; LGD – loss given default, a credit loss incurred if a counterparty of a bank defaults; and EAD – an exposure at default. These parameters are used to estimate an expected loss. In the advanced IRB approach, all parameters are determined by a bank. Many banks are not ready to fully implement the advanced IRB approach because the advanced approach also requires the modeling and determination of LGD. This chapter proposes a methodology to estimate the loss given default and to find determinants of LGD using a set of firm loan micro-data of an anonymous Czech commercial bank.

In this essay, various statistical models are applied to test empirically the determinants of LGD. We find that LGD is driven primarily by the period of loan origination, relative value of collateral, loan size, and length of business relationship. Different models employed in our analysis provide similar results; in more complex models, log-log models appear to perform better, implying an asymmetric response of the dependent variable. The models with the commonly assumed beta distribution achieved slightly worse results and hence are not deemed optimal for our data. From a policy perspective, the essay provides evidence that workout LGD is a viable option in the credit risk estimation, despite various methodological difficulties. This study attempts to provide a reasonable detail of various issues to be addressed and to propose methodological alternatives to cope with these issues.

The third essay, “Improving Service Performance in Banking using Quality Adjusted Data Envelopment Analysis,” describes the application of data envelopment analysis (DEA) to the performance evaluations of bank branches.

The service economy consists of a large proportion of developing countries’ economic activity, and its growing development has raised the importance of maximizing organizations’ productivity. Organizations are searching for a benchmarking technique to identify best practices in supporting their decisions in order to achieve the effective utilization of resources. To be able evaluate organization’s performance that using multiple resources and providing multiple outcomes it is required non-parametric benchmarking methods such as data envelopment analysis (DEA). Advantage of this non-parametric method is that it does not require specification of the production function form, and it can measure the relative efficiency by simultaneously analyzing their multiple resources with multiple outcomes. Therefore, managers are interested in supporting their decisions through the use of academic methodologies.

It is particularly difficult for service industries to improve productivity and to find substantial cost savings without sacrificing service quality. Many subjective factors affect productivity and service quality. A good example of an industry in which the quality of services is an important issue is the banking sector. In banks, such subjective factors influencing productivity include customers’ needs, behaviors when receiving the service, service providers’ judgment, and skills in providing services.

In this essay, special attention is focused on how to incorporate the quality dimension into branch efficiency. DEA will apply to a set of micro-data from a Czech commercial bank branch network. The goal of the quality-adjusted DEA model is to identify best practice branches that work efficiently and simultaneously provide services with high quality. This model avoids the productivity-quality tradeoff, which is present in the standard DEA model. The quality of services is measured by customer service, mystery shopping and calls, client information index, retention, and client product penetration.

Results show that service quality has a significant impact on branch efficiency, and it should be incorporated into DEA models and operational processes. The essay demonstrates the short-term interaction among service quality and operating branch efficiency. From a policy perspective, the essay provides evidence that there are real reserves for improvement, which can be realized through optimal resource allocations and increasing service quality. This research study discovered that the main factors of efficiency, quality, and productivity-quality tradeoffs are branch size and region, characterized by complex indicator purchasing power. There is documented evidence that larger branches are less efficient than smaller ones. In the future, it will be optimal to open small branches or to redeploy client officers from large, inefficient branches to several smaller and more efficient ones. Moreover, findings indicate that branches in the region with the highest purchasing power are not able to fully utilize their opportunities, which implies lower efficiency.

The first essay “Policy Risk in Action: Pension Reforms and Social Security Wealth in Hungary, Czech Republic and Slovakia” was jointly written with Libor Dusek from CERGE-EI. An early version of this essay was published in working papers series of IES FSV, CERGE-EI at Charles University, Prague and The Pension Institute at Cass Business School City University, London. The final work was published in Czech Journal of Economics and Finance, Vol. 58, No. 7-8, pp. 329-358, 2008. The essay was also presented on the most prestigious international conferences in the area of public finance, for example International Institute of Public Finance in Warwick (2007) and in Maastricht (2008). The data (agent level information about wages, education, gender, age; tax and social security laws, rules and calculation for pension since 1950; mortality tables), for the chapter were collected from several institutions in three different countries. In addition, author carried out parts of the empirical analysis at the Institute of Economics at the Hungarian Academy of Sciences.

The second essay “Modeling Bank Loan LGD of Corporate and SME Segments: A Case study” was jointly written with Radovan Chalupka from IES FSV Charles University. An early version of this essay was published in working papers series of IES FSV, CERGE-EI at Charles University in Prague. The final essay was published in Czech Journal of Economics and Finance, Vol. 59, No. 4, pp. 360-382, 2009. The chapter was also presented on the premier international conferences in the area of quantitative finance, for example Credit Scoring and Credit Control XI in Edinburgh and International Conference Computing in Economics and Finance in Sydney. The data for the essay were collected from an anonymous bank and it was written partially during the author’s stay in London supported by the Fund Mobility of Charles University.

The third essay “Improving Service Performance in Banking using Quality Adjusted Data Envelopment Analysis” is under the revision at the peer-reviewed working papers series of IES FSV Charles University. The data for the chapter were collected from an anonymous bank and it was written partially during the author’s stay in London supported by the Fund Mobility of Charles University.

Acknowledgement

I appreciate the financial support by the Grant Agency of Charles University, grant nos. 318/2006, 131737/2007, by the Grant Agency of the Czech Republic, grant nos. 402/05/0711, 402/05/H510, 402/08/0501, the IES Institutional Research Framework 2005-2010 under no. MSM0021620841 and Fund mobility of Charles University.

I wish to thank to Ondrej Schneider, Zeljko Sevic, Andras Simonovits, Frantisek Turnovec, Libor Dusek, Radovan Chalupka, students at research seminars at the Institute of Economic Studies and conference participants for valuable comments, discussions and suggestions that help to write a good empirical dissertation thesis.

JEL classification: H55, G32, P35
Keywords: social security, policy risk, pension reforms

Policy Risk in Action: Pension Reforms and Social Security Wealth in Hungary, Czech Republic, and Slovakia

Abstract

We provide evidence on the policy risk of social security in Hungary, the Czech Republic, and Slovakia by computing the changes in social security wealth induced by the pension reforms undertaken since the 1990s. Analyzing the impact of reforms on workers of different genders, ages, and education levels allows us to document the aggregate, inter-generational, and intragenerational aspects of the policy risk. Pension reforms reduce social security wealth by amounts that sometimes exceed several years' worth of earnings and have large redistributive effects across and within generations. Our findings imply that uncertainties about the redistributive impacts, timing, and political dynamics of reforms contribute significantly to the policy risk in addition to the inevitable demographic and economic risks.

1. Introduction

The choice between the pay-as-you-go (PAYG) and fully funded pension system is sometimes put in terms of a trade-off between return and risk. The funded system should provide a higher expected return on workers' contributions at the cost of exposing workers to investment risk (Feldstein, 2005a,b), and (Lindbeck, Persson, 2003). Since contributions are invested in stocks and bonds, which yield uncertain returns, workers face uncertainty about the level of their pension when they retire[1].

However, the PAYG systems are not risk-free either. The rules of the pension system may be changed any time as governments respond to demographic, economic, or political shocks. As a consequence of this so-called policy risk, the contributions actually paid and benefits actually received by a worker may differ substantially from what she was promised by the pension legislation at various moments in her lifetime. Appropriate comparisons between the PAYG system and privately funded system should therefore involve a comparison of two risky systems.

We provide a detailed descriptive account of the policy risk of social security in three Central European countries: Hungary, the Czech Republic, and Slovakia. We compute the impact of all major changes in pension legislation adopted since the early 1990s on the social security wealth[2] (SSW) of workers of different ages, genders, and education levels. Altogether these countries undertook ten reforms during the span of 14 years covered. The reforms naturally differed in their breadth, from minor adjustments of several parameters to full-scale reforms introducing a mandatory funded pillar.

An emerging literature has already produced some quantifications of the magnitude of the policy risk. McHale (2001) computes the change in the present value of benefits induced by pension reforms that were implemented in the G7 countries during the 1990s for average workers at age 45 and at the standard retirement age. He finds that some of the reforms reduced the present value of benefits by as much as 29 % (the Italian 1992 reform) or 26 %

[1] Feldstein and Ranguelova (2001), Feldstein, Ranguelova and Samwick (2001), and Poterba et al. (2005) produced quantitative estimates of the distribution of benefits upon retirement in a risky funded scheme, and made expected utility comparisons between the funded and PAYG schemes.
[2] Social security wealth is the expected present value of the future stream of pension benefits minus the expected present value of future contributions.

(the German 1992 reform). McHale's contribution was valuable in demonstrating that cuts in benefits do happen and can be substantial. Shoven and Slavov (2006) compute the internal rates of return from the Social Security in the United States since 1939 for an average, 10th percentile, and 90th percentile worker in 1900–1985 birth cohorts. They find "a considerable variation in the internal rates of return through time for a given birth cohort". They also find substantial differences in IRRs across cohorts. Blake (2008) shows that even private pensions in the United Kingdom have not been completely immune to policy risk, but have been less sensitive than public pensions. Holst (2005) looks at a representative worker in the cohorts that have already retired in the United States and Germany and computes the discrepancy between the SSW that they were promised at age 55 and the SSW that they were promised when they actually retired. He also makes the first attempt to explain the deviation between realized and promised SSW by demographic variables.

Our approach is methodologically similar to McHale (2001) except that our definition of SSW deducts the present value (PV) of contributions from the PV of benefits. In our opinion it is appropriate to deduct the contributions, as they are an important component of the worker's lifetime wealth. For example, a reform that only raises the contributions clearly makes a worker worse off even though the PV of the benefits remains unchanged.

The main contribution of this essay is in providing a more comprehensive picture of the policy risk. The preceding literature generally computes the changes in SSW for a representative worker in selected cohorts. It thus captures only the aggregate component of the policy risk (the risk that a reform makes workers worse off on average) and partly the intergenerational component (the risk that a reform will affect one cohort differently than others). However, most pay-as-you-go systems also redistribute income within cohorts. This introduces an intragenerational component to the policy risk, i.e., the risk that a reform will affect one's income group or gender differently than others. We document both the intergenerational and intragenerational impacts of the pension reforms, as we carry out our analysis separately for men and women with different levels of education and for all pre-retirement cohorts.

Although each of the reforms had unique impacts, our results do allow several generalizations. Most importantly, pension reforms produce large shifts in SSW and as such create substantial uncertainty. In seven of the ten reforms covered, there were some workers whose SSW declined by an amount exceeding the average annual earnings in their country, and in four reforms there were some workers whose SSW declined by more than twice the average annual earnings.

The reforms typically have had largely differential impact across cohorts, genders, and education levels. Seven reforms produced both winners as well as losers. As for the intergenerational redistribution, in four of the reforms older cohorts gained relative to younger ones (or at least lost less). In four other reforms older workers fared worse. While McHale (2001) observes that workers in the G7 countries aged slightly below the retirement age were essentially insulated from cuts in their SSW, this was not generally the case with the reforms studied here. Some reforms (Hungary 1997 and 1998) introduced different rules for different cohorts, and as a consequence the change in SSW between comparable workers in adjacent cohorts differed by as much as 1.6 average annual earnings. In only three reforms were the inter-generational patterns of the changes in SSW broadly consistent with optimal sharing of risk from negative demographic or economic shocks between generations.

The intragenerational component of the policy risk is also significant. As a whole, the reforms tended to be relatively beneficial to richer workers and detrimental to women. None of the reforms simultaneously benefited workers with lower education and hurt workers with higher education. Each country had at least one reform in which workers with university education benefited substantially relative to workers with low education (Hungary 1998, Czech Republic 1996, Slovakia 2004–2005). The Slovak 2004–2005 reform was extreme in this regard – for example, 50-year old men with elementary education lost 1.8 average annual earnings, while equally aged men with university education gained 4.7 average annual earnings. The desire to create a closer link between benefits and earnings is understandable. Nevertheless, the very fact that some older poor workers also experienced large cuts in SSW is troubling, as these workers generally have neither sufficient savings to cushion the cuts in benefits nor enough years of remaining working life to build them up.

Another observation concerns the political dynamics. Both radical reforms that introduced a funded pillar were quickly followed by another reform that mitigated some of its aspects. Should these two cases be generalized, such reversals make SSW more volatile, as one reform breeds yet another reform. On the other hand, they make SSW less volatile as long as the reversal reform brings SSW closer to the level it was at prior to the initial reform. Workers should somewhat discount the rules laid out by the first reform when making plans for the future.

Switching to a mixed system in Hungary and Slovakia did not increase the SSW of almost any workers. This surprising byproduct of our analysis can be attributed to two factors. First, both countries also promised generous PAYG benefits to high--wage workers when they introduced the funded pillar. Second, the returns on savings in the pension funds appear to be low, due to a combination of overly conservative investment strategies and high fees charged by the funds. The policy lesson is that the rules governing pension funds are indeed critical in order to provide workers with a high return on their savings and to keep the administrative costs low. Our calculations indicate that neither Hungary nor Slovakia have set the rules well enough to realize the potential of the funded system.

2. *Methodology and Data*

Social security wealth (SSW) is defined as the difference between the present value of expected future benefits and contributions promised to workers under the current pension legislation. We compute the impact of each reform on the SSW of all cohorts that either were working at the time of the reform or were born but not yet working, and within each cohort we carry out the computation separately for men and women and for representative workers with different levels of education: elementary, lower secondary (apprenticeship), upper secondary (high-school with a school-leaving exam), and college/university.

The SSW for each cohort (*a*) at the time of the reform (*T*) is calculated according to the following formula:

$$SSW(a,T) = -\sum_{t=T}^{R-1}\left[w_{a,t}\,\frac{C_e + C_r}{(1+r)^{t-T}}\,S(t|T)\right] +$$
$$+\sum_{t=R}^{a+100}\left[\frac{B(a,R)}{(1+r)^{t-T}}\prod_{k=R+1}^{t}(1+i_k)\,S(t|T)\right] + \sum_{t=R}^{a+100}\left[\frac{A(a,R)}{(1+r)^{t-T}}\prod_{k=R+1}^{t}(1+j_k)S(t|T)\right]$$

where R is the year of retirement, t is the current year, C_e and C_r are the employee and employer contributions, respectively, to both the PAYG and funded pillars, B is the value of the initial benefit from the PAYG pillar, A is the value of the initial benefit from the funded pillar, r is the discount rate, w is the gross nominal wage, $S(t|T)$ is the probability of surviving until year t conditional on being alive at T, and i and j are the rates at which the benefits from the PAYG and funded pillars are indexed. Calculating SSW involves three basic steps. First, the discounted value of future contributions is calculated from a projected path of wages and the contribution rates specified by the current legislation. Second, the initial benefit is computed according to the formula prescribed in the legislation. Third, the discounted value of benefits is computed using the current indexation rule and a projected path of variables that affect the indexations[3]. To put the results in perspective we normalize the change in SSW by the average annual earnings in the economy in the year of the reform. A change in SSW by -1.0 units hence means that the worker lost SSW equivalent to the annual earnings of the average worker[4].

Computing the social security wealth required a number of assumptions about the wage profiles of workers, the evolution of certain variables in the future, and the returns on savings in pension funds. These assumptions and the data used to construct them are described in detail in *Appendix B*. Our general principle is that we attempt to compute SSW under the legislation as it is written on the books and as actually implemented. For example, if a particular reform was passed with an understanding that some additional changes would be made in the future, we ignore those envisioned but unlegislated changes[5]. Similarly, the rates of return on savings in pension funds are to a large extent affected by the regulation of funds' portfolio choices and performance embedded in the pension legislation. Our assumptions about the future returns are based on the actual portfolio choices and performance of the funds instead of on arbitrary stock and bond market indices, which are indicators of potential, rather than actual, returns.

Our "representative" workers start working at age 20, work without interruption (men) or with an interruption devoted to child care (women) until the standard retirement age, and at each age they earn the wage that is predicted by the wage profile specific to their gender, education level, and calendar year. The wage profiles were estimated for each country from large individual-level datasets. The length of life is probabilistic, and the future taxes and benefits are discounted by the survival probability.

While the above assumptions are natural they inevitably have some limitations. Analyzing only representative workers does not fully characterize the impact of the reforms across the income distribution. Likewise, the assumption that the worker's wages follow a typical wage profile and that the worker is employed without interruption leaves out a part of the intragenerational component of the risk. Two workers with identical lifetime earnings may be affected differently by a particular reform if they differ in their individual wage profiles or working histories. By assuming that workers work until the standard retirement age we do not analyze the impact of changes in early retirement options that were part of some of the reforms and undoubtedly affected the workers who chose to exercise them. Analysis of

---

[3] The formula computes nominal SSW; we also discount all money flows by accumulated inflation to obtain real (as of the time of the reform) SSW.

[4] The level of SSW of a worker who is at the beginning of her working career also indicates the degree of redistribution built into the PAYG system. If it is positive, the system effectively provides a net transfer to the worker, while if it is negative, the system effectively taxes the worker.

[5] The Hungarian 1998 reform is the most significant case; see section 3.2.

such finer impacts of the reforms would be worthwhile but would require detailed data on individual working histories that were not available to us.

We assume that people had perfect foresight about the future evolution of the relevant variables at the time of the reform. That is, the actual wages, inflation, and returns on assets in pension funds until 2005 are used as the expected wages, inflation, and returns at the time of the reform. For 2006 onwards, we assume a 3% growth rate of real wages for all education categories and genders, and a 2% inflation rate[6]. To project the future returns on savings in Hungarian pension funds, we compute their average returns since the time they were established (1998). Slovakia introduced pension funds too recently to infer their historical returns. We therefore set the expected future returns equal to the average historical return on the portfolios that the funds currently hold. The fees charged by the funds are deducted from the gross returns.

3. Results

Below we describe the main outcomes of the reforms as they concern different aspects of the policy risk.[7] The key characteristics of all the reforms are summarized in *Appendix C*. *Tables H.1–10, C.1–2, and S.1–4* present the main result – the change in SSW – separately for men and women of different levels of education and birth cohorts, for all the reforms. We report results averaged over cohorts born during 5-year intervals, due to space limitations[8]. When the pension system has two pillars, separate tables are reported for workers in the PAYG pillar and the mixed pillar. The *figures* illustrate the impact of selected reforms on selected genders and education levels, and also separate the overall impact into a change in contributions and a change in benefits. (Tables and Figures see in *Appendix A*.)

3.1 Aggregate Risk

Pension reforms do produce large shifts in SSW and as such create substantial uncertainty. In seven of the ten reforms covered, there were some workers whose SSW declined by an amount equal to or greater than the average annual earnings in their country, and in four reforms there were some workers whose SSW declined by more than twice the average annual earnings. At the extreme, the Slovak 2004–2005 reform cut the SSW of women with elementary education born between 1955 and 1959 by 4.5 average annual earnings (*Table S.1*). Examples of other reforms with a large negative impact on SSW include the Hungarian 1993 reform (which postponed the eligibility age for women by 5 years and thus cut the SSW of all women (*Table H.1*), the Hungarian 2007 reform (which raised the employer contributions and altered the benefit formula, resulting in a decline in SSW for all workers, in particular by more than 2 average annual earnings for men and women with university education born after 1980 – *Tables H.9 and H.10*), and the Czech 1996 reform (which postponed the eligibility age for men by 2 years and women by 5 years and made the benefit formula more regressive, resulting in a cut in SSW by more than average annual earnings for most workers (*Table C.1*).

---

[6] These are roughly the rates of wage growth and inflation currently experienced by all countries.

[7] Readers who wish to see the results discussed chronologically by each country and reform are referred to an earlier version of the paper (Dusek, Kopecsni, 2008).

[8] Detailed results for individual cohorts are available upon request.

The three countries experienced declines in fertility and improvements in life expectancy during the 1990s. Since defined benefit systems do not have automatic adjustments to such shocks, measures such as higher contributions, postponed eligibility age, lower benefits or less generous indexations of benefits must be explicitly legislated – an inevitable source of aggregate policy risk. Governments should attempt to allocate the risk optimally across generations when they legislate the changes. Optimal risk sharing generally requires (Gordon, Varian, 1988), and (Ball, Mankiw, 2007) that the burden from a negative shock is distributed across all generations, although in absolute terms the younger generations should bear a proportionately greater burden.

Several reforms were adopted with the motive of improving financial sustainability, and their impact on SSW broadly emulates the optimal risk sharing pattern (with certain exceptions). Specifically, the Hungarian 1997 reform reduced the SSW of all men by 0.16–1.66 average annual earnings, and by less for most women (*Table H.2*). The Hungarian 2007 reform, adopted with a clear motivation to cut structural budget deficits, reduced the SSW of men with elementary education by 0.26–0.99 average annual earnings and the SSW of men with university education by 0.61–2.45 average annual earnings. The cuts in SSW were gradually larger for younger cohorts and somewhat smaller in absolute terms for women (*Tables H.9–10*). Likewise, both Czech reforms (adopted in 1996 and 2002–2003) made approximately proportional cuts to the SSW of all workers with more than 10 years to go until retirement, and gradually smaller cuts to workers close to retirement, except that they hurt women more than men (see *Tables C.1* and *C.2* and *Figures C.1–C.3*).

Some features of certain reforms are clearly inconsistent with optimal responses to aggregate shocks. The current retirees are insulated from benefit cuts[9], contrary to the prescriptions of the literature on optimal risk sharing within a social security system (Bohn, 2001), and (Diamond, 1997). There were also reforms that *increased* SSW approximately proportionally for all workers at least within the PAYG pillar (the Hungarian 1999 and 2003 reforms, which cut contribution rates and provided additional benefits, respectively; see *Tables H.5–H.8*). Such reforms make economic sense if they respond to a positive shock. That clearly was not the case of Hungary in 1999 and 2003, as the two reforms were shortly preceded and followed by reforms which had a negative and much larger impact on workers' SSW. The political mechanism hence produces additional risk by promising more generous benefits even in the presence of a negative shock, necessitating an additional reform which cuts SSW by more than what would be necessary to improve the financial balance in a single reform.

Even if reforms always uniformly reduced SSW, people would still be exposed to additional risk about their timing. For example, shocks to fertility arrive gradually and take many years until they explicitly affect the financial balance of the PAYG system. Politicians may procrastinate before they implement the necessary reform. Worker are exposed to the risk of whether the reform is adopted before or after they retire, and how severe the cuts in SSW would be if it is adopted before they retire (postponing the reform longer implying more severe cuts).

[9] The reform legislations never cut benefits to current retirees and the changes in the indexation rules were generally beneficial to them. The Hungarian 1998 reform was the only exception, as it provided for a gradual switch from net wage indexation to Swiss indexation (50 % CPI and 50 % net wage growth). As a consequence, those already retired saw the present value of their benefits cut by as much as 20 %.

3.2 Intergenerational Risk

The reforms have largely differential impacts across cohorts, genders, and education levels. In seven reforms there were both workers whose SSW increased and workers whose SSW fell. Such a pattern is generally inconsistent with efficient allocation of risk.

We observe a very diverse pattern of intergenerational redistribution. In four of the reforms the older cohorts gained relative to the younger ones (or at least lost less). In four other reforms it was the other way round.

Reforms may have a differential impact across cohorts by introducing different rules for different cohorts or by changing the general rules in a way that is relatively more beneficial to some cohorts than others. The Hungarian 1998 reform is an extreme example of the former. Among other adjustments it created different sets of rules for workers who were to retire before 2012 and after 2012. The benefit formula sets the initial benefit as a certain percentage of the worker's average net earnings until 2012 and as a fraction of average gross earnings from 2013. It was also planned that benefits would become taxable at the same time; however, the corresponding change in the income tax code has not been implemented. This rather ambiguous provision creates additional uncertainty over whether benefits will be taxable at all after 2013, and if so, what the income tax rates will be at that time[10].

The impact of the Hungarian 1998 reform on different cohorts is clearly visible in *Figures H.2 and H.3.* Men with university education retiring before 2013 (the 1942–1950 cohorts) saw an increase in the present value of their benefits of approximately 49 % due to faster indexation of income brackets in the benefit formula. The same men born just after 1950 experienced an 80% increase (*Figure H.3*). In SSW terms, university-educated men and women born in the 1950s gained between 3.2 to 3.5 average annual earnings, while those just slightly older gained between 2.4 to 2.8 (*Table H.3*). The differential impact was less pronounced for workers with lower education, since the gap between gross and net earnings is smaller for them (*Figure H.2* shows that men with elementary education born before 1950 experienced a small cut in the PV of benefits, while those born in the early 1950s experienced a small increase).

Another example of a reform feature that reflects intergenerational risk is the Hungarian 1997 reform, which postponed the retirement age for men and women gradually to 62. However, it shifted the retirement age back by 1 year for women born between 1942 and 1944. The reform was clearly beneficial to these "privileged" women, whose SSW rose by as much as 2 average annual earnings (*Figure H.1*). In contrast, the SSW of women born between 1950 and 1954 fell by 0.45 (upper secondary education) and 0.96 (university education) average annual earnings (*Table H.2*).

Similarly, the Czech 1996 reform helped women just before retirement age (by raising their SSW by as much as 1.6 average annual earnings) while hurting all other women (*Table C.1*). This effect was due to a combination of a gradual increase in retirement age, which had only a minor impact on women who were close to retiring, and changes in the benefit formula that turned out to be relatively more beneficial to older women with historically lower earnings.

[10] We do not subtract any income tax when we compute the benefits after 2012, since our goal is to evaluate the impact of the reforms as they were actually legislated.

The Slovak 2004–2005 reform had a particularly strong intergenerational pattern. It gradually increased the retirement age from 55 to 62 years for women[11] and from 60 to 62 years for men. It cut contribution rates and allowed the opt-out of 9 % of the wage into the newly established private pillar. The new benefit formula made the benefit linear in the worker's average earnings over his entire working history since 1994, up to a cap beyond which workers with more than 3 times the average earnings do not receive higher benefits. The reform provided for a transitory period, initially legislated to last until 2006, during which the benefits were in fact regressive in the worker's lifetime earnings but becoming gradually less regressive over time. An additional provision that positively affected younger parents was a 0.5% deduction from contributions for every child aged below 26 as long as the child was studying.

The mix of these adjustments was much more beneficial to younger workers relative to older workers (*Table S.1*). Men with elementary education born after 1975 gained (their SSW increased by 0.35–0.71 average annual earnings), while those born in the 1960s or earlier lost (their SSW fell by 0.8–2.38 average annual earnings, progressively more for older cohorts). Such differences are equally pronounced among women, especially poorer ones. Women with elementary education just before retirement age (the 1950–1954 cohorts) lost as much as 4.29 average annual earnings. Losses of a similar magnitude are observed for all women with elementary or lower secondary education born during 1950–1964, and are gradually smaller for younger cohorts.

The differential treatment of different cohorts is illustrated in *Figures S.1–S.4*. The retired cohorts were affected only by a change in indexations and the PV of their benefits rose by 11–14 %. Within each education level, the PV of benefits changed by an almost equal percentage for almost all working cohorts, while the percentage change in the PV of contributions differs by cohorts – those very close to retirement (1945–1948) saw a large increase (by 20–75 %), while those just at the beginning of their working careers experienced a 32% reduction.

The negative impact of the Slovak 2004–2005 reform on older workers demonstrates how substantial the policy risk is and how uncertain SSW can be even for workers with just a few years to go until retirement. If we think of this reform (unusually radical as it was in many respects) as a realization from a tail of a distribution of possible reforms, one could argue that the policy risk is in some sense greater than the investment risk in the funded pillar. It seems inconceivable that an amount worth four years of earnings would vanish from the accumulated savings in a pension fund during the last few years before retirement. Moreover, large cuts in the SSW of older workers, and especially poor ones, impose higher costs on such workers than cuts of equal magnitude suffered by younger and richer workers. The former usually have neither sufficient savings to cushion the cuts in benefits nor enough years of remaining working life to build them up.

A few other reforms also hurt workers aged slightly below the retirement age. Specifically, the Hungarian 1993 reform reduced the SSW of women of pre-retirement ages by between 0.4 and 1.3 average annual earnings; the 1998 reform was more severe, as it reduced the SSW of both men and women of pre-retirement ages by between 1.1 and 2.5

[11] This is the case for women with two children. For women with no children the eligible age increased from 57 years, for those with 1 child from 56 years, for those with 3–4 children from 54 years and for those with 5 or more children from 53 years.

average annual earnings. The Czech 1996 and Slovak 2006 reforms also hurt workers of pre-retirement ages, although by less than the younger cohorts. McHale's (2001) observation that such workers were essentially insulated from cuts in their SSW does not appear to be a general phenomenon.

To summarize, the intergenerational impacts of many reforms are not consistent with optimal risk sharing and rather indicate that the political process generates additional intergenerational risk.

3.3 Intragenerational Risk

In the funded scheme, shocks to the stock market return raise or reduce the savings held by workers from a certain cohort by the same percentage irrespective of their wages or gender[12]. In contrast, the policy risk of the PAYG scheme contains an intragenerational component, as reforms may not have a uniform impact across income levels and genders. Most reforms have a greater impact (positive or negative) on the SSW of workers with higher education, in part due to a higher absolute level of contributions and benefits. More interestingly, none of the reforms simultaneously benefited workers with lower education and hurt workers with higher education. Each of the countries had at least one reform in which workers with university education benefited substantially (in absolute or at least percentage terms) relative to workers with low education (Hungary 1998, Slovakia 2004–2005, and, to a lesser extent, Czech Republic 1996).

The Slovak 2004–2005 reform was extreme also in this regard, as it completely eliminated the redistribution from rich to poor workers that was explicit in the benefit formula with only a 3-year transition period. The stark difference between the impacts of the reform on workers with different education levels is illustrated in *Figures S.1 and S.2*. The PV of benefits fell by 36 % for almost all working men with elementary education, while it increased by between 61 and 71 % for almost all working men with university education.

Figures S.3 and S.4 depict the differential impact in SSW terms. SSW increased by at least 4 annual average earnings for all male cohorts with university education born in 1947 or later, and it increased by 7.6 annual average earnings for the 1982 cohort (i.e. those just entering the labor market). Young men with elementary education gained comparably little (0.7 average annual earnings)[13].

The peculiar provision of the Hungarian 1998 reform (setting benefits as a fraction of gross earnings instead of net earnings starting after 2012) had a strong redistributive impact as well, since the wedge between gross and net earnings is higher for high-wage workers. The SSW of the affected men with university education rose by 1.95–3.49 average annual earnings (depending on the cohort), while the SSW of men with elementary education changed only slightly, and similar differences are observed for women (*Table H.3*). The Czech 1996 reform made the formula somewhat less regressive and also increased the number of years counted in assessed earnings. SSW declined by approximately the same amount for all workers of the same gender (1 average annual earnings for men, 1.2 for

---

[12] This statement is strictly true only if all members of the cohort always hold the same portfolio.

[13] On the other hand the pre-reform system taxed high-wage workers particularly heavily – an average man with university education who had just started working had SSW of minus 13.1 average annual earnings. The post-reform SSW of men with elementary education who have just started working is still higher than that of men with university education (-2 compared to -5 average annual earnings).

women), implying that the declines were greater in percentage terms for workers with lower education (*Table C.1*).

Five of the reforms had a distinctly differential impact on women compared to men. In three of them, women fared worse than men (Hungary 1993, Czech Republic 1996, and Slovakia 2004–2005). In two of them they fared better (Hungary 2003 and Slovakia 2006), although overall the reforms tended to make women worse off.

The large negative impacts on women in the three reforms were caused by larger postponements in retirement age for women than for men. Specifically, the Hungarian 1993 reform gradually postponed the retirement age for women from 55 to 60 while leaving men unaffected. The Czech 1996 reform gradually postponed the retirement age from 55 to 60 for women but only from 60 to 62 for men, and the Slovak 2004–2005 reform did likewise (from 55 to 62 for women and from 60 to 62 for men). Such changes reduced the implicit redistribution from men to women, as women had been retiring earlier than men before the reform and surviving longer at the same time.

The reforms as a whole tended to make the pension systems actuarially fairer and to reduce redistribution[14]. Such a trend logically followed from the highly egalitarian benefit formulas and very generous retirement ages for women that the three countries inherited from the communist era[15]. Actuarially fairer systems also reduce labor market distortions and tax evasion, a problem that had plagued the revenue side of the Slovak system before the reform. From the policy risk perspective, it is difficult to assess whether such intragenerational effects of the reforms could be anticipated or accommodated by workers. More importantly, the intragenerational effects of potential future reforms are much less predictable now than they could have been in the mid 1990s. Future reforms may be driven more by shifts in the government's preferences on the left-to-right scale than by a systematic desire to eliminate the most egalitarian features of the old system.

3.4 Early Reversals

Two of the reforms studied here were truly radical, as they introduced a fully funded pillar. In both the Hungarian 1998 reform and the Slovak 2004–2005 reform current workers were given the option of switching from pure PAYG to a mixed system, while new entrants to the labor market had to participate in the mixed system[16]. The reform legislations allowed switchers to contribute 8 % to the funded pillar in Hungary (out of the 31% total contribution by both the employee and employer) and 9 % in Slovakia (out of the 18 % total). Plus, the Slovak reform radically changed the benefit formula.

---

[14] We regressed the change in SSW on a measure of redistribution in favor of a particular worker type to formally assess whether the reforms systematically reduced redistribution. Our measure of redistribution was the ratio of the initial pension benefit to the last wage in the PAYG scheme before each reform. The unit of observation was the worker type characterized by gender, age, and education. We ran separate regressions for each country but pooled all reforms and worker types within countries, and as additional control variables we included the dummy variable for each reform, a gender dummy, and the number of years until retirement. The results show a statistically and economically significant negative relationship between the replacement ratio and the change in SSW. In the Hungarian case, workers whose replacement ratio was higher by 0.1 experienced a change in SSW that was 0.277 average annual earnings smaller (the corresponding magnitudes are 0.135 for the Czech Republic and 1.28 for Slovakia). The change in SSW for women was 0.18 average annual earnings smaller than that for men after controlling for other variables in Hungary (0.81 in the Czech Republic and 1.44 in Slovakia). The results are available upon request.

[15] (Müller, 2002)

[16] (Blake, 2008)

The new government in Hungary was opposed to private pensions and just one year later it cancelled the increase in employees' contribution to the private pillar that had been promised by the previous reform and increased the contribution rate to the PAYG pillar for workers in the mixed system. *Figures H.4 and H.5* show the differential impact on the stayers and switchers. While men in both systems experienced a 3% cut in contributions, workers in the PAYG had their benefits unaffected, while cohorts 1951 and younger in the mixed system saw the PV of their benefits decline by 2.6–6.5 % (gradually more for younger cohorts, who, due to longer accumulation of savings, have a greater gap between the benefits from the funded and PAYG pillars).

Likewise in Slovakia, a reform adopted a mere one year later substantially postponed the final date by which benefits were to become strictly linear in earnings – from 2006 to 2014. This had a particularly negative impact on men with university education (*Table S.2*), whose SSW declined by between 0.25 and 0.85 average annual earnings depending on the cohort.

The policy risk from reversals of reforms should be particularly present if a radical reform is pushed through unilaterally by the current government coalition without a broader consensus with the opposition, as was clearly the Hungarian case. Early reversals make SSW more volatile, as one reform breeds yet another reform. On the other hand, they make it less volatile as long as the reversal reform brings SSW closer to the level it was at prior to the initial reform.

3.5 Funded Pillar Appears to Reduce Policy Risk

In the mixed system the two sources of future pensions are subject to different types of risk. Workers' pension wealth is invested in a more diversified "portfolio" than under a pure system, provided the policy shocks in the PAYG pillar are not positively correlated with shocks to returns in the funded pillar. Diversification hence provides a new argument in favor of a mixed system. Policy risk, however, is present also in the funded pillar[17] since politicians may adopt legislation that hurts workers in the mixed system relative to those in the pure PAYG system (the Hungarian 1999 reform being an example). Whether the mixed system is subject to more or less policy risk becomes an empirical question.

Our results indicate that the funded pillar does reduce the overall policy risk. Almost all workers in the mixed system experienced smaller absolute changes in SSW than their counterparts in PAYG systems from the reforms adopted in Hungary and Slovakia after mixed systems were introduced[18]. Most of the subsequent reforms, after all, concerned only the PAYG pillar; therefore, their impacts were less pronounced for workers in the mixed system, since the PAYG pillar constitutes only a fraction of their SSW.

3.6 Relative Attractiveness of the Pure PAYG and Mixed Systems

Both in Hungary and Slovakia, the outcomes of workers who switched to the mixed system at the time of the radical reform reveal a surprising result – most workers should either not gain by switching, or gain only marginally. In the Hungarian 1998 reform, consider the group that supposedly has most to gain from the private pillar, i.e., men with university

[17] (Blake, 2008)
[18] The reader can compare the respective pairs in *Tables H.5–H.10* and *S.3–S.4*.

education at the beginning of their career (1975–1979 cohorts). The change in SSW is essentially the same regardless of whether they switched or stayed in the PAYG (2.28 and 2.27 average annual earnings, respectively). Since older cohorts contribute to the PAYG for a shorter time, they do not accumulate enough savings to compensate for the 25% cut in the PAYG benefit, and so they are relatively even worse off by switching (*Tables H.3 and H.4*).

In the Slovak 2004–2005 case, almost all workers who switched, including young men with university education, gained slightly less or lost more (typically by 0.1 to 0.4 average annual earnings) than their counterparts who stayed in PAYG (*Tables S.1 and S.2*).

There are two causes for this finding. First, high-wage workers have an incentive to switch from a redistributive PAYG system to avoid redistribution. However, both Hungary and Slovakia adjusted the PAYG benefit formulas at the same time in a way that was highly advantageous for high-wage workers, reducing or even eliminating the redistributive element[19]. Second, the actual net returns of the pension funds appear to be too low to make the funded pillar attractive; high fees and very conservative investment strategies are the main culprits.

The Slovak case is illustrative: The fees are regulated – a frontload charge of 1 % of the monthly contribution plus 0.07 % of the average monthly net value of assets. Such fees reduce the accumulated savings after 40 years of constant contributions by about 20 % compared to the idealized world with zero fees. Even though the so-called growth funds[20] are allowed to invest 80 % of their assets in stocks, they actually invest only 20 % because of additional regulation of the funds' performance. While our calculations assume a somewhat higher share (30 %), the resulting projected nominal return of 6.9 % is not sufficient to make switching to the mixed system attractive. The growth funds would have to invest 50 % in stocks in order to achieve the 8.1% return required to make young men with university education indifferent between staying and switching.

The relative unattractiveness of the funded pillar was hardly an intended outcome of the reforms. It rather appears to be a byproduct of a desire to radically reduce redistribution within the PAYG pillar and poor implementation of the funded pillar.

The fact that more than 50 % of eligible workers in both Hungary and Slovakia switched appears at odds with our result. One possible explanation is that they may anticipate future improvements in the net returns of the pension funds. As the funds' costs do not rise proportionately with assets, the average administrative costs will fall as workers build up their savings. Competitive pressures would then reduce fees and improve returns. The second explanation is based on portfolio diversification, i.e., workers shift some of their contributions to the funded pillar precisely to diversify away from the policy risk in the PAYG. Last, workers may realize that the current PAYG is unsustainable and therefore

[19] If the PAYG benefits in Hungary do become taxable after 2013, the gains to switching to the mixed system relative to staying in the PAYG system will be more favorable than our computations suggest.

[20] Pension fund administrators have to offer three types of funds differentiated by their risk and expected return – growth, balanced, and conservative funds. Conservative funds may invest only in bonds and money market instruments and must be secured against currency risk. Balanced funds must invest at least 50 % of their assets in bonds and money market instruments and at most 50 % in stocks. Growth funds may invest at most 80 % of their assets in stocks and at most 80 % of their investments may be left un-secured against currency risk. The worker's choice of type of pension fund is regulated in order to prevent a significant loss as the worker approaches the retirement age – among other rules, workers with less than 7 years until retirement may invest in conservative funds only.

anticipate that their PAYG benefits will be less generous than they are promised by the current legislation.

4. Conclusions

We documented the policy risk of social security by computing the changes in benefits, contributions, and social security wealth induced by pension reforms in three transition countries. Although the policy risk has various sources, it always materializes through pension reforms, when past promises are replaced by new ones. The reforms usually involve numerous adjustments to contribution and benefit formulas, which are complicated, not very transparent, and contain a large number of parameters. Such adjustments may affect people of different ages and earnings histories differently, often in ways that may not have been recognized or anticipated by the legislators.

Our findings confirm that the policy risk is real and can be substantial. We also show that the PAYG system exposes workers to aggregate as well as intergenerational and intragenerational risk. The policy risk of the PAYG system as documented here provides a new rationale for a pension system that combines the PAYG and funded pillars. A mixed system in effect follows the old investors' recommendation: "Don't put all your eggs in one basket." Finding the optimal balance between the two pillars requires an appropriate quantitative comparison of the risks, one that would characterize the policy risk in a similar way to how stock market risk has traditionally been characterized. Making such a comparison represents a challenge, as the data-generating process driving the changes in SSW induced by pension reforms is fundamentally different from the data-generating process driving stock market fluctuations. Our work may be regarded as the necessary first step towards making such a comparison.

APPENDIX A
TABLES AND FIGURES

TABLE H.1 Reform 1993, pay-as-you-go, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elementary	Lower	Upper	University	Elementary	Lower	Upper	University
1935–39	-0.01	-0.03	-0.19	-0.23				
1940–44	0.11	0.07	-0.03	-0.16	-0.40	-0.53	-0.90	-1.25
1945–49	0.27	0.25	0.15	0.09	-0.45	-0.46	-0.79	-2.06
1950–54	0.23	0.24	0.11	0.10	-0.45	-0.50	-0.83	-2.15
1955–59	0.19	0.19	0.06	0.06	-0.43	-0.52	-0.87	-2.07
1960–64	0.16	0.16	0.00	0.02	-0.41	-0.53	-0.88	-2.02
1965–69	0.13	0.12	-0.06	-0.01	-0.40	-0.54	-0.89	-1.95
1970–74	0.09	0.08	-0.08	0.03	-0.39	-0.53	-0.85	-1.80
1975–79	0.06	0.05	-0.09	0.04	-0.38	-0.52	-0.82	-1.67
1980–84	0.04	0.02	-0.11	0.02	-0.39	-0.52	-0.81	-1.57
1985–89	0.02	0.00	-0.12	0.01	-0.39	-0.51	-0.79	-1.50
1990–94	0.01	-0.01	-0.12	0.00	-0.38	-0.50	-0.76	-1.44

TABLE H.2 Reform 1997, pay-as-you-go, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elementary	Lower	Upper	University	Elementary	Lower	Upper	University
1935–39	0.04	0.02	-0.07	-0.34				
1940–44	-0.69	-0.84	-1.18	-1.66	0.63	0.74	0.99	1.00
1945–49	-0.49	-0.53	-0.70	-1.09	-0.09	-0.11	-0.17	-0.59
1950–54	-0.42	-0.45	-0.59	-0.92	-0.31	-0.34	-0.45	-0.96
1955–59	-0.35	-0.38	-0.50	-0.77	-0.27	-0.29	-0.39	-0.86
1960–64	-0.31	-0.34	-0.44	-0.67	-0.24	-0.26	-0.34	-0.77
1965–69	-0.27	-0.29	-0.38	-0.58	-0.20	-0.23	-0.30	-0.69
1970–74	-0.23	-0.25	-0.34	-0.51	-0.18	-0.20	-0.26	-0.63
1975–79	-0.20	-0.22	-0.30	-0.46	-0.15	-0.17	-0.23	-0.57
1980–84	-0.18	-0.20	-0.28	-0.43	-0.14	-0.16	-0.21	-0.54
1985–89	-0.17	-0.19	-0.26	-0.41	-0.13	-0.15	-0.20	-0.51
1990–94	-0.16	-0.18	-0.25	-0.38	-0.12	-0.14	-0.19	-0.48

TABLE H.3 Reform 1998, pay-as-you-go, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elementary	Lower	Upper	University	Elementary	Lower	Upper	University
1935–39	-1.13	-1.21	-1.36	-1.46				
1940–44	-0.38	-0.16	0.69	2.08	-1.74	-1.92	-2.49	-1.91
1945–49	-0.05	0.21	1.06	2.72	-0.57	-0.37	0.57	2.55
1950–54	0.07	0.37	1.36	3.49	-0.47	-0.25	0.78	3.46
1955–59	0.04	0.31	1.17	3.24	-0.44	-0.26	0.65	3.27
1960–64	-0.03	0.20	0.88	2.77	-0.42	-0.29	0.44	2.78
1965–69	-0.08	0.10	0.65	2.44	-0.39	-0.29	0.31	2.48
1970–74	-0.12	0.03	0.54	2.37	-0.16	-0.08	0.54	2.84
1975–79	-0.12	0.02	0.48	2.27	-0.13	-0.06	0.50	2.95
1980–84	-0.08	0.03	0.45	2.14	-0.08	-0.01	0.48	2.78
1985–89	-0.04	0.06	0.45	2.04	-0.01	0.05	0.50	2.66
1990–94	0.00	0.09	0.44	1.95	0.05	0.10	0.52	2.55

TABLE H.4 Reform 1998, mixed system, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elementary	Lower	Upper	University	Elementary	Lower	Upper	University
1951–54	-0.57	-0.32	0.51	2.41	-1.24	-1.09	-0.27	2.17
1955–59	-0.47	-0.26	0.46	2.29	-1.04	-0.93	-0.21	2.16
1960–64	-0.38	-0.19	0.42	2.21	-0.85	-0.75	-0.13	2.13
1965–69	-0.31	-0.14	0.39	2.17	-0.66	-0.57	-0.02	2.18
1970–74	-0.23	-0.08	0.42	2.27	-0.33	-0.25	0.34	2.68
1975–79	-0.14	0.00	0.46	2.28	-0.21	-0.13	0.41	2.89
1980–84	-0.07	0.05	0.47	2.17	-0.14	-0.06	0.42	2.76
1985–89	-0.03	0.08	0.47	2.06	-0.06	0.00	0.44	2.64
1990–94	0.00	0.10	0.46	1.97	0.00	0.05	0.46	2.53

TABLE H.5 Reform 1999, pay-as-you-go, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elementary	Lower	Upper	University	Elementary	Lower	Upper	University
1935–39	0.00	0.00	0.00	0.01				
1940–44	0.03	0.04	0.06	0.09	0.01	0.01	0.01	0.02
1945–49	0.07	0.08	0.11	0.17	0.06	0.07	0.09	0.15
1950-54	0.10	0.11	0.16	0.26	0.09	0.10	0.14	0.24
1955–59	0.13	0.15	0.20	0.34	0.12	0.13	0.18	0.31
1960–64	0.16	0.18	0.24	0.42	0.14	0.16	0.21	0.37
1965–69	0.19	0.21	0.28	0.50	0.17	0.18	0.25	0.42
1970–74	0.21	0.24	0.31	0.56	0.18	0.20	0.27	0.46
1975–79	0.23	0.27	0.34	0.59	0.20	0.22	0.29	0.48
1980–84	0.24	0.27	0.34	0.57	0.21	0.22	0.29	0.48
1985–89	0.22	0.25	0.32	0.54	0.20	0.21	0.28	0.46
1990–94	0.21	0.24	0.31	0.51	0.19	0.20	0.26	0.43

TABLE H.6 Reform 1999, mixed system, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elementary	Lower	Upper	University	Elementary	Lower	Upper	University
1951–54	0.00	0.00	0.00	0.01	-0.03	-0.04	-0.05	-0.09
1955–59	0.01	0.02	0.02	0.04	-0.03	-0.03	-0.05	-0.09
1960–64	0.03	0.03	0.04	0.07	-0.03	-0.03	-0.04	-0.08
1965–69	0.04	0.05	0.06	0.11	-0.03	-0.03	-0.04	-0.07
1970–74	0.05	0.06	0.08	0.14	-0.02	-0.02	-0.03	-0.06
1975–79	0.07	0.07	0.09	0.16	-0.01	-0.02	-0.02	-0.05
1980–84	0.07	0.08	0.09	0.16	0.00	-0.01	-0.01	-0.05
1985–89	0.07	0.07	0.09	0.15	0.00	0.00	-0.01	-0.04
1990–94	0.06	0.07	0.09	0.14	0.00	0.00	-0.01	-0.04

TABLE H.7 Reform 2003, pay-as-you-go, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elementary	Lower	Upper	University	Elementary	Lower	Upper	University
1935–39	0.24	0.26	0.33	0.40				
1940–44	0.37	0.41	0.53	0.69	0.31	0.36	0.47	0.57
1945–49	0.36	0.40	0.50	0.68	0.42	0.47	0.61	0.81
1950-54	0.29	0.33	0.43	0.62	0.35	0.39	0.52	0.76
1955–59	0.24	0.27	0.35	0.52	0.30	0.33	0.44	0.66
1960–64	0.20	0.22	0.28	0.41	0.27	0.29	0.38	0.56
1965–69	0.16	0.18	0.23	0.33	0.23	0.25	0.32	0.48
1970–74	0.13	0.15	0.19	0.29	0.20	0.22	0.29	0.44
1975–79	0.10	0.12	0.15	0.24	0.18	0.19	0.26	0.42
1980–84	0.08	0.10	0.12	0.20	0.16	0.17	0.23	0.38
1985–89	0.08	0.09	0.11	0.19	0.15	0.16	0.21	0.36
1990–94	0.07	0.08	0.11	0.18	0.14	0.15	0.20	0.34

TABLE H.8 Reform 2003, mixed system, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elemen-tary	Lower	Upper	Univer-sity	Elemen-tary	Lower	Upper	Univer-sity
1950-54	0.24	0.27	0.36	0.53	0.29	0.32	0.43	0.66
1955–59	0.21	0.23	0.31	0.46	0.26	0.29	0.39	0.60
1960–64	0.18	0.20	0.26	0.39	0.24	0.26	0.35	0.54
1965–69	0.15	0.17	0.22	0.34	0.23	0.24	0.32	0.50
1970–74	0.13	0.15	0.19	0.31	0.21	0.22	0.30	0.48
1975–79	0.12	0.13	0.17	0.28	0.19	0.21	0.28	0.46
1980–84	0.10	0.12	0.15	0.25	0.18	0.19	0.26	0.44
1985–89	0.09	0.11	0.14	0.24	0.17	0.18	0.24	0.41
1990–94	0.09	0.10	0.13	0.23	0.16	0.17	0.23	0.39

TABLE H.9 Reform 2007, pay-as-you-go, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elemen-tary	Lower	Upper	Univer-sity	Elemen-tary	Lower	Upper	Univer-sity
1935–39	0.00	0.00	0.00	0.00				
1940–44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1945–49	-0.26	-0.29	-0.37	-0.66	-0.21	-0.24	-0.30	-0.47
1950–54	-0.25	-0.28	-0.38	-0.61	-0.24	-0.26	-0.35	-0.63
1955–59	-0.33	-0.38	-0.53	-0.85	-0.31	-0.34	-0.47	-0.86
1960–64	-0.46	-0.53	-0.72	-1.22	-0.43	-0.47	-0.64	-1.16
1965–69	-0.58	-0.67	-0.90	-1.58	-0.54	-0.58	-0.80	-1.44
1970–74	-0.70	-0.81	-1.07	-1.91	-0.64	-0.69	-0.95	-1.69
1975–79	-0.81	-0.94	-1.22	-2.19	-0.74	-0.79	-1.08	-1.89
1980–84	-0.92	-1.04	-1.35	-2.40	-0.80	-0.85	-1.15	-2.00
1985–89	-0.99	-1.12	-1.45	-2.45	-0.86	-0.92	-1.23	-2.06
1990–94	-0.97	-1.09	-1.40	-2.33	-0.84	-0.90	-1.19	-1.96

TABLE H.10 Reform 2007, mixed system, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elemen-tary	Lower	Upper	Univer-sity	Elemen-tary	Lower	Upper	Univer-sity
1951–54	-0.21	-0.24	-0.34	-0.52	-0.20	-0.22	-0.30	-0.55
1955–59	-0.33	-0.38	-0.53	-0.85	-0.31	-0.34	-0.47	-0.86
1960–64	-0.46	-0.53	-0.72	-1.22	-0.43	-0.47	-0.64	-1.16
1965–69	-0.58	-0.67	-0.90	-1.58	-0.54	-0.58	-0.80	-1.44
1970–74	-0.70	-0.81	-1.07	-1.91	-0.64	-0.69	-0.95	-1.69
1975–79	-0.81	-0.94	-1.22	-2.19	-0.74	-0.79	-1.08	-1.89
1980–84	-0.92	-1.04	-1.35	-2.40	-0.80	-0.85	-1.15	-2.00
1985–89	-0.99	-1.12	-1.45	-2.45	-0.86	-0.92	-1.23	-2.06
1990–94	-0.97	-1.09	-1.40	-2.33	-0.84	-0.90	-1.19	-1.96

TABLE C.1 Reform 1996, Change in SSW as a fraction of annual average earnings

	Men				Women			
Cohort	**Elemen-tary**	**Lower**	**Upper**	**Univer-sity**	**Elemen-tary**	**Lower**	**Upper**	**Univer-sity**
1936–39	-0.09	-0.14	-0.32	-0.33				
1940–44	-0.91	-0.90	-1.06	-0.98	1.60	1.40	1.01	0.75
1945–49	-1.16	-1.10	-1.27	-1.23	-0.73	-0.81	-0.76	-0.87
1950–54	-1.16	-1.12	-1.27	-1.25	-1.05	-1.22	-1.24	-1.38
1955–59	-1.16	-1.12	-1.24	-1.21	-1.11	-1.28	-1.28	-1.40
1960–64	-1.16	-1.14	-1.21	-1.15	-1.14	-1.30	-1.25	-1.37
1965–69	-1.15	-1.14	-1.19	-1.11	-1.18	-1.32	-1.23	-1.35
1970–74	-1.12	-1.11	-1.15	-1.07	-1.17	-1.30	-1.20	-1.32
1975–79	-1.08	-1.07	-1.11	-1.02	-1.14	-1.26	-1.16	-1.27
1980–84	-1.05	-1.03	-1.08	-1.00	-1.13	-1.24	-1.14	-1.25
1985–89	-1.01	-0.99	-1.04	-0.96	-1.09	-1.20	-1.10	-1.21
1990–94	-0.96	-0.95	-1.00	-0.91	-1.06	-1.16	-1.05	-1.16
1995–96	-0.93	-0.92	-0.96	-0.89	-1.05	-1.15	-1.05	-1.15

TABLE C.2 Reform 2002–03, Change in SSW as a fraction of annual average earnings

	Men				Women			
Cohort	**Elemen-tary**	**Lower**	**Upper**	**Univer-sity**	**Elemen-tary**	**Lower**	**Upper**	**Univer-sity**
1941–44	0.01	0.01	0.01	0.01				
1945–49	-0.13	-0.14	-0.21	-0.31	-0.02	-0.02	-0.02	-0.05
1950–54	-0.48	-0.53	-0.65	-0.93	-0.61	-0.64	-0.78	-1.04
1955–59	-0.56	-0.63	-0.79	-1.11	-0.86	-0.92	-1.16	-1.51
1960–64	-0.60	-0.67	-0.86	-1.24	-0.87	-0.93	-1.19	-1.57
1965–69	-0.64	-0.71	-0.92	-1.36	-0.85	-0.92	-1.20	-1.60
1970–74	-0.68	-0.76	-0.99	-1.48	-0.86	-0.93	-1.23	-1.65
1975–79	-0.72	-0.80	-1.05	-1.59	-0.84	-0.90	-1.20	-1.60
1980–84	-0.75	-0.83	-1.09	-1.66	-0.82	-0.89	-1.19	-1.58
1985–89	-0.74	-0.82	-1.08	-1.60	-0.80	-0.87	-1.17	-1.52
1990–94	-0.71	-0.78	-1.03	-1.53	-0.77	-0.83	-1.11	-1.45
1995–99	-0.67	-0.74	-0.97	-1.44	-0.72	-0.78	-1.05	-1.37
2000–03	-0.65	-0.72	-0.94	-1.39	-0.70	-0.76	-1.01	-1.32

TABLE S.1 Reform 2004–2005, pay-as-you-go, Change in SSW as a fraction of the annual average wage

	Men				Women			
Cohort	**Elemen-tary**	**Lower**	**Upper**	**Univer-sity**	**Elemen-tary**	**Lower**	**Upper**	**Univer-sity**
1935–39	0.37	0.39	0.41	0.42				
1940–44	0.21	0.24	0.52	0.89				
1945–49	-2.38	-1.83	-0.41	3.97	0.45	0.48	0.58	0.97
1950–54	-1.83	-1.22	0.11	4.69	-4.29	-4.03	-2.81	-0.23
1955–59	-1.29	-0.66	0.62	5.27	-4.50	-4.26	-3.20	-0.87
1960–64	-0.80	-0.14	1.10	5.89	-4.08	-3.82	-2.82	-0.43
1965–69	-0.37	0.33	1.54	6.42	-3.61	-3.33	-2.32	0.16
1970-74	0.01	0.76	1.92	6.86	-3.23	-2.89	-1.91	0.60
1975–79	0.35	1.13	2.28	7.31	-2.89	-2.51	-1.55	0.98
1980–84	0.65	1.45	2.58	7.55	-2.54	-2.12	-1.16	1.33
1985–89	0.75	1.54	2.62	7.24	-2.32	-1.90	-0.95	1.33
1990–94	0.71	1.46	2.49	6.87	-2.20	-1.80	-0.90	1.26

TABLE S.2 Reform 2004–2005, mixed system, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elemen-tary	Lower	Upper	Univer-sity	Elemen-tary	Lower	Upper	Univer-sity
1953–54	-1.72	-1.16	0.14	4.59				
1955–59	-1.38	-0.80	0.45	4.94	-4.61	-4.41	-3.32	-1.19
1960–64	-0.91	-0.32	0.90	5.51	-4.25	-4.05	-3.02	-0.88
1965–69	-0.49	0.14	1.33	6.05	-3.80	-3.59	-2.56	-0.35
1970–74	-0.10	0.57	1.72	6.53	-3.41	-3.17	-2.15	0.08
1975–79	0.26	0.95	2.08	6.96	-3.07	-2.80	-1.83	0.35
1980–84	0.56	1.28	2.37	7.15	-2.71	-2.42	-1.45	0.61
1985–89	0.67	1.38	2.41	6.85	-2.50	-2.18	-1.24	0.63
1990–94	0.63	1.30	2.29	6.50	-2.37	-2.07	-1.18	0.59

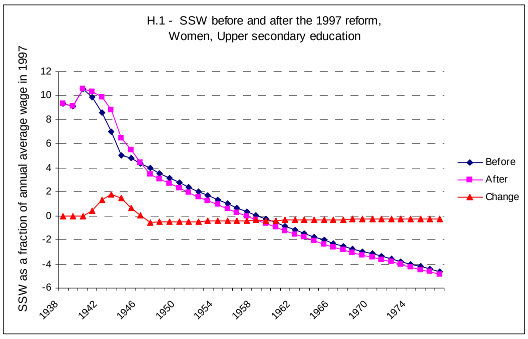
TABLE S.3 Reform 2006, pay-as-you-go, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elemen-tary	Lower	Upper	Univer-sity	Elemen-tary	Lower	Upper	Univer-sity
1935–39	0.00	0.00	0.00	0.00				
1940–44	0.00	0.00	0.00	0.00				
1945–49	0.30	0.19	0.02	-0.85	0.00	0.00	0.00	0.00
1950–54	0.08	0.09	-0.05	-0.26	0.44	0.40	-0.03	-0.16
1955–59	0.00	0.03	-0.10	-0.25	-0.03	0.00	-0.09	-0.04
1960–64	-0.07	-0.05	-0.19	-0.45	-0.07	-0.02	-0.08	-0.13
1965–69	-0.11	-0.11	-0.21	-0.41	-0.09	-0.02	-0.09	-0.13
1970–74	-0.12	-0.14	-0.19	-0.32	-0.03	-0.01	-0.03	-0.02
1975–79	-0.14	-0.16	-0.19	-0.36	0.00	0.00	0.00	0.00
1980–84	-0.13	-0.16	-0.19	-0.35	0.00	0.00	0.00	0.00
1985–89	-0.13	-0.15	-0.18	-0.34	0.00	0.00	0.00	0.00
1990–94	-0.12	-0.14	-0.17	-0.32	0.00	0.00	0.00	0.00

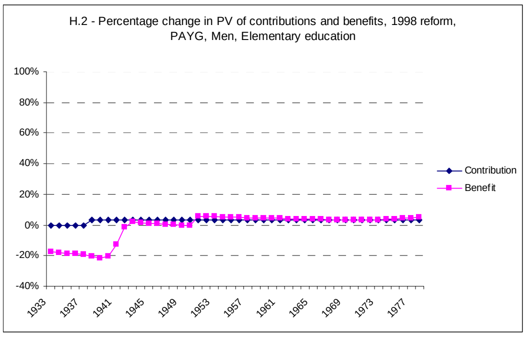
TABLE S.4 Reform 2006, mixed system, Change in SSW as a fraction of the annual average wage

Cohort	Men				Women			
	Elemen-tary	Lower	Upper	Univer-sity	Elemen-tary	Lower	Upper	Univer-sity
1953–54	0.02	0.06	-0.06	-0.10				
1955–59	0.00	0.02	-0.08	-0.21	-0.02	0.00	-0.08	-0.03
1960–64	-0.06	-0.05	-0.15	-0.37	-0.05	-0.01	-0.07	-0.10
1965–69	-0.10	-0.10	-0.18	-0.35	-0.06	-0.02	-0.07	-0.09
1970–74	-0.12	-0.13	-0.18	-0.31	-0.02	-0.01	-0.02	-0.01
1975–79	-0.14	-0.16	-0.19	-0.36	0.00	0.00	0.00	0.00
1980–84	-0.13	-0.16	-0.19	-0.35	0.00	0.00	0.00	0.00
1985–89	-0.13	-0.15	-0.18	-0.34	0.00	0.00	0.00	0.00
1990–94	-0.12	-0.14	-0.17	-0.32	0.00	0.00	0.00	0.00

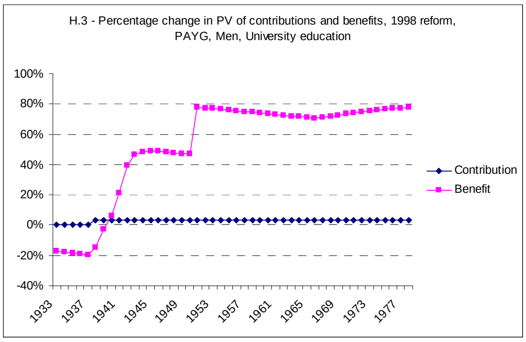
Figures – Hungary



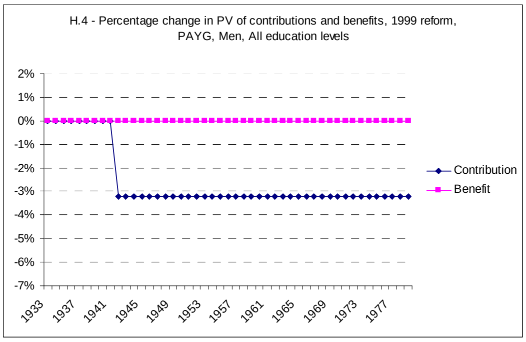
H.1 - SSW before and after the 1997 reform, Women, Upper secondary education



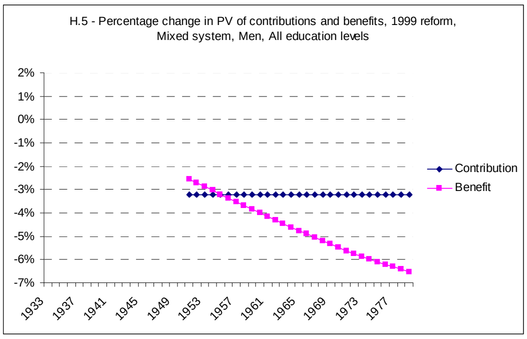
H.2 - Percentage change in PV of contributions and benefits, 1998 reform, PAYG, Men, Elementary education



H.3 - Percentage change in PV of contributions and benefits, 1998 reform, PAYG, Men, University education

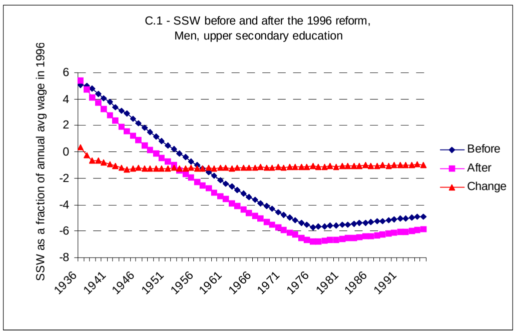


H.4 - Percentage change in PV of contributions and benefits, 1999 reform, PAYG, Men, All education levels

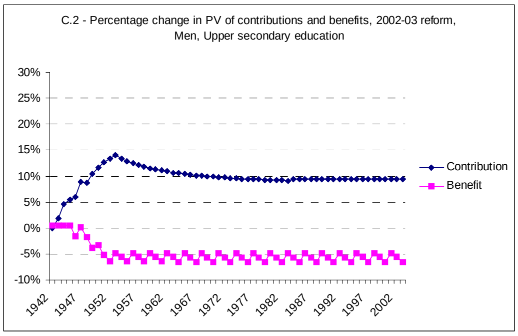


H.5 - Percentage change in PV of contributions and benefits, 1999 reform, Mixed system, Men, All education levels

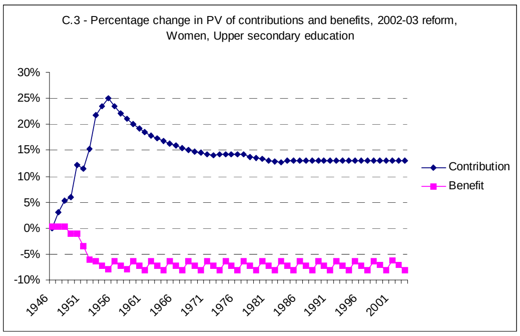
Figures – Czech Republic



C.1 - SSW before and after the 1996 reform,
Men, upper secondary education

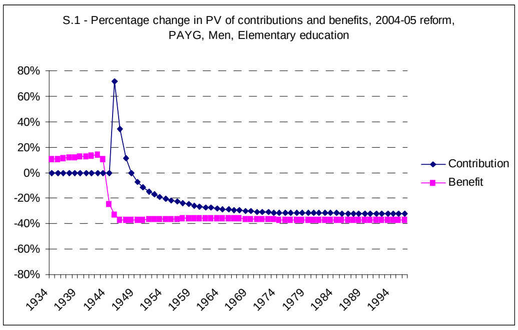


C.2 - Percentage change in PV of contributions and benefits, 2002-03 reform,
Men, Upper secondary education

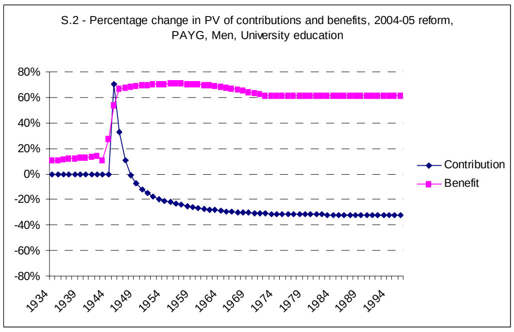


C.3 - Percentage change in PV of contributions and benefits, 2002-03 reform,
Women, Upper secondary education

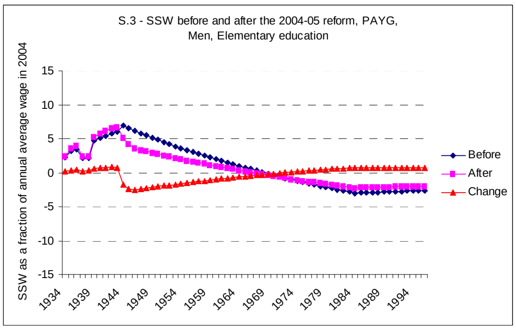
Figures – Slovakia



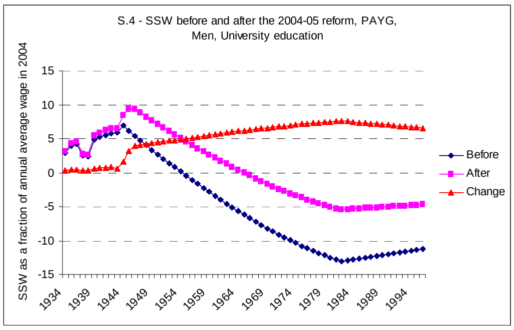
S.1 - Percentage change in PV of contributions and benefits, 2004-05 reform,
PAYG, Men, Elementary education



S.2 - Percentage change in PV of contributions and benefits, 2004-05 reform,
PAYG, Men, University education



S.3 - SSW before and after the 2004-05 reform, PAYG,
Men, Elementary education



S.4 - SSW before and after the 2004-05 reform, PAYG,
Men, University education

APPENDIX B

DETAILED ASSUMPTIONS

Wage profiles

Our "average" workers start working at age 20, work full-time until the standard retirement age[21], and earn the wage that is predicted by the wage profile specific to their gender, educational category, and calendar year. The wage profiles are estimated from individual level cross-sectional datasets described below and have the standard form

$$\log w_{ijt} = \alpha_{jt} + \beta_{1jt} a_{ijt} + \beta_{2jt} a_{ijt}^2 + u_{ijt}$$

where w is the monthly wage, subscript i denotes an individual, j denotes the worker's gender and educational category, t denotes the year, a is the worker's age, and α, β_1, and β_2 are the parameters that we estimate. The profiles were estimated on a sample of workers aged between 20 and the standard retirement age who worked at least 6 months in a given year. The regression estimates and the corresponding wage profiles are available upon request.

We constructed the wage profiles from individual-level datasets that were best suited to the task in each country. All of them contain basic information about each worker (gender, age, education level) and sufficient information about his/her employment status and labor income (either the monthly wage or the annual/quarterly wage and the number of weeks/months worked, from which the monthly wage can be imputed). For Hungary, we used the Harmonized Hungarian Wage Survey of the Public Employment Service. The survey was collected at the firm level in 1986 and 1989 and annually from 1992 to 2003 and contains data on 100,000–200,000 employees depending on the year. For the Czech Republic, we used the Czech Microcensus, a representative household survey conducted once every 4 or 6 years by the Czech Statistical Office. The surveys that we use were collected in 1992, 1996, and 2002[22] and cover approximately 44,000, 64,000, and 19,000 individuals in the respective years. For Slovakia, we used the TREXIMA dataset, a representative survey of firms. Collected in 2001, it contains quarterly data on 350,000 employees.

Since the samples allow us to estimate the wage profiles only for some years[23] while we need to have profiles for all years since 1988 (Hungary), 1986 (Czech Republic) or 1984 (Slovakia), we impute the profiles for the remaining years. We assume that the coefficients on the age and age squared are the same as in the nearest adjacent year for which the profile was

[21] Workers with university education start working at age 22. Women have 2 children; the first one is born at the average age of the first childbirth in the respective country and year (information collected from the population statistics published by each country's statistical office), and the second one two years later. Women spend 5 years (Czech Republic, Slovakia) or 4 years (Hungary) in childcare without earning labor income.

[22] Unfortunately, the 1988 microcensus was not usable for our purpose, since all observations are recorded at the household level and not the individual level. Even though it does report the earnings of the head of household and his spouse, it does not allow us to identify the gender of workers who live in households other than traditional families of married couples.

[23] The Hungarian Wage Survey is not available for 1987–1988, 1990–1991, and 2004+. Moreover, the surveys from 1993, 1998–1999, and 2002 appeared to contain data problems, since the estimates of the wage profiles in these years produced estimates that were substantially different from the estimates for adjacent years and, more importantly, were economically implausible.

estimated[24]. Then we adjust the intercept α such that the average fitted wage in the sample is equal to the actual average wage in the year for which the wage profile is being imputed[25].

The 2004–2005 Slovak reform allowed one of the parents to deduct 0.5 % for every child from their PAYG contributions. We assume that the deduction is claimed by men, since they earn more on average.

Future projections

Certain assumptions about the future were required to project future benefits and contributions. The length of life is probabilistic and future money flows are discounted by the survival probability. We had survival probability tables for all countries (unfortunately without a finer breakdown by education categories) until 2004. For 2005 onwards, we assume that the survival probabilities are the same as in 2004[26].

We assume that as of the time of the reform people had perfect foresight about the evolution of all economic variables that affect future taxes and benefits (aggregate and individual wage growth, inflation, survival probabilities). That is, the future wages, inflation rates and survival probabilities that are expected at the time of the reform are equal to the wages and inflation rates that were actually realized up to 2005, and for 2006 onwards we assume a 3% growth rate of real wages for all education categories and genders and a 2% inflation rate.

The rate of return on savings in pension funds is the key parameter affecting the benefits from the funded pillar. It is to a large extent determined by the regulation of funds' investments and fees. Our choice of rate of return is an estimate of the net rate of return that the pension funds, as actually established and regulated by Hungarian and Slovak law, are expected to deliver to their clients. That is, we avoid using an average historical return on some "optimal" stock and bond portfolio as commonly done in simulations of benefits from the funded pillar (e.g. (Feldstein, Ranguelova, 2001), since that approach would give the level of benefits that the funded pillar could provide rather than did provide.

For Hungary, the expected real return on savings is calculated as the weighted average of the real net return[27] of all Hungarian pension funds during 1998–2005, which was 2.7 %[28] Pension funds in Slovakia were established too recently to project future returns from historical returns. Instead, we compute the expected future returns as the average historical returns on the portfolios that the growth funds currently hold. Specifically, we calculate the average historical return for each of the major bond and stock indices in which the funds currently invest, and then compute the average of these returns weighted by their share in the average growth fund's portfolio[29]. The resulting projected nominal rate of return after

[24] For example, the coefficients on age and age squared estimated from the Czech 2002 Microcensus were used to generate wage profiles for 2000–2004.

[25] The data sources for average wages by gender and education level are reported in detail in an earlier version of the paper (Dušek, Kopecsni, 2008, p. 25).

[26] This assumption probably underestimates the true survival probabilities, since life expectancies have been increasing in all three countries since the 1990s and are expected to increase in the future. However, we were not able to obtain specific projections of future survival probabilities.

[27] That is, after deducting fees.

[28] Source: (Czajlik, Szalay, 2006)

[29] Specifically, the expected returns are computed from the returns on the following indices over the periods indicated: UX 1991–2007, PX 1995–2007, SLOVN SK 1999–2007, VIX 1990–2003, MXEU 1995–2007, FTSE 1990–2003, DAX 1990–2007, and SPX 1990–2007. The funds' stock portfolio is composed of stock

deducting fees is 6.9 percent[30]. As workers approach retirement age they may prefer a gradual switch to a completely risk-free portfolio (and in Slovakia they are in fact required to switch to more conservative funds). We therefore assume that the above mentioned returns apply only from the beginning of employment until 15 years before retirement. Afterwards workers rebalance the portfolio each year such that the real return linearly decreases to zero by the age of retirement.

The pension funds offer unisex life annuities, which we computed as follows: First, the share of men and women upon retirement in the population is weighted by their wages. Next, the population structure in the future is projected by applying the mortality tables separately for the male and female parts of the population. The two projections are combined to obtain the evolution of the unisex population. The annuities are computed by applying the actuarial formulas to this unisex population. Finally, as the annuities are subject to Swiss indexation the technical interest rate was modified by the magnitude of this indexation.

Computing future indexations of benefits in the Czech Republic and Slovakia required additional assumptions. The legislation before the first reform did not prescribe any indexations, yet it is implausible to assume that the benefits or the system parameters would never be indexed. In fact, the benefits had been indexed in an ad-hoc manner with a clear goal to preserve their real value. Therefore, we assume that once granted, benefits would have been indexed for inflation, and the income brackets in the benefit formula would be indexed for wage growth. Under these assumptions, the replacement ratio remains at a similar level (48–50 % in the Czech Republic, 30–35 % in Slovakia) as it was during the years just preceding the reform[31]. After the 1996 reform, Czech law prescribed minimum indexations, but the government frequently provided more generous increases[32]. Therefore, until 2006 we assume perfect foresight and compute the benefits as they were actually indexed, and only after 2006 we index them conservatively by the minimum prescribed by the legislation.

---

indices in the Visegrad countries (20 %), the EU-15 countries (50 %), and the United States (30 %). Data on the portfolio compositions were taken from the funds' annual reports.

[30] The growth funds currently invest 80 % of their assets in bonds, which appears to be an overtly conservative strategy, particularly if the legislation restricts them to investing at most 80 % in stocks. Even though other regulations give funds incentives to invest in stocks below the maximum limit, several fund managers admit in official reports that they do plan to increase the share of stocks in the near future. In our computation we therefore assume that they will invest 30 % in stocks.

[31] In addition, prior to 1995, the new benefits were computed according to the old formula but were increased immediately (by 32 % in 1995) to make up for the inflation that had accumulated since 1990. We assume that such increases in newly granted benefits would continue into the future with the same purpose of compensating for the reduction in the real value of past wages that enter the benefit due to inflation. We increase the new benefits by 32 %, and further increase them by the ratio of the price index at the time of retirement to the average price index during the 5 years preceding retirement.

[32] (Dušek, 2007)

APPENDIX C

MAIN FEATURES OF THE REFORMS

	Hungary (1993)
Pension Scheme	PAYG system
Retirement Age	Men: 60, Women: 55 to 60 gradually
Contribution Rate	Employer: 24.5 %, Employee: 6 %
Assessed Earnings	Average net monthly earnings during 4 years with highest earnings in the 5 years before retirement ➔ average net monthly earnings from 1988 until the year of retirement
Benefit Formula	The benefit is set as a certain fraction (pension accrual) of average net earnings during the period considered. The benefits are regressive in average net earnings but less regressive after the reform.
Indexation Rule	Net wage indexation
	Hungary (1997)
Pension Scheme	PAYG system
Retirement Age	Men: 60 to 62 gradually, Women: 55 to 62 gradually except cohorts 1942–1944 whose retirement age was shifted back by 1 year
Contribution Rate	Employer: 24.5 % to 24.0 %, Employee: 6 %
Assessed Earnings	No change
Benefit Formula	Higher pension accrual was applied
Indexation Rule	No change
	Hungary (1998)
Pension Scheme	The reform split the mandatory PAYG scheme into a public PAYG and privately funded pillar. The workers already employed had the option to switch from the public to the mixed system and more than 50 % of eligible workers did switch. For new entrants to the labor market, participation in the mixed system was compulsory.
Retirement Age	No change
Contribution Rate	PAYG scheme – Employer: 24 % to 23 % (1999), to 22 % (2000), Employee: 6 % to 7 % (1998), to 8 % (1999), to 9 % (2000) Mixed scheme – Employer: 24 % to 23 % (1999), to 22 % (2000), Employee: (PAYG pillar): 1 % Employee (Funded pillar): 6 % (1998), 7 % (1999), 8 % (2000)
Assessed Earnings	No change
Benefit Formula	The benefit formula should switch from the net to the gross principle after 2013, meaning that the benefit will then be set as a fraction of average gross earnings instead of net earnings. It was also planned that the benefit would become taxable. The degression is gradually eliminated in the calculation. The switchers in the mixed system will have their benefit from the public pillar reduced by 25 % in such a way that their accruals are 75 % of stayers' accruals. In addition, they will receive unisex annuities from savings in the pension fund.
Indexation Rule	Net wage indexation ➔ Swiss indexation, gradually since 2001
	Hungary (1999)
Pension Scheme	Multi pillar system – public PAYG pillar with privately funded pillar
Retirement Age	No change
Contribution Rate	PAYG scheme – Employer: 23 % to 22 % (1999), 22 % to 21 % (2000), Employee: 8 % (1999), 9 % to 8 % (2000), Mixed scheme – Employer: 23 % to 22 % (1999), 22 % to 21 % (2000), Employee (PAYG pillar): 1 % to 2 %, Employee (Funded pillar): 7 % to 6 % (1999), 8 % to 6 % (2000)
Assessed Earnings	No change
Benefit Formula	No change
Indexation Rule	No change
	Hungary (2003)
Pension Scheme	Multi pillar system – public PAYG pillar with privately funded pillar
Retirement Age	No change
Contribution Rate	PAYG scheme – Employer: 18 %, Employee: 8 % to 8.5 %, Mixed scheme – Employer: 18 %, Employee (PAYG pillar): 2 % to 1.5 %, Employee (Funded pillar): 6 % to 7 %
Assessed Earnings	No change
Benefit Formula	Gradual introduction of an additional monthly benefit (13th monthly pension) within PAYG pillar. Pensioners would effectively receive their benefit 13 times a year from 2006 onwards.
Indexation Rule	No change
	Hungary (2007)
Pension Scheme	Multi pillar system – public PAYG pillar with privately funded pillar

Retirement Age	No change
Contribution Rate	PAYG scheme – Employer: 17 % to 21 % (2007), 16 % to 21 % (2009), Employee: 8.5 %, Mixed scheme – Employer: 17 % to 21 % (2007), 16 % to 21 % (2009), Employee (PAYG pillar): 0.5 %, Employee (Funded pillar): 8 %
Assessed Earnings	No change
Benefit Formula	For workers who would retire between 2008 and 2012 the employees' pension and health care contributions and the employees' contribution to the employment fund will be deducted from the net earnings entering the benefit formula in a way that will reduce the benefit. Earnings during the whole life will be indexed to the level of the individual's last working year, while before the reform earnings in the last three working years were not indexed at all.
Indexation Rule	No change
	Czech Republic (1996)
Pension Scheme	PAYG system
Retirement Age	Men: 60 to 62 gradually, Women: 55 to 59 gradually
Contribution Rate	Employer: 19.5 %, Employee: 6.5 %
Assessed Earnings	Earnings from 5 years with the highest earnings during the 10 years prior to retirement ➜ average monthly earnings from the 30 years of employment preceding retirement since 1986
Benefit Formula	The new benefit formula introduced a flat component of the benefit (same for all retirees) and at the same time made the variable component (which depends on the worker's average lifetime earnings) less regressive; the ceiling on the maximum benefit was abolished
Indexation Rule	Indexation ad hoc ➜ indexation to the consumer price index and at least once every two years also for at least 33 % of real wage growth, but the government has the discretion to provide more generous indexation.
	Czech Republic (2002–2003)
Pension Scheme	PAYG system
Retirement Age	Men: 60 to 63 gradually, Women: 55 to 61 gradually by 2013
Contribution Rate	Employer: 19.5 % to 21.5 %, Employee: 6.5 %
Assessed Earnings	No change
Benefit Formula	No change
Indexation Rule	Benefits have to be adjusted annually and the minimum increase has to include inflation plus at least 33 % of real wage growth
	Slovakia (2004–2005)
Pension Scheme	The reform split the mandatory PAYG scheme into a public PAYG and privately funded pillar. The mixed system is mandatory for new entrants to the labor market. Workers aged below 52 had a choice to switch from pure PAYG to a mixed system, and 60 % of workers had switched by 2006.
Retirement Age	Men: 60 to 62 gradually, Women: 55 to 62 gradually
Contribution Rate	PAYG scheme – Employer: 20.6 % to 16 % (2004), to 14 % (2005), Employee: 5.9 % to 4 %, one of the parents can deduct an additional 0.5 % in contributions for every child, Mixed scheme – Employer (PAYG pillar): 5 % (2005), Employer (Funded pillar): 9 % (2005), Employee: 4 %, one of the parents can deduct an additional 0.5 % in contributions for every child
Assessed Earnings	Earnings from 5 years with the highest earnings during the 10 years prior to retirement ➜ entire working period since 1994, which in turn should be at least 10 years
Benefit Formula	The new benefit formula made the benefit linear in the worker's average earnings over his entire working history since 1994, up to a cap beyond which workers with more than 3 times the average earnings do not receive higher benefits. The formula set the benefit as the worker's average earnings times the number of working years times the actual pension value. During a transitory period until 2006 the benefits were regressive in the worker's lifetime earnings but were gradually becoming less regressive over time. The PAYG benefits for switchers are cut proportionally to the number of years they have participated in the mixed system. In addition, they will receive unisex annuities from savings in the pension fund.
Indexation Rule	Indexation ad hoc ➜ Swiss indexation
	Slovakia (2006)
Pension Scheme	Multi pillar system – public PAYG pillar with privately funded pillar
Retirement Age	No change
Contribution Rate	Employee: the 0.5% deduction in contributions for every child was abolished
Assessed Earnings	Entire working period since 1994, which in turn should be at least 10 years ➜ entire working period since 1984, which in turn should be at least 20 years
Benefit Formula	Gradual adjustment in the benefit formula to create a stronger link between earnings and benefits. This was supposed to have been fully phased in by 2006, but was prolonged until 2014, after which the benefits should indeed be linear in earnings.
Indexation Rule	No change

REFERENCES

Augusztinovics M, Gal RI, Matits A, Mate L, Simonovits A, Stahl J (2002): The Hungarian Pension System before and after the 1998 Reform. In: Fultz E (Ed.): *Pension Reform in Central and Eastern Europe*, Vol. I:25–93.

Ball L, Mankiw NG (2007): Intergenerational Risk Sharing in the Spirit of Arrow, Debreu, and Rawls, with Applications to Social Security Design. *Journal of Political Economy*, 115(4).

Blake D (2008): What is a Promise from the Government Worth? Quantifying Political Risk in State and Personal Pension Schemes in the United Kingdom. *Economica*, 75:342–361.

Bohn H (2001): Social security and demographic uncertainty. In: Campbell JY, Feldstein M (Eds.): *Risk Aspects of Investment-Based Social Security Reform*. Chicago, University of Chicago Press: 203–246.

Czajlik I, Szalay Gy (2006): A magánnyugdíjpénztárak működése és szabályozása. *Magyar Nemzeti Bank tanulmányok*, no. 48.

Diamond P (1997): Insulation of Pensions from Political Risk. *NBER Working Paper*, no. 4895. Published also in: Valdez S (Ed.): *The Economics of Pensions: Principles, Policies, and International Experience.* Cambridge University Press, Cambridge.

Dušek L, Kopecsni J (2008): Policy Risk in Action: Pension Reforms and Social Security Wealth in Hungary, Czech Republic and Slovakia. *Discussion Paper*, no. PI-0812, The *Pensions Institute, Cass Business School*, London.

Dušek L (2007): Political risk of social security: The case of the indexation of benefits in the Czech Republic. *CERGE-EI working paper*, no. 318.

Feldstein M, Ranguelova E (2001): Individual Risk in an Investment-based Social Security System. *American Economic Review*, 91(4):1116–1125.

Feldstein M, Ranguelova E, Samwick A (2001): The Transition to Investment-Based Social Security when Portfolio Returns and Capital Profitability Are Uncertain. In: Campbell JY, Feldstein M (Eds.): *Risk Aspects of Investment-Based Social Security Reform*. University of Chicago Press.

Feldstein M (2005a): Reducing the Risk of Investment-Based Social Security System. *NBER Working Paper*, no. 11084

Feldstein M (2005b): Structural Reform of Social Security. *NBER Working Paper*, no. 11098

Gordon RH, Varian HR (1988): Intergenerational risk sharing. *Journal of Public Economics*, 37(2): 185–202.

Holst R (2005): Social Security and Policy Risk: Evidence and Effects on Wealth Accumulation. *Working paper, University of Chicago.*

Lindbeck A, Persson M (2003): The Gains from Pension Reform. *Journal of Economic Literature*, XLI:74–112.

McHale J (2001): The Risk of Social Security Benefit-Rule Changes. In: Campbell JY, Feldstein M (Eds.): *Risk Aspects of Investment-Based Social Security Reform*. University of Chicago Press: 247–282.

Melichercik I, Ungvarsky C (2004): Pension Reform in Slovakia: Perspectives of the Fiscal Debt and Pension Level. *Finance a úvěr - Czech Journal of Economics and Finance*, 54(9-10):391–404.

Müller K (2002): From the State to the Market? Pension Reform Paths in Central-Eastern Europe and the Former Soviet Union. *Social Policy and Administration*, 36(2):156–175.

Poterba J, Rauh J, Venti S, Wise D (2005): Utility Evaluation of Risk in Retirement Saving Accounts. In: Wise D (Ed.): *Analyses in the Economics of Aging*. Chicago, University of Chicago Press: 13–52.

Shoven JB, Slavov SN (2006): Political Risk versus Market Risk in Social Security. *NBER Working Paper*, no. 12135.

JEL Classification: G21, G28
Keywords: credit risk, loss given default, fractional responses, ordinal regression, quasi-maximum likelihood estimator

Modeling Bank Loan LGD of Corporate and SME Segments: A Case Study

Abstract
Loss given default (LGD) is one of key parameters to estimate credit risk in an internal rating based approach considered in The New Basel Capital Accord. The aim of this essay is to find determinants of LGD using a set of firm loan micro-data of an anonymous Czech commercial bank. We find that LGD is driven primarily by the period of loan origination, relative value of collateral, loan size and length of business relationship. Different models employed in our analysis provide similar results; in more complex models, log-log models appear to perform better, implying an asymmetric response of the dependent variable.

1. Introduction

The New Basel Capital Accord (Basel Committee on Banking Supervision, 2006) has been created with an objective to better align regulatory capital with the underlying risk in the bank's credit portfolio. The New Accord motivates international banks to develop and use internal risk models for calculating credit risk capital requirement. It allows banks to compute their regulatory capital in two ways: (1) using a revised standardised approach based on the 1998 Capital Accord which applies regulatory ratings for risk weighting assets or (2) using an internal rating based (IRB) approach where banks are permitted to develop and employ their own internal risk ratings.

The IRB approach is based on three key parameters used to estimate credit risk: PD – a probability of default of a borrower over a one-year horizon, LGD – loss given default, a credit loss incurred if a counterparty of a bank defaults, and EAD – an exposure at default. These parameters are used to estimate an expected loss, which is a product of PD, LGD and EAD. There are two possible variants of IRB, the foundation and the advanced approach. The difference between them lies in the way of estimating parameters. In the foundation approach only PD is estimated internally, LGD and EAD are based on supervisory values. On the other hand, in the advanced approach, all parameters are determined by a bank.

Most banks are prepared to use the foundation approach, since they have already built internal models to estimate PD. However, many banks are not ready to fully implement the advanced IRB approach, because the advanced approach also requires to model and determine LGD.

This research contributes to propose a methodology to estimate the loss given default and then apply it to a set of loan micro-data to small and medium sized enterprises (SMEs) and corporations. The data was provided by an anonymous Czech commercial bank (the "Bank"). The access to a unique database of loans enables us to show empirically economic determinants of LGD. The focus of this chapter is on identification of LGD drivers using various statistical approaches.

Based on the literature, we propose and apply three different statistical modeling techniques in order to estimate determinants of LGD — (1) generalised linear models using symmetric logit and asymmetric log-log link functions for ordinal responses, as well as

(2) for fractional responses using beta inflated distribution and (3) quasi-maximum likelihood estimator. Moreover, several ways how to measure predictive performance are suggested.

Our essay is organised as follows: the second section is a brief literature review, the third section discusses the key regulatory issues regarding LGD, such as definition of default and measurement of LGD, the fourth section focuses on characteristics of LGD from a modeling perspective and on description of data, the fifth section analyses and discusses typical risk drivers, the next section depicts the regression methodology used, whereas the last three sections provide results, goodness-of-fit performance measures and conclusions, respectively.

2. Literature Review

The banks which are using the advanced IRB approach need to consider common characteristics of losses and recoveries. These basic characteristics are bimodality, seniority and type of collateral, business cycles, industry and the size of loan.

Loss given default[33], defined as 100% minus a percentage of recovered exposure during a workout process, tends to have a bimodal distribution. Bimodality implies that most loans have LGD close to 0% (full recovery), or there is 100% LGD (no recovery at all). Bimodality makes parametric modeling of recovery difficult and requires a non-parametric approach (Renault and Scaillet, 2004).

The second important issue is collateral of defaulted claims and a place in the debt structure. Bank loans are typically at the top of the debt structure, generally implying higher recovery rates than bonds. LGD tends to be lower (i.e. recovery rate tends to be higher) when a claim is secured by collateral of high quality. Asarnov and Edwards (1995), Carey (1998) and Gupton et al. (2000) confirm that seniority and collateral do matter. They primarily use the data from Citibank and Moody's.

There is strong evidence that LGD in recessions is higher than during expansions, for instance according to Carey (1998) and Frye (2000). Employing Moody's data they show that during recessions recoveries are lower by one third.

Other studies by Grossman et al. (2001) and Acharya et al. (2003) argue that industry type is another important determinant of LGD. Results of Altman and Kishore (1996) provide evidence that some industries such as utilities (30% average LGD) do better than others (e.g. manufacturing 58%).

The most ambiguous key characteristic is the size of loan. Asarnov and Edwards (1995) and Carty and Lieberman (1996) find no relationship between LGD and the size of loan in the U.S. market. Thornburn (2000) obtains similar negative result for Swedish business bankruptcies. However, Hurt and Felsovalyi (1998) show that large loan defaults exhibit lower recovery rates. They attribute it to the fact that large loans are often unsecured, and they are provided to economic groups that are family owned.

Currently, bank loan LGD is not explored well by theoretical and empirical literature. Although several empirical academic studies have analyzed credit risk on corporate bonds, only a few studies have been applied to bank loans. The reason for this is that since bank loans are private instruments, little data is publicly available.

[33] The same applies to recovery rate defined as 100% minus a LGD percentage. To retain consistency among results, we have recalculated recovery rates of some studies to LGD.

Asarnow and Edwards (1995) analyzed 831 defaulted loans at Citibank over the period 1970–1993 and they show that the distribution of LGD is bimodal, with concentration of LGD on either the low or high end of the distribution. Their average LGD is 35%. Carty and Lieberman (1996) measured the recovery rate on a sample of 58 bank loans for the period 1989–1996 and reported skewness toward the high end of price scale with an average LGD of 29%. Gupton et al. (2000) reported lower LGD of 30% for senior secured loans as compared to unsecured loans (48%), based on 1989–2000 data sample consisting of 181 observations. The above studies focused on the U.S. market. Hurt and Felsovalyi (1998) who analyze 1,149 bank loan losses in Latin America over 1970–1996 found an average LGD of 32%. Another study by Franks et al. (2004) calculated recovery rates of 2,280 defaulted companies whose data was taken from 10 banks in three countries over the period 1984–2003. They found a country-specific bankruptcy regime, which indicates significantly different recovery rates. Average LGD is 25% for the UK, 39% for Germany and 47% for France. The results of these studies are sensitive to the analyzed data sample, regulatory framework of workout process and modeling techniques used and therefore it is hard to compare them directly.

The paper by Dermine and Neto de Carvalho (2006) is the first study to apply the workout LGD methodology on a micro-data set from Europe. They estimate LGD for a sample of 374 corporate loans over the period 1995–2000. The estimates are based on the discounted value of cash flows recovered after the default event and the estimated LGD is 29%. They find that beta distribution does not capture the bimodality of data and using multivariate analysis they identify several significant explanatory variables.

3. *Key Regulatory LGD Issues*

The following are the key LGD issues (Schuermann, 2004): (1) definition and measurement, (2) key drivers and (3) modeling and estimation approaches. In this section we describe some of these characteristics of LGD which are important for the empirical part of our essay.

LGD is typically defined as a ratio of losses to an exposure at default. There are three classes of LGD for an instrument. These are market, workout and implied market LGD. Market LGD is observed from market prices of defaulted bonds or marketable loans soon after an actual default event. Workout LGD is derived from the set of estimated cash flows resulting from a workout and collection process, properly discounted to the date of default. Thirdly, implied market LGD is derived from the risky but not defaulted bond prices, using a theoretical asset pricing model. In this essay only workout LGD is considered. A recent study by Seidler and Jakubík (2009) deals with implied market LGD in the Czech economic context.

3.1 Definition of Default

There is no standard definition of default. Different definitions are used for different purposes. Even the international rating agencies, like S&P, Moody's and Fitch, use different default definitions. However, a measured loss at default depends on the definition, so it is important to make the definition that is used clear.

According to the Bank for International Settlements (BIS), default is a situation when an obligor is unlikely to pay credit obligations or the obligor is past due more than 90 days on any material credit obligation. We follow this definition.

Table 1 – Different Measurement of LGD on the Portfolio Level

i is a default observation, y is the year of default, there are n_y defaults in each year and a total of m years of observations, LR is the loss rate or LGD for each observation

	Default count averaging	Exposure weighted averaging
Default weighted averaging	$LGD = \dfrac{\sum_{y=1}^{m}\sum_{i=1}^{n_y} LR_{i,y}}{\sum_{y=1}^{m} n_y}$	$LGD = \dfrac{\sum_{y=1}^{m}\sum_{i=1}^{n_y} EAD_{i,y} \times LR_{i,y}}{\sum_{y=1}^{m}\sum_{i=1}^{n_y} EAD_{i,y}}$
Time weighted averaging	$LGD = \dfrac{\sum_{y=1}^{m}\left(\dfrac{\sum_{i=1}^{n_y} LR_{i,y}}{n_y}\right)}{m}$	$LGD = \dfrac{\sum_{y=1}^{m}\left(\dfrac{\sum_{i=1}^{n_y} EAD_{i,y} \times LR_{i,y}}{\sum_{i=1}^{n_y} EAD_{i,y}}\right)}{m}$

3.2 Measurement of LGD

There are four ways of measuring long-term average LGD at the portfolio level, using default weighted averaging vs. time weighted averaging and default count averaging vs. exposure weighted averaging. *Table 1* shows these four options.

The time weighted averaging is less desirable as it smoothes out high LGD years with low ones; therefore, it can underestimate the LGD. Thus, the default weighted averaging is used in practice. The default count averaging is recommended for the non-retail segment and it is employed in our analysis along with the default weighted averaging. On the other hand, the exposure weighted averaging is frequently used for retail portfolios.

3.3 Economic Loss

The loss used in LGD estimation for regulatory purposes is the economic loss. When measuring the economic loss, all relevant factors should be taken into account, such as material discount effects and material direct and indirect costs associated with collection of the exposure[34]. Direct (external) costs include the fees paid to an insolvency practitioner, costs of selling assets, costs of running a business and other professional fees. Indirect (internal) costs are the costs incurred by a bank for recovery in the form of intensive care and workout department costs. Economic loss should also consider costs of holding non-performing assets (funding costs) over a work-out period. The funding costs should be reflected in an appropriate discount rate, which includes a risk premium of the underlying assets[35]. Moreover, it is also important to understand effectiveness of the workout process in time, particularly to make appropriate assumptions for modeling LGD[36].

To estimate internal costs, several methods are possible. Aggregate workout costs or costs of intensive care of the workout department could be related to the (1) aggregated amount of exposure, (2) aggregate recovery amount or (3) to the number of defaults in a given period. The reasoning for the first alternative is that more costs are allocated to events with larger exposure. However, the amount recovered is even more important, so the option where higher

[34] Taking into account these factors distinguish an economic loss from an accounting loss.
[35] The issue of an appropriate discount rate is discussed in the following section.
[36] The analysis of workout period length and time distribution of recoveries is presented later in the next section.

costs are related to higher recoveries could be more appropriate[37]. The third case suggests that workout costs are more or less constant, regardless of the size of an exposure or recovery of a particular file. In this study internal costs of the workout process are estimated as 1.8%[38] relative to the recovered amount based on past experience. Actual external costs were available for each default case, so they are used in our analysis.

3.4 Downturn LGD

The last important regulatory issue we would like to discuss is downturn LGD. Basel II requires reflecting economic downturn conditions when estimating LGD. Downturn LGD cannot be less than a long-run default-weighted average. To estimate downturn LGD based on own historical data, banks need to have at least seven years long period of high quality dataset. In most Central European commercial banks this condition is currently not met. For this reason, several options are available to achieve an indication of downturn LGD:

– use a different discount factor;
– work with default weighted LGD instead of exposure weighted or time weighted LGD;
– take into consideration non-closed files, where a recovery is lower;
– use macroeconomic factors within several stress scenarios; or
– choose 5 worst years out of last 7 years.

To estimate downturn LGD, non-closed files are recommended to be included in the model, until there are enough of long periods available. In this study, non-closed files are included also, hence the estimated LGD can be considered as an indication of downturn LGD.

4. Data Sample and Selected Modeling Issues

In this section we describe a portfolio that is analyzed in the essay and we discuss associated issues of effective workout period and discount rate.

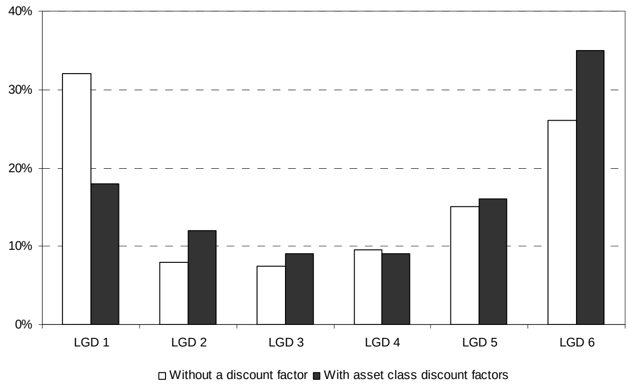
4.1 Data sample

The original data sample is based on all available historical closed files for 1989–2007[39] and all open defaulted issues. In the first step we picked all closed files. Then we decided to enhance the dataset and so we also included those non-closed files whose recovery period was longer than the effective recovery period. As shown in the analysis presented later in this section, after twelve quarters of a workout process, a recovery increases only slightly. Hence, such files might be considered as closed. We are aware of the fact that our estimation of LGD could be overestimated[40].

Additionally, we decided to split the sample into two parts; the first subsample includes the cases closed within a year, whereas the second part contains defaults with a longer recovery period. Observations with a very short workout period likely represent special cases that are different from a normal workout process. These might be either “technical defaults” when a client falls in the definition of default for having temporary past due obligations (LGD close

[37] For these two alternatives it is preferable to set a floor and a ceiling for minimum and maximum internal costs
[38] Dermine and Neto de Carvalho (2006) receive similar 1.2% internal costs.
[39] In early years of this period, however, not all defaults were recorded and some information was missing. Moreover, recent defaults are not closed and workout periods are short, so this data is not included in our dataset. The majority of quality data is for the period 1995–2004.
[40] On the other hand, as we have noted, employing this approach is an indication of downturn LGD.

Figure 1 The Effect of a Discount Factor on LGD[a]



Note: [a] LGD grades 1 to 6 are based on Moody's grades and are described in the next section.

to 0%) or cases of frauds with LGD close to 100%. Possibly different determinants of LGD might be important for each subsample, so we analyze the whole sample and each of the subsamples separately. The overall LGD is 52%, for files closed within a year the figure is 16%, while for the second subsample LGD amounts to 60%[41]. Our results are in line with other empirical studies.

Due to a relatively low number of observations closed within a year, in this study we only focus on LGD determinants of cases resolved outside a year. It might be important to determine ex ante which cases are likely to fall in each subsample. We performed a logistic regression analysis in order to find factors which govern whether a defaulted case is likely to be settled within a year, or its workout period is expected to be longer. The same explanatory variables as to explain determinants of LGD were experimented with, however, no conclusive factors were found. We believe that technical defaults or frauds can be more easily detected by an expert judgment.

Altogether there are several hundred data points[42]. For each default case an amount of cash flows received from the workout process[43] and their timing is available together with other data collected by the workout department, such as exposure at default, type and amount of collateral, type of loan, a year of loan origination, etc.

The observations are aggregated at the level of counterparty. Date of default is determined on counterparty level and is the same for all contracts related to that client. Contract drivers are assigned to a client level in different ways. Exposures on all contracts of a certain client are summed up and create a total exposure at the client. Similarly, for each default case the amounts of cash flow received from the workout process on all contracts are aggregated and create a total cash flow received during the workout process. In the case of loan origination, the oldest loan was taken into consideration.

To account for bimodality we used an option to map continuous LGD to a number of LGD grades. In each of these classes, data is more normally distributed than overall LGD. We use LGD grades based on Moody's[44].: LGD_1 is 0–10%, LGD_2 is 10–30%, LGD_3 is 30–50%, LGD_4 is 50–70%, LGD_5 is 70–90% and LGD_6 is 90–100%. The frequency counts based on

[41] For comparison with studies which include only closed files, the entire sample can be divided into closed files (LGD of 34%) and open files (LGD of 67%). The LGD of closed files outside a year is 45%.

[42] A more exact number is not presented to preserve confidentiality of the Bank.

[43] The cash flows from the workout process equal recovered amount minus direct costs of recoveries.

[44] Alternatives are the other major rating agencies such as S&P and Fitch with similar LGD grades.

these grades as already depicted in *Figure 1,* which reveals a binomial pattern of the LGD distribution.

LGD can be less than 0%, implying that a bank ultimately recovers more than 100% or more than 100%, e.g. as a result of high workout costs which exceed recoveries. LGD needs to be cut off to avoid distortions. In this essay LGD is censored between 0 and 1, similarly to many other publications.

4.2 Effective Length of Workout Period

The estimation of a workout period length and analysis of recoveries in time is important from both regulatory and modeling perspective. The recovery period starts when client defaults or when workout department undertakes a file. The recovery period ends when the file is officially written-off or when the counterparty recovers and gets back to the portfolio of performing loans. Nowadays, most issues in Central European commercial banks are non-closed because of a relatively short period since transition to market economy and emergence of first defaults. Some of the non-closed files can be included in a sample of closed files if the estimated amount to be recovered is not significant. For these cases the length of the workout period can be considered:

– until non-recovered value is less than 5% of EAD;
– one year after default (mainly used in retail);
– +25% upper percentile from the distribution of length of workout period;
– until effective recovery period (useful for non-retail).

In this research, the last option is used; the effective workout period is estimated on the basis of cumulative recovery rate analysis. A cumulative recovery rate was calculated in order to show dynamic evolution of recovery rates over time[45], i.e. the time distribution of recovery rates. This enables us to analyze the evolution of recovered amount and to identify reasons of a non-efficient workout process. Count weighted and exposure weighted average cumulative and marginal recovery rates are calculated quarterly after the default date (see *Appendix*).

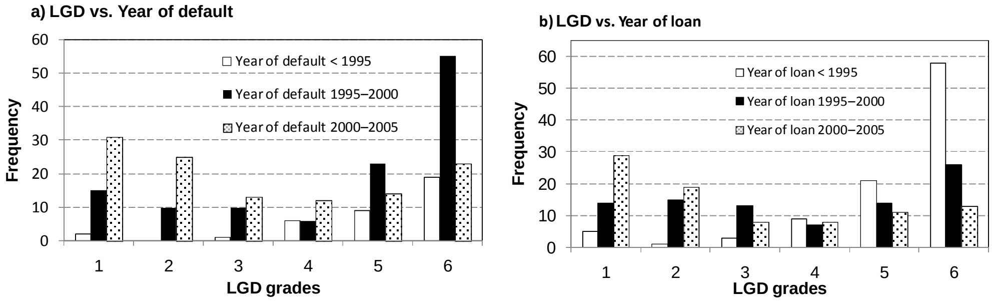
Based on this analysis we can conclude that the workout process is effective until the end of the third year. After the third year of recovery process there are only minor recovered amounts, mainly due to earlier defaulted counterparties with a long recovery period. Therefore, in this essau we considered a file with a workout period longer than three years to be a closed file.

4.3 Discount Rate

In order to calculate LGD for a particular client, ex-post realized cash-flows have to be discounted back to the date of default. Although there is no agreement about which discount rate to choose[46], we consider systematic asset risk class approach proposed by Maclachlan (2005) as a preferred option to derive a risk premium of the discount rate[47]. Risk premium for a particular client is determined by the class of collateral used to secure its claim. This

[45] The methodology is based on a univariate mortality-based approach applied in Dermine and Neto de Carvalho (2006); the calculation does not include internal and external costs.
[46] A summary of LGD discount rate issues can be found e.g. in Chalupka and Kopecsni (2008).
[47] A discount rate can be simply defined as a sum of a risk-free rate and a risk premium.

Figure 2 LGD in Different Periods According to the Year of Default and the Year of the Loan



approach enables to distinguish between various risks based on different sources of net cash flows.

We assigned different risk premiums as follows:

– 0 basis points – cash collateral;
– 240 basis points – residential real estate and land;
– 420 basis points – movables and receivables;
– 600 basis points – commercial real estate, stocks and unsecured loans;
– 990 basis points – guarantees and promissory notes[48].

The effect of application of discount rate based on the systematic asset risk class approach is shown in *Figure 1*.

As a benchmark option we consider a flat risk premium of 940 bps derived from ex post defaulted traded loans from a study by Brady et al. (2007). This study calculates a flat 940 bps premium for bank debt, which seems much more conservative to the systematic asset risk class approach in which only the last category has a higher risk premium.

In our calculations we also tested flat LGD premiums in the range of 0–9%, increasing the premium by 1% resulted in an increase of LGD by approximately the same percentage point. This relatively small effect is due to a relatively short average workout period and significant portion of payments received in early years. The resulting average LGD from applying the systematic asset risk class approach is similar to the flat premium of 5%.

5. Analysis of Typical Risk Drivers

Figure 5 in *Appendix* shows[49] the distribution of the portfolio and average LGD according to factors typically discussed in the literature – EAD, industry, age of the counterparty at the moment of default, collateral value to a loan and year of default. Columns marked in the right axis indicate frequencies and the lines are the average LGDs with values shown in the left axis. In general, the features of our data sample are consistent with the characteristics described in the literature[50].

[48] As a client generally has more than one type of collateral, we weight the risk-premiums based on the percentage of particular collateral out of an exposition at default (EAD) to arrive at a composite discount rate for the client.
[49] A more detailed discussion is presented in the paper by Chalupka and Kopecsni (2008).
[50] The sample also consists of non-closed files whose recovery period is currently longer than effective recovery period. For this reason graphs do not show the definite figures.

5.1 Year of Default and Loan Origination

We hypothesize that year of default and loan originations are important determinants of LGD. Because of small number of observations in individual years we grouped cases in consecutive years into three periods. The first period contains cases which originated or defaulted in 1994 and earlier. The remaining cases are divided in another two groups, the first one containing observations which originated or defaulted in 1995–2000, the other covering the period 2000–2005. All three subsamples are aimed to cover different stages in development of the Czech banking sector or different cycles of the Czech economy.

The first period can be characterized by transformation phase in the Czech economy with low prudence in loan origination process and subsequent problems with high proportion of non-performing loans. In 1990–1994 the government took steps to solve the problem in large banks by the Consolidation program, in which apart from other measures the non-performing loans of some banks were taken over by the government consolidation agency (Jaroš, 2000). As depicted in *Figure 2b*, the loans originated in this period have the highest (worst) LGD compared to the other periods. Because some cases defaulted in the second or third period, the pattern is not so clear when LGD is analyzed according to year of default (*Figure 2a*). Nevertheless, the overall pattern is as expected.

The second and third period represent different macroeconomic cycles of the Czech economy. While the period 1995–2000 recorded a modest growth in the level of real HDP with two negative rates in 1997 and 1998, in 2000–2005 the Czech economy experienced stable growth ranging from 1.9 to 6.3% (ČSÚ 2009). The positive economic conditions appear to translate into relatively low LGD in 2000–2005 (*Figure 2a* and *Figure 2b*).

In line with the results of this analysis we decided to use three periods (two dummy variables[51]) based on the year of default as another explanatory variable. Additionally, we also used a dummy variable for the year of default with the cases which defaulted before 1995.

6. Regression Methodology

This section describes how the regression models employed in our analysis are carried out. Before performing calculations the data is processed in several steps. Firstly, missing data is handled in the following ways:

- Observations with missing data are excluded from the dataset. This option was used in the cases when data necessary for modeling was missing, such as a collateral value.
- Missing data are added, replaced by an average or median value of the portfolio, replaced by a lower or higher cut-off. The age of counterparty is an example.
- Missing data are not replaced, neither those observations are excluded. This data is not essential for modeling; an illustrative factor is an industry where the missing industry was coded as one level along with the data where information on the industry was available.

Outliers are detected on the basis of factor distribution and an expert judgment. An appropriate conservative cut-off value is applied, which is determined on the basis of the median, quantiles and power statistics[52] of the factor.

In order to receive a more powerful model, different types of data need different transformation and adjustment. For continuous factors normalizing is applied after elimination

[51] The first dummy for loans originating in 1995–2000 and the other for loans started after 2000.
[52] The power statistic is measured as an accuracy ratio defined in Sobehart and Keenan (2007).

of outliers. This is useful when the variables such as EAD (with a wide range of 0 to hundreds of millions in currency units) or factors such as age of counterparty (with a narrow range of 0–30 in years) are included in the model.

For some categorical factors a transformation into dummy variables is carried out. For collateral type and industry, grouping similar categories into one class is employed. We have used four collateral type classes based on the risk aspect of the collateral, similar to the classes used in the calculation of the discount rate:

– Class A: low risk – cash, land and residential real estate
– Class B: lower average risk – movables and receivables
– Class C: upper average risk – commercial real estate
– Class D: high risk – securities and guarantees

There are 30 industry groups in the datasets, we grouped them into fewer categories based on two classifications. Additionally, we "compressed" the alternative industry classification even further by having only two groups, the first one containing the "new industries" (Financial Services, Life Sciences and Healthcare, Technology, Media and Telecommunications and Business and Consumer Services) and the rest being the "traditional industries".

6.1 Explanatory Variables Used

From a statistical modeling point of view, factors are divided into continuous factors (can be of any value), categorical factors (can be only of certain number of values) and dummy factors (can be of two values – zero and one). However, from a practical point of view, factors are divided into four main categories. We list the variables that are available for our analysis and in *Table 2* we show those determinants of recovery which are actually used in the models.

Counterparty related factors[53]: industry classification, the company's age at the time of default, year of default, year of company origination, year of loan origination and length of business connection at the time of default.

Contract related factors[54]: type of the contract, exposure at the time of default, interest rate on the loan, tenure and number of different type of contracts.

Collateral related factors: collateral type, collateral value by type, aggregate collateral value, collateral value relative to the EAD, collateral value as a percentage of aggregate collateral value, a number of collaterals and diversification as a number of different collaterals.

Macroeconomic factors[55] are not analyzed, because the dataset is relatively short.

[53] Other possible counterparty related factors are a reason of default; a legal form of the company, a counterparty segment (SME, corporation, small tickets, etc.) size of the company, probability of default one year before default, length of time spent in default, intensity of business connection as distance from the domicile, financial indicators such as profitability, liquidity, solvency, capital market ratio, structure of the balance sheet, stock return volatility.

[54] Other possible contract related factors are seniority of the loan, size of the loan, type of approval and type of monitoring process

[55] Possible macroeconomic factors are default rates, interest rate, GDP growth, inflation rate, industry concentration

6.2 Multivariate Analysis

Three different generalized linear models are applied in order to estimate determinants of the LGD – the first two models employ fractional responses either assuming beta inflated distribution or a more general model estimated by the quasi-maximum likelihood estimator, the last one uses ordinal responses of dependent variable. In all three cases logit and log-log link functions are applied to capture both, a symmetric (logit) and an asymmetric case (log-log). As a benchmark, we firstly used a classical linear regression model to fit the data.

6.3 Models with Fractional Responses Using Quasi-Maximum Likelihood Estimator

Since LGD is a continuous variable typically bound within the interval [0,1], we need to map the limited interval of LGD on the potentially unlimited interval of LGD scores (**β’x**). Generalized Linear Models (GLM) with an appropriate link function can be used for this procedure (McCullagh and Nelder, 1989). The quasi-maximum likelihood estimator (QML)

Table 2 Recovery Rate Determinants Used in the Models (type of variable and expected correlation with recovery rate)

Recovery rate determinants	Type	Correlation
Counterparty related factors		
Age of a counterparty	Continuous	Positive
Length of business relationship	Continuous	?
Year of default before 1995	Dummy	Negative
Year of loan origination 1995–2000	Dummy	Positive
Year of loan origination after 2000	Dummy	Positive
New industries	Dummy	?
Industry not specified	Dummy	?
Contract related factors		
Exposure at default	Continuous	?
Number of loans	Categorical	?
Investment type of loan	Dummy	?
Overdraft type of loan	Dummy	?
Revolving type of loan	Dummy	?
Purpose type of loan	Dummy	?
Collateral related factors		
Collateral value of A relative to EAD	Continuous	Positive
Collateral value of B relative to EAD	Continuous	Positive
Collateral value of C relative to EAD	Continuous	Positive
Collateral value of D relative to EAD	Continuous	Positive
Number of different collaterals	Categorical	Positive

described below does not assume a particular distribution and it is hence more flexible to fit the data than a model using a particular distribution.

If we denote the transformation function as G (.), the logit link using the logistic function is

$$G(\alpha+\boldsymbol{\beta}'\mathbf{x})=\frac{\exp(\alpha+\boldsymbol{\beta}'\mathbf{x})}{1+\exp(\alpha+\boldsymbol{\beta}'\mathbf{x})}$$

the log-log link using the extreme value distribution for dependent variable (the standard Gumbel’s case) is

$$G(\alpha + \boldsymbol{\beta}'\mathbf{x}) = e^{-e^{-\alpha+\boldsymbol{\beta}'\mathbf{x}}}$$

and the complementary log-log link is

$$G(\alpha + \boldsymbol{\beta}'\mathbf{x}) = 1 - e^{-e^{\alpha+\boldsymbol{\beta}'\mathbf{x}}}$$

To estimate this GLM we use the non-linear estimation procedure which maximizes a Bernoulli log-likelihood function[56]

$$L_i(a, \mathbf{b}) = y_i\left[\log G(a + \mathbf{b}'\mathbf{x}_i)\right] + (1 - y_i)\log\left[1 - G(a + \mathbf{b}'\mathbf{x}_i)\right]$$

where a and **b** are estimated values of α and $\boldsymbol{\beta}$.

6.4 Models with Fractional Responses Using a Beta Distribution

Beta distribution has been also used to model LGD, for example in the commercially available application LossCalc by Moody's (Gupton and Stein, 2002). This approach assumes that LGD has a beta distribution. As the values of the distribution itself are bound within the range [0,1] a link function has to be used to map LGD scores onto this interval. As LGD of 0% or 100% are the values which are normally observable and hence have non-zero probabilities p_0 and p_1, we have used the inflated beta distribution[57] with the location, scale and two shape parameters, μ, σ, ν and τ respectively, that allows for 0 and 1 defined by

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \begin{cases} p_0 & if \quad y = 0 \\ (1 - p_0 - p_1)\dfrac{1}{B(\alpha, \beta)} y^{\alpha-1}(1-y)^{\beta-1} & f\ 0 < y < 1 \\ p_1 & f\ y = 1 \end{cases}$$

for $0 \le y \le 1$, where $\alpha = \mu(1 - \sigma^2)/\sigma^2$, $\beta = (1 - \mu)(1 - \sigma^2)/\sigma^2$, $p_0 = \nu(1 + \nu + \tau)^{-1}$, $p_1 = \tau(1 + \nu + \tau)^{-1}$ so $\alpha > 0$, $\beta > 0$, $0 < p_0 < 1$, $0 < p_1 < 1 - p_0$.

The parameter estimates of this GLM model were produced using a maximum likelihood.

6.5 Models with Ordinal Responses

As an alternative technique, we have modeled 6 discrete LGD grades defined earlier using ordinal regression instead of continuous dependent variable used in the previous models.[58] These models might be more appropriate if we expect default cases to be homogenous within the LGD grade but to be different between grades, either by having a different response to factors (different $\boldsymbol{\beta}$) or a different likelihood that a default case will fall into a particular grade (a different intercept). The ordinary regression model using cumulative logit link function is defined as:

$$\begin{aligned} logit\left[P(Y \le j \mid \mathbf{x})\right] &= \log\frac{P(Y \le j \mid \mathbf{x})}{1 - P(Y \le j \mid \mathbf{x})} \\ &= \log\frac{\pi_1(\mathbf{x}) + \ldots + \pi_j(\mathbf{x})}{\pi_{j+1}(\mathbf{x}) + \ldots + \pi_J(\mathbf{x})} \qquad j = 1, \ldots, J-1 \end{aligned}$$

[56] For further technical details and practical applications see Papke and Wooldridge (1996).

[57] This definition of beta inflated distribution is based on (Stasinopoulos et al., 2008).

[58] Since LGD grades are a dependent variable in ordinal regressions, we have used recovery rates (not LGD) as a fractional response to have the same sign of estimated coefficients in both cases.

Each cumulative logit uses all J response categories. A model for logit[P($Y \le j$)] alone is an ordinary logit model for a binary response in which categories 1 to j form one outcome and categories j + 1 to J form the other. A model that simultaneously uses all cumulative logits is

$$logit\left[P(Y \le j \mid \mathbf{x})\right] = \alpha_j + \boldsymbol{\beta}'\mathbf{x} \qquad j = 1,\ldots,J-1$$

Each cumulative logit has its own intercept. The α_j are increasing in j, since P($Y \le j|\mathbf{x}$) increases in j for fixed x, and the logit is an increasing function of this probability. This model has the same effects $\boldsymbol{\beta}$ for each logit and since we consider the same effects in each grade as appropriate, we have used it; we allow for different intercepts.

To fit this special case of the GLM, we let (y_{i1}, …, y_{iJ}) to be the binary indicators of the response for the subject i. The likelihood function (e.g. Agresti (2002)) is

$$\prod_{i=1}^{n}\left[\prod_{j=1}^{J}\pi_j(\mathbf{x}_i)^{y_{ij}}\right] = \prod_{i=1}^{n}\left[\prod_{j=1}^{J}\left(P(Y \le j \mid \mathbf{x}_i) - P(Y \le j-1 \mid \mathbf{x}_i)\right)^{y_{ij}}\right]$$

$$= \prod_{i=1}^{n}\left[\prod_{j=1}^{J}\left(\frac{\exp\left(\alpha_j + \boldsymbol{\beta}'\mathbf{x}_i\right)}{1+\exp\left(\alpha_j + \boldsymbol{\beta}'\mathbf{x}_i\right)} - \frac{\exp\left(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x}_i\right)}{1+\exp\left(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x}_i\right)}\right)^{y_{ij}}\right]$$

It is minimized as a function of different intercepts α_j and common slope coefficients $\boldsymbol{\beta}$ for each LGD grade.

The complementary log-log link for ordinal regression model is defined as

$$\log\left\{-\log\left[1 - P(Y \le j \mid \mathbf{x})\right]\right\} = \alpha_j + \boldsymbol{\beta}'\mathbf{x}, \quad j = 1,\ldots,J-1$$

With this link, $P(Y \le j)$ approaches 1 at a faster rate than it approaches 0.

The log-log link

$$\log\left\{-\log\left[P(Y \le j \mid \mathbf{x})\right]\right\} = \alpha_j + \boldsymbol{\beta}'\mathbf{x}, \quad j = 1,\ldots,J-1$$

is appropriate when the complementary log-log link holds for the categories listed in reverse order.

7. Results of Models

In order to select the most appropriate model, some commonly used procedures are followed. Continuous variables are plotted against LGD (and against LGD grades for ordinal responses) to get "a feel" of the underlying relationship[59]. Similarly, categorical variables are tabulated to form an expectation of a potential relationship. Moreover, a frequency table provides information whether there are enough counts for each cell to estimate the effect reliably[60]. Thirdly, correlation[61] and power statistics (*Table 5 in Appendix*) using each explanatory variable separately is performed to see the effect of each variable independent of the other effects. All potentially plausible variables are then put together in a regression model. Afterwards, the variables not contributing significantly to the explanatory power of

---

[59] If a relationship was not monotonous because of a few extreme observations, an appropriate cut-off was applied. Provided that the relationship was unclear and no transformation of a variable was suitable, the variable was not used in the analysis.

[60] This is important for ordinal regression as we have six grades and we must have enough observations for each explanatory variable in each grade. If necessary, categorical drivers with numerous levels were grouped.

[61] Kendall's tau rank correlation coefficient (Kendall, 1990) is utilised as it does not require normally distributed variables to calculate p-values like parametric Pearson's correlation and it is also preferred to more popular Spearman's rank correlation when a data set is small with large number of tied ranks (Field, 2005).

a model are gradually eliminated from the model (backward elimination) based on Akaike (AIC) and Schwarz information criteria (SIC). For models where continuous dependent variable is estimated and a specific distribution is assumed, worm plots for residuals (van Buuren and Fredriks, 2001), and QQ-plots[62] were utilized to have a visual indication of normality of residuals.

8. *Summary of Findings*

The results in *Table 3* reveal interesting findings[63]. As expected, collaterals of class A and C have a positive and strong effect on recovery rates. These collaterals represent land, residential real estate, cash and commercial real estate. Hence there is no surprise if we find a strong positive relationship, higher proportion of collateral as a percent of EAD increases recovery or likelihood of recovery. Similarly, the year of loan origination has a significant impact on LGD. Both periods (loans originating in 1995–2000 and after 2000) have a negative influence on LGD when compared to loans extended before 1995. As we have already outlined, the period before 1995 was characterized by several problems in the Czech banking sector with a high level of non-performing loans. Weak efficiency of workout process was thus very likely, translating into high average LGD. The situation improved in the period 1995–2000; the low LGD after 2000 is further fuelled by a positive macroeconomic climate. EAD is the next variable significant in almost all the models. The correlation with recoveries is negative and recoveries tend to be lower with the higher loans. The first explanation from the literature could be a weaker link between the management and company results in the case of big companies having high bank loans. This contradicts the assumption that a bank intensifies the enquiry of the creditworthiness and the monitoring of the borrower with high loans. The second explanation could be high leverage of big companies and violation of the absolute priority rule. If a big company defaults, there are many creditors competing for the company's assets, so the recovery for a bank can be small. Length of business relationship appears in half of the models and it has a strong negative effect on the recovery rate. For clients with long relationship, lower recoveries can be explained by the less prudent attitude towards the "familiar" clients. The effect of other variables is not so unambiguous and the results are different for different models, the specifics are discussed for each class of models.

8.1 Classical Linear Regression Model

The linear regression model as the benchmark is the simplest case in which a continuous recovery rate variable is regressed on a linear combination of explanatory variables. The major drawback of this method is that the predicted values can be outside the range [0,1]. Rather surprisingly, the simple linear model is able "to remove" bimodality of the whole sample as shown by residuals which are inside the bounds of confidence intervals of the worm plot, close to normal quantile in the QQ-plot (as depicted in *Figure W1* – see the webpage of this journal).

Apart from the common factors discussed in the previous subsection, there are no other significant variables.

---

[62] In the QQ-plots sample values are plotted against theoretical values predicted by a distribution.

[63] We report only results based on SIC; AIC yields similar results, but allows more variables to be included in the models.

Table 3 Summary of Significant Determinants for All Models (*p*-values are in the brackets)

Model	Exposure at default – EAD	Collateral class A as % of EAD	Collateral class B as % of EAD	Collateral class C as % of EAD	Collateral class D as % of EAD	Age of a counterparty	Length of business relationship	No of different collateral classes	Year of default before 1995	New industries	Industry not specified	Number of loans	Investment type of loan	Overdraft type of loan	Revolving type of loan	Purpose type of loan	Loan origination 1995–2000	Loan origination after 2000
Linear model		0.469 (0.000)		0.491 (0.000)			-0.214 (0.001)										0.236 (0.000)	0.389 (0.000)
Fractional response (logit link)		2.518 (0.001)		2.677 (0.001)			-1.235 (0.013)										1.288 (0.000)	1.968 (0.000)
Fractional response (log-log link)		1.784 (0.002)		1.688 (0.005)													0.797 (0.000)	1.290 (0.000)
Fractional response (complementary log-log link)		1.538 (0.000)		1.640 (0.000)			-0.951 (0.012)										1.035 (0.000)	1.458 (0.000)
Fractional response beta (logit link)	-1.228 (0.001)			0.974 (0.018)							-0.766 (0.000)						1.107 (0.000)	1.696 (0.000)
Fractional response beta (log-log link)	-0.628 (0.000)			0.723 (0.014)							-0.442 (0.000)						0.619 (0.000)	1.048 (0.000)
Fractional response beta (complementary log-log link)	-1.086 (0.000)										-0.637 (0.000)						0.936 (0.000)	1.350 (0.000)
Ordinal response (logit link)	-2.228 (0.003)	2.787 (0.000)		2.673 (0.000)			-1.117 (0.006)				-0.909 (0.001)						1.676 (0.000)	2.643 (0.000)
Ordinal response (complementary log-log link)	-1.488 (0.008)	1.321 (0.000)		1.481 (0.000)			-0.746 (0.006)				-0.540 (0.004)						1.034 (0.000)	1.478 (0.000)

8.2 Models with Fractional Responses Using Quasi-Maximum Likelihood Estimator

Applying the asymmetric log-log link function[64] to regress continuous recovery rates estimated by quasi-maximum likelihood yields slightly better results than the logit link function. A period of loan origination, along with collateral of class A and C are the major determinants of LGD. EAD which is strongly negatively correlated to the collateral classes appears not to be a significant factor[65]. Length of business relationship is also significant in two of the three models.

8.3 Models with Fractional Responses Using a Beta Distribution

The fit employing the inflated beta distribution and the logit link is very similar to the log-log link[66]. Logit fits the data reasonably (*Figure W2* – see the webpage of this journal), although the fit is worse for higher recovery rates.

Compared to the other models, collateral of class A (residential real estate) is not significant. The model identifies a non-specified industry as an additional important factor. These singularities can be explained by the assumption of beta distributed errors. The other factors are in line with the previous results.

8.4 Models with Ordinal Responses

As for ordinal response models, increasing the length of business relationship measured as a period between the date of bank account opening and the date of default decreases the likelihood of low LGD, similarly to the classical linear model. Also consistent to previous results, a loan originated after 1995 has lower probability of high LGD. Correspondingly to the beta distribution model, a non-specified industry is a statistically significant determinant of LGD also.

9. Comparing Goodness-of-Fit of the Models

Goodness-of-fit summary measures offer an overall indication of the model fit. Out of parametric performance measures, mean square error (MSE), mean absolute deviation (MAD) and correlation between observed and modeled LGD have been evaluated to compare predictive power of suggested models. The outcomes from correlation are listed in *Table 4*.

MSE, MAD and correlation coefficient measure model performance parametrically and are sensitive to model calibration. In contrast, a power statistic is a non-parametric measure that focuses on the ability to discriminate "good" outcomes from "bad" outcomes without being sensitive to the calibration. It indicates a model's power and ranges from zero to one. It provides information about different aspects of model performance not registered by the above mentioned measures.

The power statistic[67] is commonly used for PD models where there are only two possibilities – a default or a non-default. However, in LGD models the dependent variable is continuous. In order to apply the power statistic, it is necessary to define what is considered as

---

[64] Determinants of recovery rates for the logit and complementary log-log function are shown in the summary table in *Appendix*.

[65] However, a larger sample would enable to distinguish the effects of highly correlated factors better and we expect EAD to have an impact.

[66] The log-log and complementary log-log links are again shown in the summary table in *Appendix*.

[67] The Power statistic is described for instance in Gupton and Stein (2002).

Table 4 Performance Measures – Kendall's Tau Rank Correlation and Power Statistic

Model	Correlation	*P*-value	Power Statistic	SE
Linear model	0.423	0.000	63.6%	4.2%
Fractional response (logit link)	0.382	0.000	60.9%	4.7%
Fractional response (log-log link)	0.388	0.000	61.3%	4.5%
Fractional response (complementary log-log link)	0.362	0.000	58.3%	4.3%
Fractional response beta (logit link)	0.423	0.000	63.6%	4.1%
Fractional response beta (log-log link)	0.416	0.000	62.9%	4.5%
Fractional response beta (complementary log-log link)	0.416	0.000	61.4%	4.1%
Ordinal response (logit link)	0.429	0.000	65.6%	4.5%
Ordinal response (complementary log-log link)	0.425	0.000	65.3%	3.8%

"good" and what is considered as "bad". In the paper by Chalupka and Kopecsni (2008) three different alternatives are proposed.

Alternatively, ordinal power statistic can be applied to LGD models. This statistic (*Table 4*) measures the ability of a certain model to differentiate between any numbers of rating categories in the correct order; hence it is used particularly for the ordinal response models. The results are similar to the previous case.

The linear model, fractional models using quasi-maximum likelihood and ordinal response models perform similarly and relatively well in absolute terms.

9.1 Robustness Tests

The studied sample was divided into two halves based on three different criteria to test robustness of our results. Firstly, the sample was ordered according to the date of default and each second observation was put into one subsample and the rest of observations into the other subsample. Secondly, observations were arranged according to LGD grades, and again every second observation was put into one sample and the rest into the other. Thirdly, observations were randomly assigned into two subsamples. For each division we re-performed calculations of the models and checked whether subsamples had the same significant drivers as the whole sample. We can conclude that results of subsamples are consistent with the whole sample; the major drivers, as well as the signs of coefficients remained the same. Only magnitudes changed somewhat; especially in the case of the random division.[68]

Additionally, we have verified the models' out-of-sample predictive power. Employing the results of the subsamples in each of the three divisions, the estimated coefficients were applied to predict LGD in the other half of each subsample. The resulting power statistics yield satisfactory results ranging from 54 to 60%.

[68] The estimated coefficients of the subsamples are available from the authors upon request.

10. Conclusions

In this essay, various statistical models were applied to test empirically the determinants of LGD. We have found that the main drivers are the period of loan origination, relative value of collateral, loan size and length of business relationship. Different models provided similar results. As for the different links in more complex models, log-log models in some cases performed better, implying an asymmetric response of the dependent variable. All models performed relatively well when the overall fit of the different models was assessed. However, the models with the commonly assumed beta distribution achieved slightly worse results and hence are not deemed optimal for our data.

From a policy perspective, the essay provides evidence that workout LGD is a viable option in the credit risk estimation, despite various methodological difficulties. In this study, we try to provide a reasonable detail of various issues to be tackled and propose methodological alternatives to cope with these issues.

There are several ways in which our research can be improved. Firstly, a similar study can be done on a larger sample of data and hence some of the effects could be estimated more precisely. Secondly, correlation of recovery rate and probability of default, effects of macroeconomic factors and downturn LGD should be thoroughly analyzed for a complete LGD model.

APPENDIX

Figures

Figure 3 Average Cumulative Recovery Rate

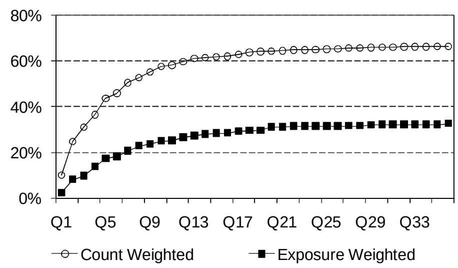


Figure 4 Average marginal recovery rate

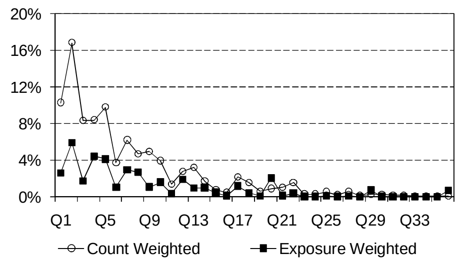
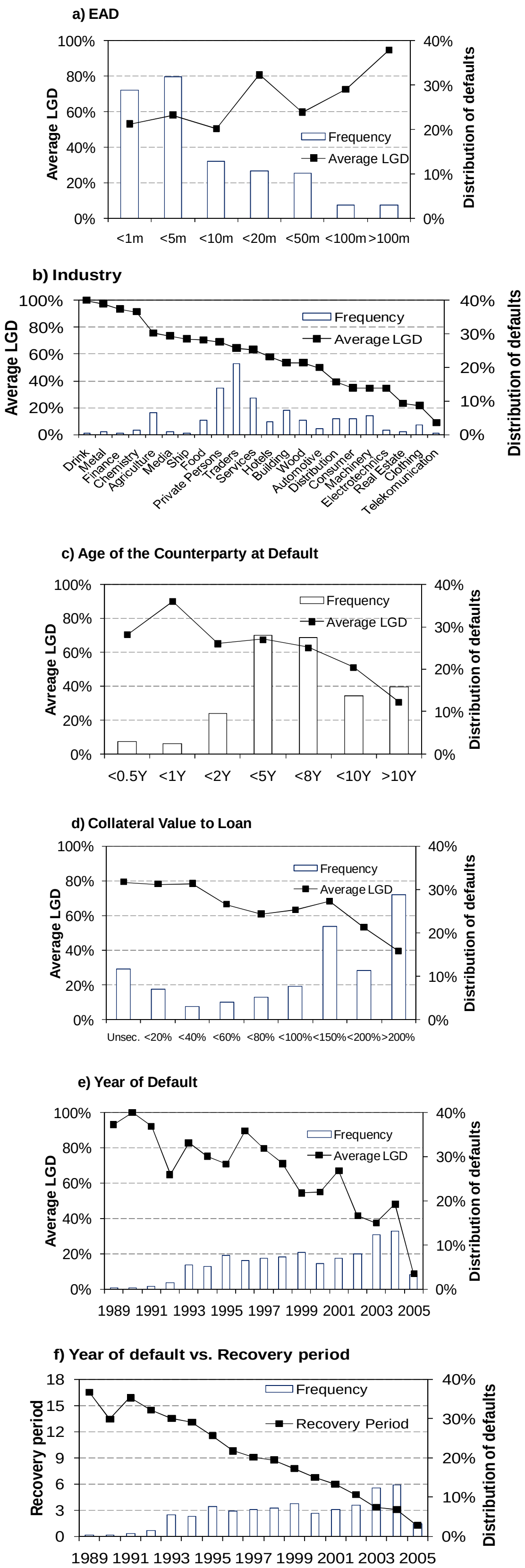


Figure 5 Characteristics of Typical Risk Drivers



TABLES

Table 5 Power statistic of individual factors

Factor	Power statistic	SE
Exposure at default – EAD	17%	6%
Collateral class A as % of EAD	-18%	6%
Collateral class B as % of EAD	-5%	4%
Collateral class C as % of EAD	-18%	6%
Collateral class D as % of EAD	-2%	6%
Age of a counterparty	-30%	6%
Length of business relationship	20%	6%
Number of different collateral classes	-29%	5%
Year of default before 1995	14%	3%
New industries	0%	3%
Industry not specified	5%	5%
Number of loans	-3%	5%
Investment type of loan	-2%	5%
Overdraft type of loan	-9%	4%
Revolving type of loan	-11%	4%
Purpose type of loan	5%	6%
Loan origination 1995–2000	-7%	6%
Loan origination after 2000	-33%	5%

REFERENCES

Acharya VV et al. (2003): Understanding the Recovery Rate of Defaulted Securities. *Center for Economic Policy Research Discussion Papers,* no. 4098.

Agresti A (2002): *Categorical Data Analysis.* 2nd Edition. Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, New Jersey.

Altman E, Kishore V (1996): Almost everything you want to know about Recoveries on Defaulted Bonds. *Financial Analyst Journal*, 52(6):57–64.

Asarnov E, Edwards D (1995): Measuring Loss on Defaulted Bank Loans: A 24 year study. *Journal of Commercial Lending*, 77(7):11–23.

Basel Committee on Banking Supervision (2006): *International Convergence of Capital Measurement and Capital Standards.* Bank for International Settlements, Basel.

Brady B et al. (2007): *Discount Rate for Workout Recovery: An Empirical Study*. Available at SSRN: http://ssrn.com/abstract=907073.

Buuren S van, Fredriks M (2001): Worm plot: a simple diagnostic device for modeling growth reference curves. *Statistics in Medicine*, 20(8):1259–1277.

Carey M (1998): Credit Risk in Private Debt Portfolios, *Journal of Finance*, 53(4):1363–1387.

Carty L, Lieberman D (1996): Defaulted Bank Loan Recoveries, *Global Credit Research, Moody's.*

Chalupka R, Kopecsni J (2008): Modeling Bank Loan LGD of Corporate and SME Segment: A Case Study. Charles University in Prague, Faculty of Social Sciences, *IES Working Paper*, no. 27.

Dermine J, Neto de Carvalho C (2006): Bank Loan Losses-Given-Default, A case study. *Journal of Banking and Finance*, 30(4):1219–1243.

Field A (2005): *Discovering Statistics Using SPSS.* 2nd Edition. Sage Publications Ltd, London, UK.

Franks J et al. (2004): A comparative analysis of the recovery process and recovery rates for private companies in the UK, France and Germany. *Risk Solutions*, *Standard & Poor's.*

Frye J (2000): Depressing recoveries. *Risk*, 13(11):108–111.

Grossman R et al. (1997): Syndicated Bank Loan Recovery Study, *Fitch Research.*

Grossman R et al. (2001): Bank Loan and Bond Recovery Study: 1997–2000, *Fitch Loan products special report.*

Grunert J, Weber M (2005): Recovery Rates of Bank Loans: Empirical Evidence for Germany. University of Mannheim, *Working Paper*

Gupton GM et al. (2000): Bank Loan Loss Given Default. *Global Credit Research, Moody's.*

Gupton GM, Stein RM (2002): LossCalc: Model for predicting loss given default (LGD). *Global Credit Research*, *Moody's.*

Hurt L, Felsovalyi A (1998): Measuring loss on Latin American defaulted bank loans, a 27-year study of 27 countries. The *Journal of Lending and Credit Risk Management*, 81(2):41–46.

Jaroš T (2000): Credit Crunch (teorie a česká praxe). *Diploma thesis*, Charles University in Prague.

Kendall MG (1990): *Rank Correlation Methods*. 4th Edition. A Charles Griffin Book, London, UK.

Maclachlan I (2005): Choosing the Discount Factor for Estimating Economic LGD. In: Altman E, Resti A, Sironi A (Eds.): *Recovery Risk, The Next Challenge in Credit Risk Management*, Risk Books , 285–305.

McCullagh P, Nelder JA (1989): *Generalized Linear Models*. 2nd edition. Chapman and Hall, London.

Papke LE, Wooldrigdge JM (1996): Econometric Method for Fractional Response Variables with an Application to 401 (K) Plan Participation Rates. *Journal of Applied Econometrics*, 11(6):619–632.

Renault O, Scaillet O (2004): On the Way to Recovery: A Non-parametric bias free estimation of recovery rate intensities. *Journal of Banking and Finance*, 28(12):2915–31.

Rigby R.A. and Stasinopoulos D.M. (2005): *Generalized additive models for location, scale and shape, (with discussion)*, Applied Statistics, 54(3):507–554.

Schuermann T (2004): What do we know about Loss Given Default? In: Shimko, D.: *Credit Risk Models and Managament*, Risk Books, London, UK.

Seidler J, Jakubík P (2009): Implied Market Loss Given Default in the Czech Republic: Structural-Model Approach. *Finance a úvěr-Czech Journal of Economics and Finance*, 59(1):20–40.

Thornburn K (2000): Bankruptcy auctions: costs, debt recovery and firm survival, *Journal of Financial Economics*, 58(3):337–368.

JEL Classification: C14, C44, C61, L8
Keywords: quality adjusted DEA, branch performance, scale efficiency, return to scale

Improving Service Performance in Banking using Quality Adjusted Data Envelopment Analysis

Abstract
The goal of this research is to describe the application of data envelopment analysis (DEA) to the performance evaluations of bank branches. Special attention is focused on how to incorporate the quality dimension into branch efficiency. DEA will apply to a set of micro-data from a Czech commercial bank branch network. In the banking sector, providing services quality is one of the key focuses. Therefore, the quality dimension should be incorporated into the DEA model. The goal of the quality adjusted DEA model is to identify best practice branches that work efficiently and at the same time provide services with high quality. This model avoids productivity-quality tradeoff, which is present by the standard DEA model. The quality of services is measured by customer service, mystery shopping and calls, client information index, retention, and client product penetration. Main determinants of efficiency and quality level are branch size and region via purchasing power.

1. Introduction

The service economy consists of a large proportion of developing countries' economic activity, and its growing development has raised the importance of maximizing organizations' productivity. Organizations are searching for a benchmarking technique to identify best practices in supporting their decisions in order to receive effective utilization of resources.

Organizations frequently use simple aggregate ratio analysis as a measure of their productivity. According to Camanho and Dyson (1999), Beamon (1999) and Reynolds (2004), ratio analyses are not sufficient[69] to measure productivity for organizations using multiple resources and providing multiple outcomes. To evaluate such organization's performance, it needs more sophisticated, non-parametric benchmarking methods. Further advantage of the non-parametric method is the fact that it does not require specification of the production function form, which is required by parametric methods. Therefore, managers are interested in supporting their decisions through the use of academic methodologies, Brazdik and Druska (2005).

The difficulties are further enhanced when the relationship between the inputs and outputs are complex and involve unknown tradeoffs as it is argued by Zhu (2009). It is particularly difficult for service industries to improve productivity and find substantial cost saving without sacrificing service quality. There are many subjective factors that affect productivity and service quality. A good example of an industry in which the quality of services is an important issue is the banking sector. In banks, such subjective factors influencing productivity include customers' needs, behavior in receiving the service, service provider's judgment, and skills in providing service.

This research proposes a methodology to describe the application of a non-parametric benchmarking method, data envelopment analysis (DEA), for performance evaluations of bank branches. The advantage of DEA is its ability to measure the relative efficiency of branches by simultaneously analyzing their multiple resources with multiple outcomes. Based on the literature, there are proposals and applications for three different methods to incorporate the quality dimension into branch efficiency. Empirical results are discussed using a set of micro-data from a Czech commercial bank branch network (the bank).

[69] The drawback of ratio analysis is its univariate nature.

In the banking sector, providing services quality is a key focus. Therefore, the quality dimension should be incorporated into the DEA model. The goal of the quality adjusted DEA model is to identify best practice branches that work efficiently and at the same time provide high-quality services. This model avoids productivity-quality tradeoff, which is present in the standard DEA model. The quality of services is measured by customer service, mystery shopping and calls, client information index, retention, and client product penetration.

At the end of last century in the Czech Republic, banks focused solely on the growth of new business volume and on new customer acquisition. Recently, due financial crises, however, they are encouraged to optimize their resources as well. They identified that with cost optimization, it is possible to receive further improvements. Moreover, it becomes more important to maintain customer retention to have valuable customers via selling more products to existing customers. Success is only possible through high-quality service.

The essay is organized as follows. The following section contains a brief literature review on DEA research with special attention on studies with quality measurements. Section three discusses the details of how bank branch network provide services for clients, focusing on the input and output specification according to the motivation system and long-term strategy. Section four gives an overview of the theoretical DEA framework. It also specifies three different methods to incorporate the quality dimension. The fifth part summaries the results obtained by non-parametric methods, and the last section draws conclusions with policy implications.

2. *Literature review*

The original CCR model (Charnes, Cooper, and Rhodes, 1978) is the first DEA model that evaluates technical efficiency in a multiple-input and multiple-output framework. After that, the DEA technique has become a widely used approach for efficiency analysis in many public and private sectors[70] like universities, non-profit organizations, hospitals, and banks. Emrouznejad et al. (2008) presents the most extensive listing of DEA research, covering its 30 years of history, theoretical developments, and empirical applications.

During the 1990s, DEA method has been frequently used to evaluate the performance of financial and banking organizations. Oral and Yolalan (1990) and Oral et al. (1992) investigate in their empirical studies the relationship between branch efficiency and its profits. Further, Giokas (1991) was the first to evaluate branch efficiency with respect to size. It was followed by studies by Drake and Howcroft (1994), Tulkens (1995) and Schaffnit et al. (1997). Drake and Howcroft (1994) reported that more efficient branches had lower cost-income ratios. They utilize data from a UK bank branch network. Schaffnit et al. (1997) use data from a large Canadian bank to show that branch efficiency has a positive effect on profit. An efficiency review of financial institutions is described in Berger and Humphrey (1997).

Several studies have solely analyzed the efficiency of bank branches. Their comprehensive branch performance review was published by Camanho and Dyson (1999), where authors also describe an application of DEA used in the performance assessment of

[70] There are several McKinsey working papers dealing with operational excellence in several industries, which are based on McKinsey 7S framework. The main source of academic work on the 7S model has to be the papers of Waterman et al. (1980, 1982). One of the elements of this framework is the strategy, which is discussed in Lynch (2005) more in details.

Portuguese bank branches and show how DEA can complement profitability measures. Later, Sevcovic et al. (2001) focus on the problem of a suitable choice of efficiency measures, and they show how these measures can influence results. A dataset was provided by one of the leading banks in Slovakia. Most recently, Irsova (2009) compares two methods in bank efficiency, the stochastic frontier approach and DEA, which are supported by the meta-regression part including several studies on the United States and transitional countries.

Above mentioned DEA papers are dealing with efficiency from general perspective. The quality dimensions are getting part of efficiency analysis later on and they are discussed in the following section. Callen (1991) early identifies that most DEA studies do not consider the quality of services or products. Excluding quality can result in applying methods that increase efficiency by reducing quality. Quality in many areas is critical, but is not included in DEA models. These studies assume that quality is homogenous among investigated units or quality is independent of efficiency. Only few DEA studies explicitly address quality.

First, Sherman and Ladino (1995) used DEA to substantially improve its branch productivity and profits while maintaining service quality. Athanassopoulos (1997, 1998) in his DEA studies of a Greek bank branch network considers the relationship of DEA productivity scores with quality. Bank branch operations are demonstrated by the effort made by management to pursue the banks' corporate objectives, which consist of the tangible part described by the operating efficiency and the intangible part characterized by the quality of the provided services. Effort effectiveness is estimated by embodying three quality dimensions—approachability, location, and telephone service. These independent quality measures are developed based on customer surveys and the statistical relationship between quality and the outputs in the DEA model. The study, however, does not combine operating efficiency and quality into the effort effectiveness; the DEA scores are calculated without quality adjustment.

Further, Soteriou and Zenios (1999) gain superior insight by simultaneously analyzing the design of operations together with the quality of the provided services and profitability, rather than by benchmarking these three dimensions separately. Other measures of service quality in banking are discussed in Athanassopoulos and Giokas (2000).

Above mentioned results request research to find ways to properly combine quality and efficiency Sherman and Zhu (2006). Only few DEA studies explicitly address quality, and those that consider it have not fully adjusted for quality. This essay suggests how to enhance and fully adjust the standard DEA method by quality dimension. It also evaluates how the results change due to different quality measurements.

3. *Banking sector providing services for clients*

Each DEA model is constructed to solve a concrete requirement. Therefore, formulation of DEA problems require an understanding of the production process, assumes deep industry knowledge, organization strategies with key motivation elements, as well as identification of the appropriate input (resources) and output (outcomes) factors.

3.1 Strategy

The bank's long term goal is to grow business profit through selling deposit and loan products[71]. However, it is hard to manage new volumes. Branches have only limited control over new volumes that are determined by external factors, mainly by sales the potential of the region. Earning long-term profit growth in such a competitive industry is possible only by also focusing on other essential components. Therefore, banks' recent strategy has also focused on rationalization of existing branches, cost optimization, and redeployment of surplus staff to new ones (step I). In addition, special focus is on the quality of service is provided to clients in order to meet their needs (step II).

The key activity of the bank is based on the operation of the branch network, which represents the main contact point between customers and management of the bank. Officers in branches sell various types of deposit and loan products to generate profits. Therefore, branches and their employees are service providers. They have to understand customers' needs, sell appropriate products, and provide high-quality services in order to receive loyalty and make customers more valuable. In order to operate efficiently, branches need to solve not only cost minimizing strategy, but also attract customers by offering high-quality services.

3.2 Branch network

Organization within the branch network and production process is a follows. There are large, medium, and small branches based on the number of employees[72]. In branches, there are four types of client officers. Universal client officers are responsible for teller activities (standard transactions such as deposits, withdrawals, and bank checks). Officers deal with general and simple customer queries (opening bank account, travel insurance, payment, and credit card administration). Advisers deal with more complex activities according to its specialization.[73] Personal and firm bankers advise the most valuable clients, caring for their product portfolios. Finally, branch directors manage client officers and attend to the most important issues.

3.3 Performance evaluation

The bank uses two different methods to analyze the performances of its branches. The first is based on the volume of new business[74] within a year. Specialists measure savings and loan volumes separately on retail and firm portfolios. Savings contain all major deposit and investment products—current and savings accounts, term deposits, investment funds, pension funds, housing savings, and single and regular life insurance. Loans include all products with loan characteristics—consumer loans, credit cards, overdraft, housing loans, mortgages, investment loans, revolving, factoring, and leasing. Measured values are compared with the

---

[71] It is not a common practice that banks measure RAROC or RORAC on the level of an individual retail and SME branch. Profit before tax that includes credit provisions is, however, frequently used as a KPI on a branch level. As regards income, it is mostly broader and consists of NII from assets and liabilities, net fee and commission income and income from financial operations.

[72] Small branches have up to 10 employees, medium branches have up to 20 employees, and large branches have more than 20 full time employees.

[73] For example, retail investment advisers offer services in investments of funds, and firm loan advisers help firms find the most appropriate loan for their business.

[74] New business volume is measured as the difference between a stage at the end of the year t and at the end of the year $t-1$.

plan determined by top management and then the weighted averages of ranks in each category describe the final branch performance rank.

The second method emphasizes branch activities. Activities are defined as the number of sold products (investment, housing loans and mortgages, non-specified loans, and SME loans) within a year and net increase in the number of active clients per number of branch employees. It has two dimensions—actual stage and growth. Branches receive rankings in each category. Some branches are new to high growth in these factors but have a poor actual stage. Most of the branches already have a very good actual stage, but they also have slow growth in several indicators. The best branches have a very good actual stage and very high growth in most of variables, and they serve as best practice branches for others. On the other hand, opposite branches need a special focus because they have a poor actual situation and poor growth. It is necessary to discover the reasons for external factors or poor management.

The advantage of the second method is that it better reflects officers' effort. While in a city with high purchasing power, on one investment deal, a branch can receive new volume of several million CZK, but effort from the officers is the same as in a small village for the investment in a volume of several thousand CZK. The more active officers have a branch with a better ranking through this method.

Results of both methods are entering the motivation system for client officers, as their bonuses depend on them. The motivation system should reflect the company's long-term strategy and should fairly reward the employees' efforts in this direction. The management of the bank identified that long-term strategies should not be based solely on financial indicators (first method), but also how it is received and how much effort is needed (second method).

3.4 Standard DEA application

The above mentioned methods do not take into consideration the employee structure of branches and external factors such as the region's purchasing power. They are unable to discover the source of inefficiency and how to deal with it in order to attain an efficient environment. Furthermore, even the second method does not take into account the quality of services. Considering activity alone is only a short-term issue. To maintain excellent long-term results, it is necessary to know more about the clients and their needs, increase product penetration, and have high client retention. This can only be reached through high-quality service. Models excluding quality dimensions assume that quality is homogenous through the branches.

Appropriately defined, the DEA model is able to solve some of the above mentioned weaknesses in the current performance measurements. The goal of the proposed standard DEA model is to find out the optimal resource allocations and minimize branch costs. It contains the following input and output factors.[75]

The best indicator for branch resources is branch size. To estimate branch size, the number of branch employees is used because personal costs are a major part of overall branch costs. In total, there are three input (resource) variables: UCO FTE—universal client officers and a branch director, Retail FTE—advisors and personal bankers for retail clients[76], and SME FTE—advisors and firm bankers for non-retail clients. FTE means the number of full

[75] Full definitions of these factors are in *Appendix, Table A1*

[76] Retail are all physical persons. SME are firms and physical entrepreneurs with annual turnover up to EUR 10 million. Corporate clients are above this threshold and are not be included in DEA analysis.

time employees, which is adjusted to account for maternity leave, holidays, part-time workers, illness, and training. It explains how many full-time employees were present in a certain period in the branch.

There are four output[77] measures—retail loans (consumer loans, credit cards, overdraft, housing loans, and mortgages), retail savings (current and saving accounts, term deposits, investment funds, pension funds, housing savings, and single and regular life insurance), SME loans (overdraft, investment loans, revolving, factoring, and leasing), and SME deposits (current and saving accounts, term deposits, and investment funds).

There are three different specifications. First, output factors are measured as new volumes within a year. This method favors branches in large cities, where clients are more likely to invest higher amounts, buy mortgages with higher values, or where bigger firms that are searching for large investment loans are located. Second, output factors are measures as above, but the volumes are divided by each branch's town purchasing power index[78] in order to eliminate the effect of that external factor—discrimination of otherwise equally good officers employed in the region with low purchasing power. Results expected to be more homogenous.

Third, output factors are measured as the number of new sold products, which reflects the client officers' activities. The motivation behind this specification is that to measure what client officers are able to influence. They are able to use their services to influence certain clients to buy a mortgage at their bank instead of at a competitor's, but they are not able to influence the volume of the mortgages. We believe that client officers are able to influence the number of sold products with their services. Therefore, output in the model is measured by the number of sold products, which should be a result of high quality service, number of meetings with clients, and other officer efforts.

3.5 Quality adjusted DEA application

The standard DEA model, however, is not quality adjusted and assumes that quality is homogenous among branches. However, this is not our case. Therefore, the basic DEA model is enhanced by the quality dimension. There are four quality measurements: service quality index, client information index, product penetration index, and client retention. The essay demonstrates short-term interaction among service quality and operating branch efficiency[79].

The banking market is very competitive, and banks can no longer grow rapidly just through new acquisitions. It becomes more important to attain valuable customers through selling more products to existing customers and maintaining customer retention. Both of them are only possible through high-quality service. Therefore, the bank's actual strategy is focused more on quality.[80] Service quality is considered very important because of high competition in the market and the value of retaining customers.

[77] Manager business objectives (MBOs) are building upon 4 elements in the bank: profit including risk element, volumes as a proxy for market share, activities of the sales force (measured through number of meetings and sales in pieces) and the last is quality of services or customer satisfaction. Managers have higher weigh on the profit KPI and employees more on activities and customer satisfaction.

[78] It is a complex index that takes into consideration several external factors, such as unemployment, cost of living, etc., and therefore it is the most appropriate indicator.

[79] In reality, however, service quality has a substantial effect on branch efficiency rather on long-term prospects. Due to a lack of data for long-term forecast, we estimate only short-term interactions.

[80] A recent situation in the financial markets further confirms that quality of services is an important element. Customers require more explanations about investment products. They put their savings where they feel more

If you know more about your customers, you can better manage customer relationships and you can have a better idea of what customers need and their interests. Consequently, you can sell them more appropriate products. Customers will be more satisfied, will be more loyal, and will return. Their churn will be lower when the bank utilizes long-term business growth. Therefore, it is important to measure and be under the control of these indicators. There are several measurements of how banks currently control and try to increase service quality.[81]

First, the bank creates a service quality index, which has three parts: customer service, mystery shopping, and mystery call. Each of them is focused on the quality of officers' willingness and proficiency. Customer service is a certain type of meeting between a client officer and an existing customer in order to maintain the customer relationship, to identify customer needs, and finally, to increase the probability to sell a new product. Client officers should proactively address clients and thoroughly prepare in advance for the meeting based on available information about clients, their past needs, and interests. Correctly done, customer service meetings encourage branch sales results. Therefore, client officers are motivated to arrange meetings with clients, and they have to fulfill a certain number of customer service meetings with their customers. Fulfillment of the branch plan is given by a score for customer service.

Mystery shopping is evaluated by a mystery shopper posing as a customer from an external consulting firm. A mystery shopper visits a branch in order to receive information about certain types of products, advice in investment, or to receive mortgages and other loan products. During the visit, he evaluates several aspect of service, mainly quality and correctness of provided information and ability to communicate well with a customer. In advance, the mystery shopper is educated on what high-service quality looks like, correct answers, and what he can and cannot do. After the visit, he fills out an evaluation form, indicating which tasks were fulfilled and which are not. Based on these figures, a branch receives another quality score. Each time the mystery shopper visits, he focuses on a different topic, product need, or client officer's seniority level. A mystery call is very similar to mystery shopping, but in this case, the mystery customer calls the branch. Client officers have to give a correct answer on the counterparty question and offer a personal meeting at the branch. Based on the behavior of the officer and the accuracy of the answers, the branch receives a third quality score. Finally, these three quality scores are put together to create a complete service quality index, which is evaluated on a monthly basis.

Second, there is the client information index. Client officers should put information about clients into the internal system, like phone numbers, emails, ID cards, education, job, incomes, and expenses. The client information index expresses how much information is recorded in the internal system about the branch customers. Of course, there is a causality problem between owned products by customers and available customer information at the banks. Selling certain products like mortgages is conditioned to deliver a lot of special information from customers regardless of how active the officers are. Basic products such as current accounts or savings do not need any additional information from clients to deliver. However, more active officers should receive more information from customers regardless of which products they own.

---

safety or search for more appropriate mortgages that fit their needs and are flexible. If they do not see high-quality service or confidence, they quickly change banks or just leave their savings at home. Customers start to value the quality of services.

[81] A full definition of quality indicators are in *Appendix, Table A2*

Third, there is the product penetration index.[82] It is very important to have customers with more than one product. Customers with more products are less likely to leave the bank. Therefore, the bank's long-term objective is good cross selling. Client officers are motivated to sell mortgages together with life insurance and possibly credit cards. With customers who just open a current account, officers attempt to sell them debit cards with advantaged travel insurance for the whole family. In this way, there is lower probability that customer is going to conduct business with competitors; customers will be more loyal, and the bank will generate higher profits.

Forth, there is client churn or retention. Monitoring the client churn (and the reasons for leave) is inevitable. Active customers are the most important assets. Clients who are dissatisfied with quality service are more likely to leave. Therefore, client retention is a good estimation of service quality.

All of these aspects contribute to the overall performance of the branches, and they are controlled fully (SQI index) or partially (client information index, product penetration index, and client retention) by client officers. Therefore, they should be incorporated into the DEA model. This chapter demonstrates interaction among service quality and operating branch efficiency.

4. *Methodology*

Section three gives an overview of the theoretical DEA framework. It also specifies three different methods of how quality is possible to incorporate. The proposed methodology follows Sherman and Zhu (2006) using real data from a branch network.

4.1 Standard DEA framework

DEA is a linear programming technique for measuring the relative efficiency of a homogenous set of Decision Making Units (DMUs, which in this study are branches) by analyzing their multiple inputs with multiple outputs. It identifies a subset of efficient best practices branches through a piecewise linear envelopment of observed data. For the rest of the branches, the magnitude of their inefficiency is measured by the distance from the envelope of best practice branches. DEA derives a summary measure of efficiency for each branch. It also derives what would be the optimal combination of input and output for inefficient branches. This means that DEA allow us to not only say whether a certain branch is efficient or not, but also which inputs and outputs are the sources of inefficiency.

The original CCR model (Charnes, Cooper and Rhodes, 1978) is the first DEA model, which evaluates technical efficiency in a multiple input and multiple output framework. The CCB model assumes constant return to scale, i.e., outputs increase by the same proportion as inputs. This assumption is appropriate only when all DMUs operate at an optimal scale. In general, however, this is not true in many sectors. The banking sector is a good example because there is a significant difference between small and large branches' activities. This indicates the existence of a variable return to scale. Therefore, in this essay, the BCC model (Banker, Charnes and Cooper, 1984) is applied to estimate efficiency, which assumes a variable return to scale. During the 1990s, the DEA technique became a widely used approach

[82] The penetration index is based on Finalta definitions. Each of the following products is counted with equal weight: current account, saving account, term deposit, investment fund (including pension savings), life insurance, consumer finance (including consumer loan, overdraft, and credit card), and mortgages.

for efficiency analysis in many public and private sectors, such as universities, non-profit organizations, hospitals, and banks. We use input-oriented[83] (cost-minimizing) BCC models; the envelopment model and its dual specifications are demonstrated in *Table 1*.

Table 1 Input-oriented BBC model with variable return to scale, Envelopment model and its dual problem Multiplier model

	Envelopment model	Multiplier model
Input – oriented DEA	$\min \theta - \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+)$ $s.t.$ $\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta x_{io}, i = 1,2,...,m$ $\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro}, r = 1,2,...,s$ $\sum_{j=1}^{n} \lambda_j = 1$ $\lambda_j \geq 0, j = 1,2,...,n$ $s_i^- \geq 0, i = 1,2,...,m$ $s_r^+ \geq 0, r = 1,2,...,s$	$\max \sum_{r=1}^{s} \mu_r y_{ro} + \mu$ $s.t.$ $\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + \mu \leq 0, j = 1,2,...,n$ $\sum_{i=1}^{m} v_i x_{io} = 1$ $\mu_r \geq \varepsilon > 0, r = 1,2,...,s$ $v_i \geq \varepsilon > 0, i = 1,2,...,m$
Efficient target	$x_{io}^{/} = \theta^* x_{io} - s_i^{-*}, i = 1,2,...,m$ $y_{ro}^{/} = y_{ro} + s_r^{+*}, r = 1,2,...,s$	

In the model, there are n branches, where every branch$_j$, $j = 1,2,\ldots,n$ produces s outputs in different amounts y_{rj} ($r=1,2,\ldots s$) using m inputs in different amounts x_{ij} ($i=1,2,\ldots,m$). In addition, $\varepsilon > 0$ is a non-Archimedean element defined to be smaller than any positive real number. The presence of ε in the objective function effectively allows minimization over θ to preempt the optimization involving the slacks $s_i^-, s_r^+ \geq 0$ (Cooper, Seiford and Zhu, 2004). Branch$_o$ is efficient if and only if $\theta^* = 1$ and $s_i^{-*} = s_r^{+*} = 0$ for all i and r. Branch$_o$ is weakly efficient if and only if $\theta^* = 1$ and $s_i^{-*} \neq 0, s_r^{+*} \neq 0$ for some i and r.

The complete theoretical background of the applied DEA methods is described in more detail in studies such as Cooper et al. (2004) or Zhu (2009). The efficiency measurement used in this study was created by Tone (1993) with respect to proportional and non-proportional slacks. It was followed by Sevcovic et al. (2001) that also analyzed several efficiency measures. The next sections describe how to enhance and fully adjust the standard DEA method using a quality dimension.

[83] Expansion on the market is limited, and it is more difficult to manage output increase than optimizing resources. We assume that outputs are given exogenous variables and searching for optimal input for each branch. Therefore input oriented strategy is chosen and with this kind of strategy is possible to receive further improvements as earlier discussed.

4.2 Method I — Quality indicator as an Output in DEA model

This is the first method that reflects the quality dimension in branch performance. In this specification basic DEA model is enhanced with one more output, a quality indicator. It is true that the DEA efficiency will not decrease if additional output is included. Therefore, some branches, which were inefficient in a standard DEA model, are becoming efficient. In addition, there could be several branches that are efficient but have low quality, as measured by a certain indicator. In these cases, high productivity compensates for low quality. Quality–productivity tradeoff is present. However, in many applications, this kind of tradeoff is not acceptable.[84] Benchmark branches should have high productivity with high quality. In Model II and III, there are suggestions on how to avoid quality–productivity tradeoff.

4.3 Method II — Quality indicator as an independent factor

In Method II, quality indicator is not included to the basic DEA model, but it is treated independently. In this way, it is possible to avoid quality–productivity tradeoff. All branches have two independent dimensions—productivity (from DEA model) and quality. Each branch has its own place in the two-dimensional chart. It is necessary to set a cut-off for high productivity and a cut-off for high quality in a way that meets operation objectives. Cut-off for high productivity should be 1, and for high quality, it should be the top 20 or 50 percentile through all branches. The two-dimensional chart is split up into four quadrants: high productivity and high quality (HP-HQ), high productivity but low quality (HP-LQ), low productivity but high quality (LP-HQ), and low productivity and low quality (LP-LQ). Branches in the quadrant HP-HQ are the best practice benchmark branches. In this two-dimensional chart, it is possible to depict the relationship between efficiency and quality. However, efficiency measurement with respect to quality is not possible to quantify this model.

A similar approach was done by Camanho and Dyson (1999), where the authors situated bank branches in an efficiency–profitability matrix and analyzed the relationship between the DEA efficiency measure and profitability measure used by a bank. Soteriou and Zenios (1999) published a similar method enhanced by the quality of services in banks. In addition, Brazdik and Druska (2005) applied the DEA efficiency score–revenue performance chart for a mobile telecommunication network.

4.4 Method III — Quality adjusted DEA model

Another way to eliminate the quality–productivity tradeoff and simultaneously quantify efficiency measurement with respect to quality is to apply a quality-adjusted DEA model. A quality-adjusted DEA model is a multi-level DEA model where non-efficient branches are compared only with best practice branches that are efficient (first level) with high quality (multi level).

More precisely, at the end of each level, the efficiency score is calculated for all branches according to the DEA model that also includes the quality dimension. Those branches that are efficient but with low quality are eliminated and do not enter the next level. This iteration is finishing at that level, where all efficient branches have high quality as well. Therefore, they are benchmark branches. The inefficiency score of all other branches are

[84] Or should be within a certain limit

calculated relative to these best practice branches according to variable benchmark methods applied on e-commerce banking activities Cook et al. (2004) and later in Zhu (2009).

5. Results

5.1 Standard DEA framework

In this section, there is a summary of results obtained by non-parametric DEA models. The empirical results are received from the analysis of 185 bank branches based on their figures for the year 2007. These branches deal with individuals and small business enterprise accounts as well. Their activities are considered reasonably homogenous. The input and output specification[85] of the standard DEA model with its descriptive statistics are in *Table 2*.

Table 2 Description statistics of inputs and outputs of the standard DEA model[86], data are related to the year 2007

Variables	Obs.	Median	Mean	St. dev.	Min	Max
Inputs (in # persons)						
SME FTE	185	0.0	2.1	3.6	0.0	15.8
Retail FTE	185	1.3	2.5	3.3	0.0	19.4
UCO FTE	185	7.0	8.8	5.5	2.8	32.1
FTE	185	8.9	13.4	11.7	3.0	64.0
Outputs (in million CZK)		M – Volume *				
Retail Savings	185	58	94	135	0	1269
Retail Loans	185	74	110	124	2	898
SME Deposits	185	52	90	121	0	836
SME Loans	185	252	322	204	0	1257
Outputs (in million CZK)		M – Volume				
Retail Savings	185	56	86	124	0	1276
Retail Loans	185	72	100	99	2	677
SME Deposits	185	51	83	107	0	841
SME Loans	185	244	309	209	0	1264
Outputs (in # contracts)		M – Count				
Retail Savings	185	1257	1682	1269	206	6642
Retail Loans	185	363	444	287	61	1657
SME Deposits	185	106	150	145	4	925
SME Loans	185	48	71	63	0	310

There are three specification of the standard DEA model. First, output factors are measured as new volumes within a year (M–Volume*). Second, in order to eliminate the effect of purchasing power as an external factor, outputs are measured as new volumes within a year adjusted by regional purchasing power (M–Volume). Third, output factors are measured as the number of new products sold (M–Count). Models based on new volumes identify 33-34 branches as fully efficient and 151-152 branches as inefficient, i.e., 18% of branches are efficient (see *Table 3*). The average efficiency of all branches in the network is 74%. Model developed on number of sold products, however, identifies 54 branches as efficient, and the average efficiency is 84%. These results indicate that branches are more homogenous with respect to the number of new products sold and their variations are lower. However, the average efficiency, 74-84%, implies that there is room for improvement through optimal resource allocations.

---

[85] Defined earlier, located in *Appendix, Table A1*

[86] Volume of new business is defined as difference between end-year and start-year stage. Therefore branches could have negative grow of AUM or loans. These figures are entering to the model as zero outputs.

Table 3 Basic statistics and average efficiency of the branch network

Model	Number of efficient branches	Median	Mean	St. dev.	Min	Max
M–Volume*	33	74%	74%	19%	27%	100%
M–Volume	34	72%	74%	19%	28%	100%
M–Count	54	89%	84%	15%	49%	100%

According to the characteristics of efficient branches, it is possible to recommend the optimal branch size. Efficient branches are mainly (54-59%) small branches with 4-6 universal client officers, as shown in *Table 4*. However, among efficient branches, there are 6-12 medium sized and 7-13 large branches as well. Advisors and personal and firm bankers are only efficient in medium and large branches. Most of the efficient branches are located in Region A and Region G.[87] In Region A, there is the highest purchasing power, which has a positive external effect on branch efficiency in the model M–Volume.* On the other hand, Region G has the lowest purchasing power, where excellent management of branches over performs the negative external factor.

Table 4 Characteristics of efficient branches

Model	Number of efficient branches	Branch size	Branch category	Region
M–Volume*	33	5-6 FTEs (4)	Small (20)	Region A (10)
M–Volume	34	5 FTEs (5)	Small (20)	Region G (10)
M–Count	54	4 FTEs (7)	Small (29)	Region G (13)

in bracket are number of branches with the most frequent characteristic

Findings indicate that an optimal branch network should contain a high number small sized branches with only universal client officers and some medium and large branches focusing on personal and firm bankers and advisors' activities. Frequently, branch sizes of 4-6 FTEs indicate that the optimal branch size should be within this interval. Moreover, in the future, it will be optimal to open small branches or redeploy client officers from larger branches to several small ones.

Efficiency score was used to calculate performance rankings of the branches. The sensitivity of results with respect to input-output model specifications was evaluated by calculating the Spearman rank correlation[88] coefficients and by testing statistics for significance of rank correlation coefficients. Results in *Table 5* show that all estimated correlation coefficients are significant, but there is only moderate positive relationship between average efficiency by models M–Volume and M–Count, i.e., branches that operate efficiently with respect to sold products are not necessary operated efficiently with respect to new volumes on those products. These results suggest not a high sensitivity of input-output model specification.

Table 5 Spearman rank correlation coefficients among three standard DEA models

	Obs.	M–Volume*	M–Volume	M–Count
M–Volume*	185	1.000		
M–Volume	185	0.902 (0.000)	1.000	
M–Count	185	0.316 (0.000)	0.423 (0.000)	1.000

[87] Regions are characterized by its purchasing power. Region A has the highest purchasing power, while Region G has the lowest purchasing power

[88] Defined in Spearman (1994), and it is commonly used to compare rankings

In order to identify determinants of efficiency, a correlation analysis was done.[89] Interestingly, there is negative relationship among M–Volume, M–Count, and purchasing power (*Table 6*). This indicates that branches with higher purchasing power are less efficient because they are not able to fully utilize the region's good purchasing power, they do not sell enough products, or they do not have high enough volumes of new products. They have comparable new volumes on deposit and loan products with branches in lower purchasing power regions,[90] but after eliminating the positive effect of purchasing power, the relative value of new volumes on products tend to be lower. In particular, is true for Region A, where the purchasing power is the highest.

Table 6 Correlation coefficients among average efficiencies of three standard DEA models, external factors like purchasing power (PP) and branch size (FTE)

	Obs.	M–Volume*	M–Volume	M–Count
PP	185	0.077 (0.299)	-0.179 (0.015)	-0.432 (0.000)
FTE	185	-0.254 (0.000)	-0.246 (0.001)	0.037 (0.621)

P-values are in the brackets

Another insight gives a negative relationship between M–Volume efficiency and branch size, i.e. larger branches are less efficient in terms of new product volume. However, branch size has no influence on efficiency based on the number of sold products.

Table 7 shows the reported average efficiency results with respect to region and branch size. There are significant variations among regions and branch sizes. Region A is the only region where the average efficiency is lower in the model M–Volume than M–Volume* (p-value at t-test of means is 0.000). It is due to the highest purchasing power in the region that branches are not able to fully utilize. Branches in Region A are the least efficient according to the number of new sold products. There is room for improvement. The most efficient branches are located in Region D.

Table 7 Average branch efficiency according to regions, branch size and DEA models

	Large branches							Medium branches						
Region	M – Volume*	St. dev.	M – Volume	St. dev.	M – Count	St. dev.	Obs.	M – Volume*	St. dev.	M – Volume	St. dev.	M – Count	St. dev.	Obs.
Region A	89%	18%	83%	20%	84%	17%	10	77%	19%	63%	17%	60%	15%	7
Region B	56%	8%	57%	11%	83%	6%	3	59%	14%	65%	22%	72%	18%	5
Region C	44%	12%	48%	16%	89%	12%	6	61%	n/a	69%	n/a	94%	n/a	1
Region D	57%	11%	60%	12%	91%	16%	3	74%	17%	77%	17%	98%	3%	6
Region E	46%	11%	49%	12%	74%	9%	5	62%	8%	64%	9%	70%	15%	3
Region F	74%	20%	77%	19%	92%	10%	4	75%	22%	81%	22%	91%	13%	7
Region G	66%	27%	70%	24%	89%	18%	6	57%	18%	66%	20%	94%	7%	9
All	65%	24%	66%	22%	85%	14%	37	67%	19%	70%	19%	83%	18%	38
	Small branches							All branches						
Region	M – Volume*	St. dev.	M – Volume	St. dev.	M – Count	St. dev.	Obs.	M – Volume*	St. dev.	M – Volume	St. dev.	M – Count	St. dev.	Obs.
Region A	77%	15%	69%	14%	73%	14%	21	80%	17%	72%	18%	73%	17%	38
Region B	78%	15%	78%	14%	84%	15%	16	71%	17%	73%	17%	81%	15%	24
Region C	74%	16%	75%	18%	83%	15%	15	65%	20%	67%	21%	85%	14%	22
Region D	88%	11%	89%	10%	94%	8%	12	79%	17%	82%	16%	95%	8%	21
Region E	80%	14%	80%	14%	85%	12%	20	72%	18%	73%	18%	81%	13%	28
Region F	81%	19%	81%	18%	87%	13%	12	78%	20%	81%	19%	89%	12%	23
Region G	79%	20%	84%	20%	93%	9%	14	70%	23%	75%	22%	92%	11%	29
All	79%	16%	79%	16%	84%	14%	110	74%	19%	74%	19%	84%	15%	185

[89] Regression analysis gives similar results, therefore, we present only the correlation coefficients with p-values

[90] There is no correlation between M–Volume* efficiency and purchasing power, i.e., the correlation is 0.077.

In general, larger branches are less efficient with respect to new volume. The explanation should be in the branch organization[91] and in fact that larger branches have larger customer portfolios that include a higher proportion of less valuable clients. The exception is in Region A, where large branches are the most efficient. Behind this interesting result is the fact that in Region A, small branches are not standalone branches but are connected to one of the large branches.

5.2 Quality indicators

Four main types of quality indicators—penetration index, service quality index, client information index, and retention—are investigated in more detail.

Correlation among quality indicators

Interestingly, there is a relevant positive relationship between penetration index and retention (*Table 8*). Branches where clients have, in average, more products tend to have more loyal customers as well. There is a naturally negative correlation between all factors and product 1, which is defined as the percentage of customers with exactly one product. The most important part of the SQI index is mystery shopping (SQI II) and mystery calls (SQI III), which are highly correlated. However, there is low correlation between customer service (SQI I) and other parts of service quality index. It is because the number of customer service meetings is a rather quantitative indicator and other parts of the service quality index measure real quality service. Those branches that have high-quality service have, on other hand, less customer service meetings, which indicates a certain level of tradeoff.

Table 8 Correlation coefficients among quality indicators

	Penetration	*Product 1*	*Product 2+*	*Product 3+*	*SQI*	*SQI I*	*SQI II*	*SQI III*	*Retention*	*Information*
Penetration	1.000									
Product 1	-0.959 (0.000)	1.000								
Product 2+	0.962 (0.000)	-0.999 (0.000)	1.000							
Product 3+	0.948 (0.000)	-0.827 (0.000)	0.832 (0.000)	1.000						
SQI	0.261 (0.000)	-0.324 (0.000)	0.324 (0.000)	0.170 (0.021)	1.000					
SQI I	0.228 (0.002)	-0.250 (0.001)	0.255 (0.000)	0.175 (0.017)	0.634 (0.000)	1.000				
SQI II	0.211 (0.004)	-0.282 (0.000)	0.278 (0.000)	0.121 (0.102)	0.798 (0.000)	0.186 (0.011)	1.000			
SQI III	0.147 (0.046)	-0.209 (0.004)	0.207 (0.005)	0.082 (0.265)	0.714 (0.000)	0.270 (0.000)	0.612 (0.000)	1.000		
Retention	0.491 (0.000)	-0.473 (0.000)	0.475 (0.000)	0.472 (0.000)	0.151 (0.040)	0.053 (0.471)	0.144 (0.051)	0.098 (0.186)	1.000	
Information	0.016 (0.825)	0.048 (0.518)	-0.043 (0.564)	0.078 (0.289)	0.196 (0.007)	0.044 (0.553)	0.200 (0.006)	0.164 (0.025)	0.135 (0.066)	1.000

Penetration – penetration index, Product 1- portion of customers with exactly one product, Product 2+ - portion of customers with more than 1 products, Product 3+ - portion of customers with more than 2 products, SQI – service quality index, SQI I – customer service, SQI II – mystery shopping, SQI III – mystery call, Retention – percentage of customers who were active in the whole year, Information – client information index, p-values are in the brackets

[91] They employ more special client officers, such as personal and firm bankers or advisers who are not able to bring sufficiently valuable clients to the branch portfolio.

Correlation between quality indicators and external factors

There is a significant relationship among branch characteristics, efficiency results, and quality indicators. Branch size measured as FTE has a negative correlation with SQI, especially SQI II and III, but a positive relationship with SQI I. This indicates that at larger branches, there are lower quality services; they are focused on quantity as a number of customer service meetings. Hence, there should be a large tradeoff between quality and quantity. At these branches, the organization is not effective. There is a large hierarchy at the expense of quality. On the other hand, in small branches, client officers know each other. They can easily cooperate and help each other, which indicate higher service quality appreciated by customers, as measured by mystery shopping or mystery calls. In addition, *Table 9* shows a weak positive correlation between branch size and penetration index. The bigger the branch, the more products their clients tend to have. However, there is no significant relationship with retention.

Table 9 Correlation coefficients among branch size (FTE), purchasing power (PP), average efficiency and quality indicators

Quality indicator	*Obs.*	*FTE*	*PP*	*M–Volume*	*M–Count*
Penetration index	185	0.196 (0.007)	-0.303 (0.000)	0.179 (0.015)	0.327 (0.000)
Product 1	185	-0.113 (0.126)	0.432 (0.000)	-0.202 (0.006)	-0.394 (0.000)
Product 2+	185	0.114 (0.121)	-0.430 (0.000)	0.205 (0.005)	0.393 (0.000)
Product 3+	185	0.245 (0.001)	-0.149 (0.043)	0.138 (0.062)	0.224 (0.002)
SQI Total	185	-0.240 (0.001)	-0.515 (0.000)	0.210 (0.004)	0.317 (0.000)
SQI I	185	0.213 (0.004)	-0.307 (0.000)	0.051 (0.494)	0.186 (0.011)
SQI II	185	-0.387 (0.000)	-0.475 (0.000)	0.196 (0.008)	0.266 (0.000)
SQI III	185	-0.467 (0.000)	-0.390 (0.000)	0.194 (0.008)	0.200 (0.006)
Retention	185	0.051 (0.495)	-0.039 (0.599)	0.013 (0.864)	-0.021 (0.781)
Information	185	-0.160 (0.029)	0.242 (0.001)	0.094 (0.203)	0.017 (0.814)

p-values are in the brackets

Purchasing power has a negative correlation with penetration index and SQI. As the highest purchasing power is in region A, they have the lowest service quality and the lowest product penetration. The latter is due to a larger proportion of foreigners in region A who have just one product—a current account. Further, purchasing power has a slightly positive relationship with the client information index and no relationship with retention. Results demonstrate that client officers know their customers better in regions with higher purchasing power.

The average efficiency of M–volume and M–count models has a positive relationship with penetration index and SQI. These are the most important quality indicators that influence branch efficiency.[92] On the other hand, there is no connection among retention, client information index, and efficiency.

[92] Again, a similar result is obtained by regression analysis.

5.3 Method I — Quality indicator as an Output in DEA model

This is the first specification of the standard DEA model, where quality is included as an additional output factor. In order to test the sensitivity of results by adding one additional quality output factor, the Spearman rank correlation coefficient was calculated. All estimated correlation coefficients are significant and their value range between 0.870-0.991, which suggests the low sensitivity of model specification. This is in line with the arguments in section 4.2.

However, here it is demonstrated that this DEA model does not solve the tradeoff problem between quality and productivity.[93] The tradeoff is present and its magnitude differs with respect to a quality indicator, branch size, and region, as shown below in *Table 10-11*.[94]

Table 10 Tradeoff by quality indicator and model type

		M–Count model				M–Volume model			
Quality indicator	**Obs.**	**Efficiency**	**St. dev.**	**Effective**	**Trade off**	**Efficiency**	**St. dev.**	**Effective**	**Trade off**
Penetration	185	85%	15%	60	28%	75%	19%	42	29%
Product 1[95]	185	88%	14%	73	53%	79%	19%	45	44%
Product 2+	185	85%	15%	58	26%	75%	19%	39	36%
Product 3+	185	85%	15%	62	32%	75%	19%	40	30%
SQI	185	86%	15%	67	31%	75%	19%	41	34%
SQI I	185	87%	15%	77	36%	77%	19%	50	32%
SQI II	185	86%	15%	63	37%	76%	19%	45	31%
SQI III	185	87%	15%	73	41%	76%	19%	44	36%
Retention	185	85%	15%	61	44%	75%	19%	37	51%
Information	185	86%	15%	66	55%	75%	19%	38	47%

Efficiency – average efficiency by the DEA model, St. dev. – standard deviation of efficiency, Effective – number of effective branches, Tradeoff – percentage of effective branches with low quality

The productivity-quality tradeoff ranges between 28-55% with a std. deviation of 9% in cases of M–Count model, and it ranges between 29-51% with a std. deviation of 8% in cases of M–Volume model. There is a high tradeoff in M–Count and M–Volume models with the mystery shopping quality indicator (SQI II), which are mainly valid at large branches (*Table 11*). Similar results were seen for mystery calls (SQI III). Here, high productivity compensates for low quality. Large branches have lower quality measured by mystery shopping and mystery calls, but they are more focused on quantity. As a consequence, there is a large productivity-quality tradeoff. Interestingly, a large tradeoff in the retention indicator is driven by smaller branches. On the other hand, lowest tradeoff is assigned to a DEA model with a penetration index. Penetration index itself is a good predictor of efficiency, and therefore, they off the lowest tradeoff.

[93] Cut-off for high productivity is set at 1, and cut-off for high quality is set in the 50% percentile through all branches.
[94] Tradeoff by quality indicators and region is shown in *Appendix, Table A3.*
[95] Quality indicator Product 1 is an "opposite" indicator, the highest is the worst quality and consequently if you want to compare the tradeoff results with others then you should calculate 100% - actual tradeoff.

Table 11 Tradeoff (percentage of effective branches with low quality) by quality indicator and branch size

	M–Count model				M–Volume model			
Quality indicator	**LB**	**MB**	**SB**	**All**	**LB**	**MB**	**SB**	**All**
Penetration	20% (3)	23% (3)	34% (11)	28% (17)	11% (1)	38% (3)	32% (8)	29% (12)
Product 1	56% (9)	53% (9)	53% (21)	53% (39)	44% (4)	40% (4)	46% (12)	44% (20)
Product 2+	29% (4)	23% (3)	26% (8)	26% (15)	38% (3)	38% (3)	35% (8)	36% (14)
Product 3+	12% (2)	31% (4)	44% (14)	32% (20)	0% (0)	43% (3)	38% (9)	30% (12)
SQI	57% (8)	24% (4)	25% (9)	31% (21)	75% (6)	38% (3)	20% (5)	34% (14)
SQI I	32% (6)	39% (7)	38% (15)	36% (28)	25% (3)	50% (5)	29% (8)	32% (16)
SQI II	64% (9)	23% (9)	31% (11)	37% (23)	63% (5)	38% (3)	21% (6)	31% (14)
SQI III	73% (11)	53% (8)	26% (11)	41% (30)	100% (7)	63% (5)	14% (4)	36% (16)
Retention	25% (4)	29% (4)	61% (19)	44% (27)	38% (3)	71% (5)	50% (11)	51% (19)
Information	69% (6)	63% (5)	44% (12)	55% (23)	57% (4)	63% (3)	39% (10)	47% (17)

LB – large branch, MB – medium sized branch, SB – small branch, in brackets are number of observations

The volume of tradeoff by region is presented in the *Appendix, Table A3*. There is a large tradeoff in DEA models with a quality indicator service quality index of 83%-100% in Region A, which clearly indicates that in Region A, client officers are motivated by the quantity of the sold products, while quality takes second place. However, the lowest tradeoff is in Region F.

5.4 Method II — Quality indicator as an independent factor

In the second specification of the standard DEA model, quality is treated independently to avoid quality-productivity tradeoff. Each branch is characterized with its DEA and quality score. Based on these scores, they are in one of the four quadrants defined in the previous section. Average efficiency and average value of quality indicators within these quadrants are presented in *Table 12* and in *Appendix, Table A4.*

The best practice branches are located in Quadrant 1. Their average efficiency score is 100%, and their average value of quality indicators is above the cut-off value. Branches in Quadrant 2 are those where efficiency is 1, but the value of quality indicators is low, below the cut-off.

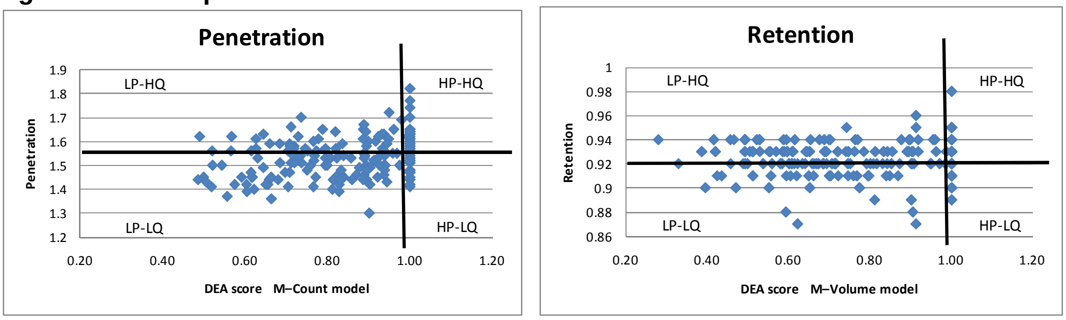
Table 12 Average efficiency and average value of quality indicators according to the quadrants

M – Count	Average efficiency					Average value of quality indicator					
Quality indicator	**1**	**2**	**3**	**4**	**All**	**1**	**2**	**3**	**4**	**All**	**Cut-off**
Penetration	100% (37)	100% (17)	80% (56)	76% (75)	84%	1.61	1.48	1.60	1.46	1.54	1.54
Product 1	100% (15)	100% (39)	76% (74)	81% (57)	84%	0.62	0.54	0.63	0.55	0.59	0.58
Product 2+	100% (39)	100% (15)	81% (56)	75% (75)	84%	0.46	0.37	0.45	0.37	0.41	0.42
Product 3+	100%(34)	100% (20)	79% (61)	77% (70)	84%	0.12	0.08	0.12	0.07	0.10	0.10
SQI	100% (34)	100% (20)	81% (59)	75% (72)	84%	0.89	0.81	0.90	0.82	0.85	0.86
SQI I	100% (28)	100% (26)	80% (61)	76% (70)	84%	0.92	0.78	0.94	0.77	0.85	0.87
SQI II	100% (32)	100% (22)	83% (63)	74% (68)	84%	0.74	0.57	0.76	0.57	0.66	0.67
SQI III	100% (25)	100% (29)	83% (67)	73% (64)	84%	0.93	0.85	0.94	0.85	0.89	0.90
Retention	100% (27)	100% (27)	77% (59)	79% (72)	84%	0.94	0.91	0.94	0.91	0.92	0.92
Information	100% (20)	100% (34)	79% (71)	76% (60)	84%	2.01	1.69	2.02	1.64	1.84	1.81

1 – high productivity and high quality (HP-HQ), 2 – high productivity but low quality (HP-LQ), 3 – low productivity but high quality (LP-HQ), 4 – low productivity and low quality (LP-LQ), Cut-off – is a cut-off value for high quality, defined as a 50% percentile value of quality indicator, in brackets are number of observations

In order to highlight which branches are the best practice branches and which are able to make improvements by increasing quality or productivity, branches are depicted in two-dimensional graphs (see *Figure 1*). In general, the correlation among efficiency and quality indicators is low, as it is reported in *Table 9*, which is also clear in *Figure 1*.

Figure 1 Scatter plots



branches are depicted in the 2-dimensions graph DEA score and quality indicators (Penetration and Retention)

This model specification allows identifying benchmark branches that will move the bank to higher productivity and quality. However, the model is unable to quantify efficiency with respect to quality.

5.5 Method III — Quality adjusted DEA model

In the final specification of the standard DEA model, productivity-quality tradeoff is completely eliminated by the multi-stage quality adjusted DEA model.[96] Here, all inefficient branches are compared to the best practice branches, which are efficient and high quality. The average efficiency is the highest at this specification due to quality adjustment, as demonstrated in *Table 13.*[97]

Table 13 Comparison of average efficiency

		No quality Standard DEA		Quality as output Method I		Quality adjusted Method III	
M–Count	**Obs.**	**Efficiency**	**St. dev.**	**Efficiency**	**St. dev.**	**Efficiency**	**St. dev.**
Penetration	185	84%	15%	85%	15%	89%	13%
Product 1	185	84%	15%	88%	14%	91%	13%
Product 2+	185	84%	15%	85%	15%	89%	13%
Product 3+	185	84%	15%	85%	15%	88%	14%
SQI Total	185	84%	15%	86%	15%	88%	15%
SQI I	185	84%	15%	87%	15%	90%	14%
SQI II	185	84%	15%	86%	15%	89%	13%
SQI III	185	84%	15%	87%	15%	91%	12%
Retention	185	84%	15%	85%	15%	91%	11%
Information	185	84%	15%	86%	15%	92%	11%

No quality – standard DEA model, Quality as output – Method I where quality indicator is an additional output factor, Quality adjusted – Method III quality adjusted DEA , Efficiency – average efficiency

When efficiency is increased, potential cost reductions decrease. Within the standard DEA model, the suggested cost saving is 16%. On other hand, the total potential cost reduction by the quality adjusted DEA model is about 10%, which is significantly lower than the amount suggested with the standard DEA model when it does not adjust for quality. It was also tested to see whether the average efficiency is the same using all methods and the results of t-tests demonstrate that there are significant differences in average efficiency on the significance level, 5%.[98] This result clearly indicates that service quality has significant

[96] All estimated Spearman rank correlation coefficients are significant and their value range between 0.753-0.961, which suggests a low sensitivity of model specification.
[97] Results for M–Volume models are in *Appendix, Table A5-A7*
[98] Due to limited space, detailed results are not reported in the paper

impact on the efficiency of branch network, and it should be incorporated into DEA models and operational processes.

Distribution of branches according to their efficiency score is shown in *Table 14*. There are 41-54 best practice branches and another 27-60 branches that are efficient but score low in service quality. There are 11-31 branches with an efficiency score below 70%, which means that costs (number of FTEs) can be reduce by at least 30% in order to operate efficiently.

Table 14 Distribution of branches based on efficiency score by Quality adjusted DEA model

M – Count			Average efficiency is below				
Quality indicator	**Best practice**	**HP-LQ**	**90%**	**80%**	**70%**	**60%**	**50%**
Penetration	48	27	26	30	13	5	0
Product 1	42	60	21	19	13	6	0
Product 2+	48	25	31	27	15	5	0
Product 3+	48	32	25	26	16	9	0
SQI	49	30	20	24	19	12	0
SQI I	54	35	19	24	12	9	1
SQI II	45	28	23	28	14	7	0
SQI III	54	44	21	21	14	4	0
Retention	41	44	22	23	14	1	0
Information	49	48	19	24	10	1	0

Best practice – number of best practice branches, HP-LQ – number of branches which have high productivity but low quality, Average efficiency is below X% – number of branches which have efficiency score below X%

Characteristics of best practice branches are monitored in *Table 15*. It indicates that best practice branches are mainly in region G, which has the lowest purchasing power; they are mainly small branches with 4-6FTEs. The proportion of small branches within the best practice branches is 39-67%, depending on the quality indicator used.

Table 15 Description of best practice branches with respect of region and branch size

M – Count		Best practice								
		Region			FTE			Branch category		
Quality indicator	**All**	**Region**	**#**	**%**	**FTE**	**#**	**%**	**Branch category**	**#**	**%**
Penetration	48	Region G	12	25%	4	6	13%	SB	22	46%
Product 1	42	Region A	9	21%	4	5	12%	SB	26	62%
Product 2+	48	Region G	13	27%	4	7	15%	SB	25	52%
Product 3+	48	Region G and D	11	23%	4	7	15%	SB	21	44%
SQI	49	Region G	12	24%	4,6	5	10%	SB	29	59%
SQI I	54	Region G and D	12	22%	4	7	13%	SB	26	48%
SQI II	45	Region G	12	27%	4,6	5	11%	SB	27	60%
SQI III	54	Region F	12	22%	4	8	15%	SB	36	67%
Retention	41	Region G	12	29%	4	4	10%	SB	16	39%
Information	49	Region A	13	27%	4	6	12%	SB	27	55%

SB – small branch, # – number of branches within benchmarks, % – percentage of best practice branches with a certain characteristic

This is a surprising result. Managers, based on ratio analysis, assumed that small branches are not efficient and large branches are considered the best performers. Also, it is documented that the best practice branches are mainly in Region G and not in Region A, as was assumed by the bank. These results, however, are in line with other DEA studies[99], such as Sherman and Zhu (2006).

[99] Some DEA studies show that optimal branch is about 6-9 FTEs. It might differ bank by bank depending on bank processes. However, most of benchmarking studies show that bigger branches (above 15FTEs) are mostly less efficient in terms of activities and customer satisfaction due likely to two reasons: lower complexity (higher manageability) and more human-client focused approach. Managers of small branches also act more as owners and not officers.

6. Conclusion

In this chapter, there were three methods applied to incorporate the quality dimension into the performance of bank branches. Quality of service is measured through service quality, product penetration, client information, and retention index. We identified that the productivity–quality tradeoff exists and it is possible to avoid through multi-level quality adjusted DEA model, where benchmark branches have not only high productivity but high service quality as well. Results show that service quality has a significant impact on branch efficiency, and it should be incorporated into DEA models and operational processes. The essay demonstrates the short-term interaction among service quality and operating branch efficiency.

From a policy perspective, the essay provides evidence that there are real reserves for improvement, an average efficiency of 74-84%, which can be realized through optimal resource allocations and increasing service quality. We discovered that the main factors of efficiency, quality, and productivity-quality tradeoffs are branch size and region, characterized by complex indicator purchasing power. There is documented evidence that larger branches are less efficient than smaller ones. Results also show that branches in the region with the highest purchasing power are not able to fully utilize their opportunities, which implies lower efficiency. Branches that operate efficiently with respect to the number of sold products do not necessarily operate efficiently with respect to new volumes on those products. In addition, branches are less homogenous with respect to new volume than by number of new products sold.

Benchmark branches are mainly small and are in the region with the lowest purchasing power despite managerial expectations. However, the results are in line with other studies conducted to DEA research. The most frequent optimal branch size, 4-6 FTEs, indicates that optimal branch size should be within this interval. In the future, it will be optimal to open small branches or redeploy client officers from large inefficient branches to several small efficient ones. Moreover, findings indicate that branch networks should contain a high number of small sized branches with only universal client officers and some medium and large branches focusing on personal and firm bankers and advisors' activities.

Most importantly, the quality indicator that explains efficiency is product penetration. Further, the quality level and magnitude of the productivity-quality tradeoff differs by branch size and region. Those branches that have high-quality service are measured by mystery shopping and mystery calls, and they have less client service meetings, which indicates a certain level of tradeoff. Findings demonstrate that large branches focus more on the information index and customer service meetings, while small branches are more interested in mystery shopping and mystery calls. Interestingly, branches in regions with high purchasing power have worse results in terms of product penetration, mystery shopping, and mystery calls than branches with low purchasing power. The largest productivity-quality tradeoff was found at large branches with respect to quality indicators, mystery shopping, and mystery calls.

APPENDIX

Table A1 Description of input and output factors used in DEA models, data are related to one year period in 2007

Inputs	**(in number of persons)**
SME FTE	Average number of FTEs, advisors and firm bankers for non-retail clients
Retail FTE	Average number of FTEs, advisors and personal bankers for retail clients
UCO FTE	Average number of FTEs, universal client officers and a branch director
Outputs	**(in mln CZK), Model M–Volume***
Retail savings	Volume of new retail savings (current and saving accounts, term deposits, investment funds, pension funds, housing savings, single and regular life insurance)
SME deposits	Volume of new SME deposits (current and saving accounts, term deposits, investment funds)
Retail Loans	Volume of new retail loans (consumer loans, credit cards, overdrafts, housing loans, mortgages)
SME Loans	Volume of new SME loans (overdrafts, investment loans, revolving, factoring, leasing)
Outputs	**(in mln CZK), Model M–Volume**
Retail savings	Volume of new retail assets under management (current and saving accounts, term deposits, investment funds, pension funds, housing savings, single and regular life insurance) divided by branch's town purchasing power index
SME deposits	Volume of new SME assets under management (current and saving accounts, term deposits, investment funds) divided by branch's town purchasing power index
Retail Loans	Volume of new retail loans (consumer loans, credit cards, overdrafts, housing loans, mortgages) divided by branch's town purchasing power index
SME Loans	Volume of new SME loans (overdrafts, investment loans, revolving, factoring, leasing) divided by branch's town purchasing power index
Outputs	**(in number of contracts), Model M–Count**
Retail savings	Number of retail deposit products - current and saving accounts, term deposits, investment funds, pension funds, housing savings, single and regular life insurance
SME deposits	Number of SME deposit products - current and saving accounts, term deposits, investment funds
Retail Loans	Number of retail loan products - consumer loans, credit cards, overdrafts, housing loans, mortgages
SME Loans	Number of SME loan products - overdrafts, investment loans, revolving, factoring, leasing

Table A2 Description of quality indicators used in DEA models

Quality indicators	**Description**
Service quality index , *SQI*	Measured by % and has three components: customer service, mystery shopping and mystery call
Customer service, *SQI I.*	Certain type of a meeting between a client officer and an existing customer in order to maintain a customer relationship, to indentify customer needs and increase probability to sell a new product. Measured as a fulfillment of the branch plan in %.
Mystery shopping, *SQI II.*	Certain type of a branch visit in order to evaluate several aspect of the service: quality and correctness of provided information and way of communication with a customer. Measured as a fulfillment of all mandatory factors in %.
Mystery call, *SQI III.*	Certain type of a phone call in order to evaluate several aspect of the service: quality and correctness of provided information and a way of communication with a customer by phone call. Measured as a fulfillment of all mandatory factors in %.
Client information index, *Information*	Average number of available client information: education, job, id card, phone number, e-mail, income and expense
Product penetration index, *Penetration*	Average number of products per customer based on Finalta definitions. Each of the following products is counted with equal weight: current account, saving account, term deposit, investment fund (including pension savings), life insurance, consumer finance (including consumer loan, overdraft, and credit card) and mortgage.
Client retention, *Retention*	Percentage of active customers at the end of the year 2006 that were still active customers at the end of the year 2007.

Table A3 Tradeoff (percentage of effective branches with low quality) by quality indicator and region – Method I

M-Count	Region A	Region B	Region C	Region D	Region E	Region F	Region G	All
Penetration	43% (3)	50% (2)	20% (1)	23% (3)	0% (0)	40% (4)	27% (4)	28% (17)
Product 1	10% (1)	30% (3)	57% (4)	69% (9)	71% (5)	58% (7)	71% (10)	53% (39)
Product 2+	57% (4)	25% (1)	20% (1)	25% (3)	0% (0)	30% (3)	20% (3)	26% (15)
Product 3+	25% (2)	50% (2)	33% (2)	15% (2)	17% (1)	50% (5)	40% (6)	32% (20)
SQI Total	83% (5)	20% (1)	29% (2)	38% (6)	17% (1)	9% (1)	31% (5)	31% (21)
SQI I	71% (5)	67% (4)	38% (3)	29% (5)	14% (1)	21% (3)	39% (7)	36% (28)
SQI II	67% (4)	17% (1)	38% (3)	38% (5)	17% (1)	50% (5)	29% (4)	37% (23)
SQI III	67% (4)	29% (2)	38% (3)	56% (9)	0% (0)	23% (3)	56% (9)	41% (30)
Retention	57% (4)	100% (4)	57% (4)	50% (6)	40% (2)	20% (2)	31% (5)	44% (27)
Information	33% (4)	75% (3)	60% (3)	46% (6)	60% (3)	73% (8)	56% (9)	55% (36)
M-Volume	**Region A**	**Region B**	**Region C**	**Region D**	**Region E**	**Region F**	**Region G**	**All**
Penetration	29% (2)	33% (1)	33% (1)	17% (1)	0% (0)	25% (2)	42% (5)	29% (12)
Product 1	17% (2)	33% (1)	50% (2)	60% (3)	67% (2)	50% (4)	60% (6)	44% (20)
Product 2+	71% (5)	50% (1)	33% (1)	20% (1)	0% (0)	25% (2)	36% (4)	36% (14)
Product 3+	29% (2)	33% (1)	67% (2)	0% (0)	33% (1)	14% (1)	42% (5)	30% (12)
SQI Total	100% (7)	0% (0)	0% (0)	40% (2)	33% (1)	0% (0)	36% (4)	34% (14)
SQI I	44% (4)	25% (1)	33% (1)	29% (2)	33% (2)	20% (2)	36% (4)	32% (16)
SQI II	100% (7)	0% (0)	0% (0)	75% (3)	33% (1)	0% (0)	25% (3)	31% (14)
SQI III	63% (5)	25% (1)	0% (0)	50% (3)	0% (0)	13% (1)	60% (6)	36% (16)
Retention	43% (3)	100% (2)	100% (3)	60% (3)	0% (0)	29% (2)	55% (6)	51% (19)
Information	25% (2)	33% (1)	0% (0)	60% (3)	100% (2)	29% (2)	80% (8)	47% (18)

in brackets are number of observations

Table A4 Average efficiency and average value of quality indicators according to the quadrants in Method II DEA model

M – Volume	Average efficiency					Average value of quality indicator					
Quality indicator	**1**	**2**	**3**	**4**	**All**	**1**	**2**	**3**	**4**	**All**	**Cut-off**
Penetration	100% (22)	100% (12)	68% (71)	69% (80)	74%	1.61	1.48	1.60	1.47	1.54	1.54
Product 1	100% (14)	100% (20)	68% (75)	69% (76)	74%	0.62	0.54	0.64	0.55	0.59	0.58
Product 2+	100% (20)	100% (14)	69% (75)	68% (76)	74%	0.47	0.38	0.45	0.36	0.41	0.42
Product 3+	100% (22)	100% (12)	67% (73)	70% (78)	74%	0.12	0.08	0.12	0.08	0.10	0.10
SQI	100% (20)	100% (14)	71% (73)	67% (78)	74%	0.90	0.82	0.89	0.81	0.85	0.86
SQI I	100% (18)	100% (16)	68% (71)	69% (80)	74%	0.93	0.80	0.93	0.76	0.85	0.87
SQI II	100% (20)	100% (14)	71% (75)	66% (76)	74%	0.74	0.56	0.76	0.57	0.66	0.67
SQI III	100% (18)	100% (16)	72% (74)	65% (77)	74%	0.94	0.85	0.94	0.85	0.89	0.90
Retention	100% (15)	100% (19)	68% (71)	69% (80)	74%	0.94	0.91	0.93	0.91	0.92	0.92
Information	100% (16)	100% (18)	69% (75)	68% (76)	74%	2.04	1.66	2.01	1.66	1.84	1.81

1 – high productivity and high quality, 2 – high productivity but low quality, 3 – low productivity but high quality, 4 – low productivity and low quality, Cut-off – is a cut-off value for high quality, defined as a 50% percentile value of quality indicator, in brackets are number of observations

Table A5 Comparison of average efficiency

		No quality Standard DEA		Quality as output Method I		Quality adjusted Method III	
M–Volume	**Obs.**	**Efficiency**	**St. dev.**	**Efficiency**	**St. dev.**	**Efficiency**	**St. dev.**
Penetration	185	74%	19%	75%	19%	78%	19%
Product 1	185	74%	19%	79%	19%	83%	18%
Product 2+	185	74%	19%	75%	19%	80%	18%
Product 3+	185	74%	19%	75%	19%	77%	19%
SQI Total	185	74%	19%	75%	19%	78%	19%
SQI I	185	74%	19%	77%	19%	81%	19%
SQI II	185	74%	19%	76%	19%	80%	19%
SQI III	185	74%	19%	76%	19%	83%	17%
Retention	185	74%	19%	75%	19%	85%	16%
Information	185	74%	19%	75%	19%	77%	19%

No quality – standard DEA model, Quality as output – Method I where quality indicator is an additional output factor, Quality adjusted – Method III quality adjusted DEA , Efficiency – average efficiency

Table A6 Distribution of branches based on efficiency score by Method III - Quality adjusted DEA model

M – Volume			Average efficiency is below				
Quality indicator	**Best practice**	**HP-LQ**	**90%**	**80%**	**70%**	**60%**	**50%**
Penetration	32	15	117	101	70	36	14
Product 1	33	37	97	77	52	29	11
Product 2+	32	20	108	91	60	27	8
Product 3+	34	14	120	100	72	41	15
SQI	30	20	112	97	71	38	14
SQI I	35	27	101	82	60	33	13
SQI II	37	24	110	94	62	33	11
SQI III	37	31	95	80	53	19	6
Retention	36	33	95	75	38	17	3
Information	26	20	118	103	75	37	14

Best practice – number of best practice branches, HP-LQ – number of branches which have high productivity but low quality, Average efficiency is below X% – number of branches which have efficiency score below X%

Table A7 Description of best practice branches in Method III with respect of region and branch size

M – Volume		Best practice								
		Region			**FTE**			**Branch category**		
Quality factor	**All**	**Region**	**#**	**%**	**FTE**	**#**	**%**	**Branch category**	**#**	**%**
Penetration	32	Region G	7	22%	4	5	16%	SB	18	56%
Product 1	33	Region A	12	36%	6	5	15%	SB	20	61%
Product 2+	32	Region G	9	28%	4	4	13%	SB	18	56%
Product 3+	34	Region G	8	24%	4	8	24%	SB	19	56%
SQI	30	Region F	10	33%	5,6	4	13%	SB	22	73%
SQI I	35	Region F	8	23%	4	7	20%	SB	21	60%
SQI II	37	Region G	10	27%	5,6	5	14%	SB	26	70%
SQI III	37	Region F	10	27%	5	6	16%	SB	27	73%
Retention	36	Region F	10	28%	4	6	17%	SB	23	64%
Information	26	Region A and F	7	27%	4	6	23%	SB	19	73%

SB – small branch, # – number of branches within benchmarks, % – percentage of best practice branches with a certain characteristic

REFERENCES

Athanassopoulos AD (1997): Service quality and operating efficiency synergies for management control in the provision of financial services: Evidence from Greek bank branches, *European Journal of Operational Research*, 98(2):300-313

Athanassopoulos AD (1998): Nonparametric frontier models for assessing the market and cost efficiency of large-scale bank branch networks, *Journal of Money, Credit and Banking*, 30(2):172-192

Athanassopoulos AD, Giokas D (2000): The Use of Data Envelopment Analysis in Banking Institutions: Evidence from the Commercial Bank of Greece, *Interfaces*, 30(2):81-95

Banker RD, Charnes A, Cooper WW (1984): Some models for estimating technical and scale inefficiencies in Data envelopment analysis, *Management Science*, 30(9):1078-1192

Beamon BM (1999): A model and measurement approach for studying retail productivity, *International Journal of Operational and Production Management*, 19(3):275-292

Berger AN, Humphrey DB (1997): Efficiency of financial institutions: international survey and directions for future research, *European Journal of Operational Research*, 98(2):175-212

Brazdik F, Druska V (2005): Too large or too small? Returns to scale in a retail network, CERGE-EI Working Paper, no. 273

Callen JL (1991): Data Envelopment Analysis: Partial Survey and Applications for Management Accounting, *Journal of management accounting research*, 3:35-56

Camanho AS, Dyson RG (1999): Efficiency, size, benchmarks and targets for bank branches: an application of data envelopment analysis, *Journal of the Operational Research Society*, 50(9):903-915

Charnes A, Cooper WW, Rhodes E (1978): Measuring the efficiency of decision making units, *European Journal of Operational Research*, 2(6):429-444

Cook WD, Seiford LM, Zhu J (2004): Models for performance benchmarking: Measuring the effect of e-commerce activities on banking performance, *Omega*, 32(4):313-322

Cooper WW, Seiford ML, Zhu J (2004): *Handbook on Data Envelopment Analysis*, Kluwer Academic Publisher, UK

Drake L, Howcroft B (1994): Relative efficiency in the branch network of a UK Bank: an empirical study, *Omega*, 22(1):83-90

Emrouznejad A, Barnett RP, Tavares G (2008): Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA, *Socio-Economic Planning Sciences*, 42(3):151-157

Giokas D (1991): Bank branch operating efficiency: a comparative application of DEA and the loglinear model, *Omega*, 19(6):549-557

Irsova Z (2009): Measuring bank efficiency, Master thesis at the Institute of Economic Studies, Charles University

Lynch R (2005): *Corporate Strategy*, Prentice Hall, UK

Oral M, Kettani O, Yolalan R (1992): An empirical study of analyzing the productivity of bank branches, *Institute of Industrial Engineers Transactions*, 24(5):166-176

Oral M, Yolalan R (1990): An empirical study on measuring operating efficiency and profitability of bank branches, *European Journal of Operational Research*, 46(3):282-294

Reynolds D (2004): An exploratory investigation of multiunit restaurant productivity assessment using data envelopment analysis, *Journal of Travel and Tourism Marketing*, 16(2-3):19-26

Schaffnit C, Rosen D, Paradi JC (1997): Best practice analysis of bank branches: an application of DEA in a large Canadian bank", *European Journal of Operational Research*, 98(2):269-289

Sevcovic D, Halicka M, Brunovsky P (2001): DEA analysis for a large structured bank branch network, *Central European Journal of Operational Research*, 9:329-342

Sherman HD, Ladino G (1995): Managing bank productivity using Data envelopment analysis (DEA), Interfaces, 25(2):60-73

Sherman HD, Zhu J (2006): *Service productivity management, Improving service performance using Data Envelopment Analysis*, Springer Science + Business Media, LLC, USA

Soteriou A, Zenios SA (1999): Operations, Quality and Profitability in the provision of banking services", *Management sciences*, 45(9):1221-1238

Spearman C (1994): The Proof and Measurement of association between two things, *American Journal of Psychology*, 15(1):72-101

Tone K (1993): An epsilon-free DEA and a new measure of efficiency, *Journal of the Operational Research Society Japan*, 36(3):167-174

Tulkens H (1993): On FDH efficiency analysis: some methodological issues and applications to retail banking, courts and urban transit, *Journal of productivity analysis*, 4(1-2):183-210

Waterman R et al. (1980): Structure Is Not Organization, *Business Horizons*, 23(3):14–26

Waterman R et al. (1982): *In Search of Excellence*, New York, Harper & Row

Zhu J (2009): *Quantitative models for performance evaluation and benchmarking*, Worcester Polytechnic Institute, Worcester, MA, USA