REGULATION AND BANKING: A SURVEY

G.E. Hebbink and H.M. Prast

Onderzoeksrapport WO&E nr 565

December 1998

De Nederlandsche Bank
Afdeling Wetenschappelijk
onderzoek en econometrie
Postbus 98
1000 AB AMSTERDAM

ABSTRACT

Regulation and banking: a survey

G.E. Hebbink and H.M. Prast

This report presents an overview of the theory of regulation in general, with special attention for the regulation of banks. Two theories of government regulation are described. The first, normative, theory uses market failures as the justification of government regulation. The second, positive, theory explains the existence of regulation as the outcome of the interaction between politicy makers and pressure groups. The second half of the report is devoted to bank regulation. It is argued, that it is the combination of withdrawable deposits and illiquid, non-marketable assets that justifies regulation. The effects of regulatory policies in the form of deposit insurance, capital requirements, bank monitoring, bank closure policy as well as the optimal design of regulatory policy are discussed. From a first look at the practice of regulation it is concluded, that the theoretical recommendations are to a certain degree already applied in practice. Examples are risk-related capital requirements and deposit insurance premiums.

Keywords: regulation, banking

JEL codes: G21, G28

SAMENVATTING

Regulering en bankwezen: een overzichtsartikel

G.E. Hebbink en H.M. Prast

Dit onderzoeksrapport geeft een overzicht van de theorie van regulering in het algemeen en van die van het bankwezen in het bijzonder. Twee theorieën ter verklaring van regulering worden beschreven. De normatieve benadering legt de nadruk op marktfalen als reden voor overheidsingrijpen. De positieve benadering verklaart overheidsregulering als de uitkomst van een wisselwerking tussen het gedrag van beleidsmakers en dat van belangengroepen. De tweede helft van het rapport is gewijd aan bankregulering. Het is betoogd, dat de voor banken typerende combinatie van direct opeisbare tegoeden enerzijds en niet-verhandelbare activa anderzijds regulering rechtvaardigt. De effecten van regulering in de vorm van depositoverzekering, kapitaalvereisten, bank monitoring, het sluiten van banken alsmede de optimale vormgeving van het bankentoezicht worden besproken. Uit een confrontatie van de theorie met de praktijk van bankregulering blijkt, dat de aanbevelingen uit de theorie tot op zekere hoogte reeds hun weerslag in de praktijk hebben gevonden. Voorbeelden zijn risico-afhankelijke depositoverzekeringspremies en dito kapitaalvereisten.

Trefwoorden: regulering, bankwezen

JEL codes: G21, G28

1 INTRODUCTION

In all market economies there is government intervention. This may take the form of legislatory policy, taxation and regulation of markets. Intervention is usually justified on grounds of promoting efficiency and equity. Financial markets are especially subject to regulation. The purpose of this report is twofold. It aims at presenting an overview of the theory on regulatory policy (and its justification) in general, as well as providing a survey of the theoretical literature on banking regulation. The report is structured as follows. In section 2, a definition of regulation is provided, with an introductory discussion of different types of regulation. Section 3 covers market failures, the main justification for government regulation at the market or firm level. Included are the concepts of Pareto efficiency, public goods, externalities, market power, and asymmetric information. An overview of the ‘economic theory of regulation’ is given in section 4. This theory applies a positive methodology to the analysis of regulation, by taking economic behaviour of regulators and interest groups as a starting point. Section 5 focuses on the regulation of banks. It pays attention to the question whether banks indeed should be subject to regulation and if so, why. Moreover it presents a number of theoretical models highlighting regulatory instruments for the financial sector as well as deriving the approppriate institutional design of regulatory policy. In section 6, some key features of the practice of banking regulation are evaluated in the light of these theoretical models. The paper ends with a summary and conclusions.

2 WHAT IS REGULATION?

Many authors define regulation differently. Here, we follow Stigler (1971), who defines regulation as “any policy which alters market outcomes by the exercise of some coercive government power.” This definition is attractive because of its conciseness, although, precisely because of this, it provokes some questions as well. First, Stigler’s definition does not give any indication why a government should intervene in market outcomes. Many authors, however, do not take such a methodologically positive view, and state that government regulation is directed at some more or less clearly defined objective. Chang (1997) e.g., adds to his definition of regulation that it is meant to serve the ‘public interest.’ Advocates of regulation often have the intention to improve economic welfare by using regulation to promote economic efficiency. This argument is based on the Pareto efficiency of competitive markets. Moreover, if lump-sum taxes and transfers are feasible, it is possible to maximize social welfare through redistribution of a Pareto efficient allocation (provided there is some well-defined social welfare function, see section 2.3). The relation between the lack of competitive markets and regulation is reflected in the definition of regulation by Laffont (1994), who links regulation to industrial organization: “regulation is the public economics face of industrial organization”, where industrial organization is “the study of economic activity at the level of a firm or an industry when the paradigm of perfect competition appears inadequate”.

A second question that comes to mind if Stigler’s definition is considered, is which types of policy are included in regulation. In general, Swann (1989) distinguishes three categories of regulation. The first form of regulation is *antitrust policy*. This is aimed at the promotion of fair competition in monopolistic industries by imposing price and entry regulations. The second main form of regulation is *economic regulation*. Sometimes economic regulation is labeled ‘old style regulation’. It typically involves price and entry controls in competitive industries. The third main form of regulation is *social regulation*. Joskow and Noll (1981) call this “qualitative” regulation. This type of regulation is concerned with externalities, like environmental, health occupational safety, and product quality regulation. Each of the three types of regulation has been introduced to correct some type of market failure, like externalities or market power, that leads to a divergence from Pareto efficiency (Swann, 1989). In section 2.3 we will return to the subject of market failures.

A third question to the Stigler definition concerns the difference between self-regulation and external regulation. Strictly speaking, Stigler’s definition excludes self-regulation, but most self-regulatory institutions have been set up with some initial government action (Swann, 1989). A similar distinction is sometimes made between regulation in a narrow sense, and taxation and subsidies. As Posner has shown (1971), however, this distinction is not always clear. Regulation can take the form of taxation and certain types of regulation may be felt to the affected industry as taxation. Both distinctions, external versus self-regulation and taxation versus other types of regulation, do not concern us here. In this paper, without further indication, we consider regulation in its widest sense.

Although the largest part of the literature on regulation focusses on the purpose of regulation, there is a large body of literature that takes a more positive point of departure. These studies try to answer general questions concerning the effects of the presence of regulatory policy and the behaviour of regulatory agencies as a class of government institutions. The effects of regulation are the subject of what has become known as the economic theory of regulation. This theory tries to give a theoretical foundation to the phenomenon of regulation based on economic behaviour of individual agents or agencies and studies its effects, including its possibly non-intended side-effects. Section 4 describes the economic theory of regulation in larger detail.

Although the theoretical developments in the economic theory of regulation are impressive, its testable implications can be ambiguous and their explanatory power controversial (Peltzman, 1989; Winston and Crandall, 1994). Moreover, empirical research in this theory is mostly applied to specific areas of regulation.

3 MARKET FAILURES AS A JUSTIFICATION OF REGULATION

3.1 Introduction

The main economic justification for goverment intervention through regulation is based on the existence of market failures. Market failures exist if markets do not work perfectly, in the sense that (1) there is perfect competition, (2) there are markets for all goods with interaction between agents through prices, and (3) all agents are fully informed. Market failures lead to inefficient allocations, in which welfare for the society as a whole can be improved, which is considered a cause for regulation by the public sector. An important counter-argument, however, is the Coase theorem, which states that Pareto efficient allocations can be reached through private negotiations. This section covers the concepts of Pareto efficiency, public goods, externalities, the Coase theorem, market power, and asymmetry of information.

3.2 Pareto efficiency

An important concept in the analysis of market failures is Pareto efficiency. The allocation of goods in an economy is *Pareto efficient* if there is no feasible alternative allocation such that all agents prefer the alternative over the existing allocation, at given initial endowments of goods (Varian, 1984).[1] There is a fundamental relationship between Pareto efficiency and competitive equilibrium. Competitive equilibrium or Walrasian equilibrium is defined by a situation where demand equals supply for all goods in an economy, individual agents take prices of these goods as given by market outcomes, and agents interact only through their mutual effects on market prices.[2] In this equilibrium the allocation of goods is Pareto efficient.[3] In other words, any divergence from a competitive equilibrium leaves room for the improvement of welfare. If the conditions for a competitive equilibrium are not fulfilled, there is some type of *market failure*. Since

[1] Under some weak hypotheses this definition is equivalent to: There is no feasible allocation where everyone is at least as well off and at least one agent is strictly better off.
[2] For a mathematically rigorous definition see, e.g., Varian (1984). Throughout this paper we skip formal definitions and proofs, which in most cases can be found in the references.
[3] This is known as the First Theorem of Welfare Economics. For more details see, for example, Varian (1984).

market failures are the cause of Pareto inefficient allocations, many economists find market failures the *prima facie* economic justification for government interventions, including regulation. In fact, market failures may justify the existence of governments itself (Mueller, 1989).

The concept of a good, used in the definition of Pareto efficiency, is very broad. It includes consumed and produced goods, including labour, distinguished by time, location and state of the world. For example, this includes contingency contracts that define the supply of a specific good in each uncertain outcome of some random event. The efficiency of a competitive equilibrium does not imply that competition leads to an optimum or ‘best’ level of welfare in the economy. Although Pareto efficiency could be considered as a normative criterion, it is indeterminate, as it does not say anything about welfare distribution. Still, any Pareto efficient allocation can be achieved as a competitive equilibrium with redistribution of initial endowments.[4] Moreover, if a social welfare function exists, it can be shown that its maximum is Pareto efficient (Varian, 1984).

Following the definition of Pareto efficiency and its relation to competitive equilibrium, we could choose to list the technical assumptions for competitive equilibrium, and check in how far they are fulfilled in reality.[5] Ideally, this could give us indications which changes are required to reach a desired Pareto efficient situation. However, instead of this technical approach, we take one step further, and list situations in which the assumptions for competitive equilibrium are not met. These cases result in inefficiencies, in which one or more individuals can be made better off, without negative effects on someone else. In a situation like that some kind of *market failure* exists. The types of market failure we consider are public goods, externalities, market power, and informational asymmetry. In general, it is clear that markets or goods may be fundamentally non-competitive, and hence Pareto inefficient.

3.3 Public goods

According to Samuelson (1954) a good is a public good if each individual’s consumption leads to no subtraction from any other individual’s consumption of that good. A public good in its most pure form must be provided in equal quantities to all individuals. In other words, it is supplied to all consumers jointly, and it is impossible to exclude others from its consumption. Well known examples are clear air, national defense, street lights, and fire protection, but systems of property rights or law and law enforcement are public goods as well. The public good in the latter examples increases cooperation and coordination among individuals, which supports the

[4] This is known as the Second Theorem of Welfare Economics (Varian, 1984).
[5] For example, goods are homogeneous, and the number of agents is large, such that prices are given to individuals.

possibility to engage in exchange of goods that benefits both parties. The existence of public goods, like law and property rights, is viewed by many authors as the reason for the emergence of some type of government system based on collective choice (Mueller, 1989).

In a large group, or a society, voluntary compliance with provision of public goods leads to free riding and under- or nonprovision of the public good (Olson, 1965). That coordinated provision of public goods is superior (in a Pareto sense) can be shown as follows. Suppose each individual would voluntarily supply G_i of the public good, and enjoys utility $U_i(X_i,G)$ from a private good X_i and the total quantity of public good supplied G. Given a budget constraint $Y_i=P_xX_i+P_gG_i$, where income Y and prices P_x and P_g are fixed, each individual takes the total supply of public goods as given and maximizes utility by choosing G_i and X_i. This yields the Cournot-Nash equilibrium condition:

$$\frac{\partial U_i / \partial G}{\partial U_i / \partial X_i} = \frac{P_g}{P_x} \tag{1}$$

That this equilibrium results in a less Pareto efficient allocation of goods than in the case of coordination can be seen by maximizing a social welfare function that adds the individual utilities U_i, with positive weights. Hence, if the weighted sum of individual welfare levels is maximized, a Pareto improvement can not be attained by separately maximizing individual welfare levels. The budget constraint now reads $\Sigma Y = P_x\Sigma X_i + P_gG$. Maximizing total welfare by choosing X_i and G_i under the budget constraint gives a Pareto optimal allocation, where the following condition exists (Samuelson, 1954):

$$\sum_i \frac{\partial U_i / \partial G}{\partial U_i / \partial X_i} = \frac{P_g}{P_x} \tag{2}$$

which can be rewritten as

$$\frac{\partial U_i / \partial G}{\partial U_i / \partial X_i} = \frac{P_g}{P_x} - \sum_{j \neq i} \frac{\partial U_j / \partial G}{\partial U_j / \partial X_j} \tag{3}$$

If G and X are normal goods, the second term on the right hand side is positive, and the left hand side in (1) is less than in (3). Consequently, with voluntary provision of public goods (equation 1) less of the public good and more of the private good is consumed than with coordination (equation 3), where coordination is Pareto superior to individual supply of public goods.

3.4 Externalities and the Coase theorem

An externality occurs when consumption or production of one individual or firm has an unintended impact on the utility or production function of another individual or firm. In fact, public goods are goods with a specific kind of externality (Varian, 1984). In an economic sense, the fact that there is an unintended impact reflects that there is no existing market or price mechanism to coordinate the externality. In this situation, price-taking profit maximizing or utility maximizing behaviour will not necessarily lead to allocative efficiency. There are several ways to find a solution to the problem of externalities. First, a system of taxes and subsidies could be set up by the government, correcting the relative prices that producers and consumers face. This option was first proposed by Pigou (1920). A problem with a tax/subsidy scheme is that it is difficult to find the required level of tax or subsidy. Another possibility is that the government sets up a private market for externalities. However, it may be prohibitively costly to set up a market like this, although there are cases in which it might work, like pollution rights. A third option is to change property rights in such a way that an efficient allocation is obtained. For example, a firm that suffers from another firm's externality may consider to buy the other firm. The total production of the new firm will then be efficient, since the negative effect of the other firm's production is internalized. However, this is not really a solution, since a feasible solution like this should be attainable through the market mechanism, because it will be perceived as such by the first firm.

Another solution to the problem of externalities that does not require government intervention is offered by Coase (1960). The famous Coase theorem states that, in the absence of transaction or bargaining costs, a Pareto optimum can be achieved by private negotiations, regardless of the distribution of property rights. The theorem can be explained by a simple model of two firms, where one of them suffers from an externality. Assume there are two firms, each producing output Y_i by employing labour L_i. Firm 2's output depends negatively on output of firm 1:[6]

$$Y_1 = F_1(L_1),\ Y_2 = G_2(Y_1, L_2)$$

$$\text{with } \frac{\partial F_1}{\partial L_1} > 0,\ \frac{\partial G_2}{\partial L_2} > 0,\ \frac{\partial G_2}{\partial Y_1} < 0$$

Substituting Y_1 in Y_2 gives $Y_2 = F_2(L_1, L_2)$, where $\partial F_2 / \partial L_1 < 0$ and $\partial F_2 / \partial L_2 > 0$.

---

[6] All second order partial derivatives are assumed to be negative.

A Pareto optimum is found by maximizing the sum of both firms' profit, where prices p and the wage rate w are given.[7] The Pareto optimum condition is:

$$p_1 \frac{\partial F_1}{\partial L_1} + p_2 \frac{\partial F_2}{\partial L_1} = p_2 \frac{\partial F_2}{\partial L_2} \tag{4}$$

This Pareto optimum is attained in a cooperative situation. In the competitive situation, the firms maximize individual profits, resulting in the condition

$$p_1 \frac{\partial F_1}{\partial L_1} = p_2 \frac{\partial F_2}{\partial L_2} \tag{5}$$

which is equivalent to (4) if there is no externality ($\partial F_2 / \partial L_1 = 0$). Clearly, in the presence of externalities the competitive solution is not Pareto optimal. To illustrate the Coase theorem, note that because of the externality firm 2's maximum profit is reduced by the difference between its maximum profit when L_1 is 0 and its maximum profit when L_1 is at firm 1's optimum level. Call this difference ρ. Suppose firm 1 is liable, and compensates firm 2 by paying ρ. Then, its profits are $p_1 F_1(L_1) - wL_1 - \rho$. Firm 2's profits are $p_2 F_2(L_1, L_2) - wL_2 + \rho$. Profits are maximized if the following conditions hold:

$$p_1 \frac{\partial F_1}{\partial L_1} - \frac{d\rho}{dL_1} = w, \text{ and } p_2 \frac{\partial F_2}{\partial L_2} = w$$

Combining these two conditions and noting that $d\rho / dL_1 = -p_2 \partial F_2 / \partial L_1$, the Pareto optimum condition (4) is found, which confirms the Coase theorem.

Although the Coase theorem is a remarkable results, stating that government intervention is not necessary to reduce the negative effects of externalities, it is important to be aware of the underlying assumptions. The most important one is mentioned already, the absence of transactions and bargaining costs. Apparently, in many cases there are huge costs involved in settling disputes between firms, e.g. legal costs. Another problem is that, although the distribution of property rights does not matter for the outcome of the negotiations, it does matter to determine which of the parties is liable and compensates the other party. In many cases, however, it is not clear how property rights are distributed. Still, the Coase theorem shows that it is possible, under

[7] This is the situation considered by Coase (1960).

certain conditions, to achieve a Pareto optimum without government interventions. Although the theorem is often used as an argument against government regulation, except for providing the necessary legal infrastructure, there is no dispute in the literature that its underlying assumptions are rather strong. Hence, there still is a case for government intervention to accommodate the effects of externalities (see, e.g., Mueller, 1989).

3.5 Market power

In a competitive market, firms take the price of their product as given and there is free entry and exit to and from the industry. Some industries, however, may be characterized by market power, i.e., firms are able to influence the price of their product. This is the case if the number of firms in the market is comparatively small. For simplicity, we restrict the discussion here to the case of a single firm. A monopolistic firm's profits attract other producers, so in the longer run, a monopoly can only exist if there are entry barriers. In general, there are two types of entry restrictions: legal and technical. Technical restrictions may be caused by economies of scale: production is only efficient at a scale that other producers cannot reach. Legal restrictions protect some kind of industry, either because the regulator explicitly wishes to limit the number of producers, or because of side-effects from regulation for other purposes. A specific explanation for legal restrictions is given by the economic theory of regulation, which explains regulation by rent-seeking activities of interest groups (see section 4, below).

As defined in section 3.2, a Pareto efficient situation exists if there is no possibility to make someone better off, without reducing someone else's welfare.[8] It can easily be seen that a monopolistic industry is not Pareto efficient. Figure 1 depicts a demand curve *D*, with a related marginal return curve *MR*. The monopolist's marginal cost curve is *MC*. At price p_m the monopolist receives rents, with a value of the area between p_m, y_m and *MC*. Compared to the competitive situation, p_c and y_c, consumers loose part of their surplus, of which rectangle *R* goes to the monopolist and triangle *L* is a deadweight loss, i.e., a loss to society. At output y_m, however, the monopolist could sell an extra unit of output at price p_m, making an extra profit. This could continue, as long as the monopolist is allowed to reduce the price for each additional output along the demand curve, thus to discriminate in pricing. Once price p_c is reached no further improvements are possible.

Regulators interested in correcting market failures, in order to approach Pareto efficient allocations, could consider two types of regulations. The first type of regulation is aimed at reducing

[8] This definition is equivalent to the one in 3.2.

entry barriers. The argument is that reduced legal or technical barriers attract other firms to the monopolistic industry. This is supposed to reduce the market power of the monopolistic firm, and reduce the welfare loss. Another type of regulation is aimed at price levels. Formulated simply, the regulator subsidizes the monopolist's output, in order to increase output and consumption of the good to a level as close as possible to the competitive level. A more advanced analysis of regulation of monopolies lies beyond the scope of this paper; see, e.g., Laffont (1993), Laffont and Tirole (1994).

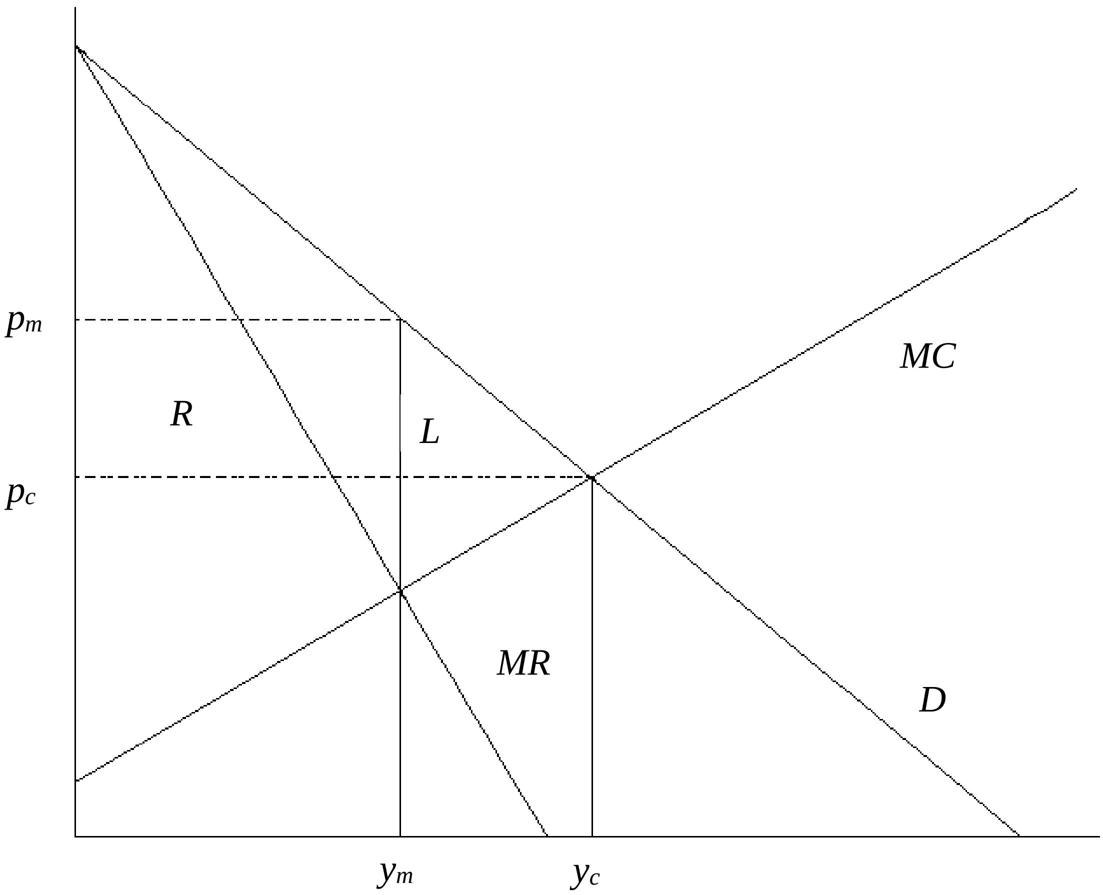

Figure 1 Monopoly power and Pareto efficiency

3.6 Informational asymmetry

3.6.1 Hidden knowledge and hidden action

As we will show in this section, informational asymmetry may lead to market failures. First, it is important to stress what is meant by information. Information can be regarded as a stock or as a flow concept. Usually, it is meant as a stock, giving an amount of knowledge (data, evidence) about the world. Information may also be viewed as a message, or a news item, which increases the stock of knowledge (Hirshleifer and Riley, 1992). Additionally, information has an economic value, since it is a scarce good. Some information is known to everyone, public information, other information is known to one or more individuals. In a perfect market all information is public. With private information, not known to everyone, at least one of the agents in an economy is less informed about some particular piece of information than the other parties.

Asymmetric information concerns situations where some participants are more informed than others. When a financial institution invests its deposits in a firm, for example, the firm is more informed about the safety of the investment than the financial institution. As another example, the depositors of the financial institution are not completely informed how much effort the financial institution makes in selecting the safest investment opportunities. In the first example, there is asymmetric information that concerns the knowledge of both parties on issues that are relevant to the transaction. This type of asymmetry is also called hidden knowledge. Hidden knowledge may result in *adverse selection*. The second example of asymmetric information is an example of hidden action. This may result in *moral hazard*. Both situations frequently arise in the regulation of financial institutions. In all cases the problem to the party that is less informed is to reach an agreement – a contract – that is aimed at reducing the effects of informational asymmetry. As an introduction to models with asymmetric information applied to the financial sector, we will first consider adverse selection and moral hazard in a more general sense.

3.6.2 Adverse selection

Adverse selection is also known as the 'lemon principle', since a seminal article by Akerlof (1970) applied asymmetric information to the market for used cars, where bad cars are known as lemons. Sellers would know the quality of their used car while the buyers do not. In general, with hidden knowledge there are two points to notice (Hirschleifer and Riley, 1992). First, if a price reflects the *average* quality of a product, there will be a tendency for the highest quality to self-select out of the market. Second, although this adverse selection may be severe, it is only in special cases that merely the lowest-quality remains present in equilibrium. The product referred to could be a used car, but other well-known examples are insurance policies and bank loans.

The concept of adverse selection can be illustrated by an example concerning credit rationing in financial markets with asymmetric information.[9] This model, developed by Stiglitz and Weiss (1981), contains the basic elements of more advanced financial market models with adverse selection like those considered, for example, in the sections below. Stiglitz and Weiss consider a

[9] For applications to the insurance market see, e.g., Laffont (1989) or Lippman and McCall (1981).

competitive banking system in which the supply of funds increases with the interest rate. There are two types of firms, type θ_1 and θ_2, both undertaking an investment project that requires the same amount of funds. Each firm i undertakes one investment project i, which has a return R_i if successful and zero if it fails. The probability of failure is q_i. Assume that $q_1 < q_2$ and $R_1 < R_2$. In other words, the θ_1-type firm is a low-risk, low-return firm, compared to the other firm. Each firm borrows the same amount of funds B with collateral C, where the interest rate is r and $B > C$. If project i is successful, firm i receives $R_i–(1+r)B$, and if it fails it loses C. Hence, the expected value of project i can be written as: $y_i \equiv (1-q_i)[R_i – (1+r)B] – q_iC$.

Assuming $(1-q_1)R_1 < (1-q_2)R_2$, it can be shown that $y_1 (r) < y_2 (r)$, for each r. Let r' and r'' be defined by $y_1 (r')=0$ and $y_2 (r'')=0$. Then the following relations exist:

(a) $r < r' \Rightarrow$ both projects will be undertaken; hence, firms of both types borrow funds;

(b) $r' < r < r'' \Rightarrow$ only project 2 will be undertaken; hence, only type θ_2 firms borrow funds.

Consequently, a higher interest rate selects the high-risk firms. In addition, a higher interest rate, ceteris paribus, increases the bank's expected revenue, z_i, which also depends on the type of project that is undertaken: $z_i \equiv q_iC+(1–q_i)(1+r)B$. The bank prefers lending money for project 1, since $z_1(r)>z_2(r)$. However, the bank does not have the same information as the firm. The bank only observes that firms require an amount B. It cannot charge different interest rates for different projects. It could have a higher return if r increases, but if r increases from a low level (below r') to a level between r' and r'', only firms of type θ_2 remain.

Now suppose that the bank knows that the proportion of type θ_1 firms is α, and that the bank finances n projects. Then the bank's expected profit is

$$\pi(r) \equiv \alpha n z_1(r) + (1-\alpha) n z_2(r) - nB$$

At a higher interest rate z_1 and z_2 rise. Also, at higher r the bank wants to finance more projects, i.e., n rises. Thus, expected profits rise as well. However, due to the higher interest rate, eventually only the type θ_2 firms remain, and expected profits decrease (because $z_1(r)>z_2(r)$). So, the bank's profit-maximizing interest rate can be below the market rate. However, every bank prefers the lower interest rate. Consequently, at the lower interest rate there is an excess demand for loans, and rationing of credit. In this model there is adverse selection due to informational asymmetry between firms and banks. Because the firms cannot take any action that affects the success probability of their investment project, there is no moral hazard in this model.

As, e.g., Tirole (1988) argues, simple models of adverse selection, like the Stiglitz-Weiss model, concern ‘one-shot relationships’. By ‘repeat purchases’ banks or other agents could acquire more information about the quality of a product. Another possibility to avoid adverse selection is to specify contracts with more parameters than just a price or interest rate. An example in the insurance market is charging different premiums depending on the extent of coverage. In general, buyers, like insurers, can use *screening* to gain information on quality or risk involved in the transaction. Sellers are interested in *signaling* the quality of their product or the probability of success of their project, because low risk and high quality sellers receive a low price.

3.6.3 Moral hazard

If there is informational asymmetry, one party’s actions may be hidden to the other party. The possibility that the more informed party takes actions that affect the expected return of the less informed party has been labeled moral hazard, although ethical considerations are not involved here. Hidden knowledge or adverse selection induces moral hazard in many cases: the asymmetry of information induces certain people to take advantage of it to the detriment of the other party without the other party knowing it (Takayama, 1994). Examples are the insurance market, where the insured may omit preventive action to reduce their risk.

In many cases, moral hazard situations can be formulated as a *principal-agent problem*. According to this approach an agent acts on behalf of a principal, who cannot completely monitor the agent’s actions. The problem is to design a contract that motivates the agent to act in his own interest (an incentive compatible contract), thereby maximizing the principal’s return. The basic structure of a principal-agent model is as follows (Tirole, 1988). Suppose that a manager is employed by risk-neutral shareholders. The manager’s (agent’s) actions are unobservable to the shareholders (principal), who are interested in maximizing profits. Profits equal returns minus wages, where returns are determined by the manager’s effort and some random variable. The shareholders reward the manager with a wage rate that depends on the level of returns. The risk-averse manager’s utility depends on his wage and his effort. Given these assumptions, the shareholders must design an incentive compatible wage structure, that induces an effort level from the manager that maximizes the shareholders’ profits. In other words, that effort level maximizes the manager’s utility.

In general, it is a complex task to solve the principal-agent problem (see, e.g., Holmström, 1979). Moreover, there are very few general results for the moral-hazard problem (Grossman and Hart, 1983). Additional assumptions, depending on the specific application of the model, are required for more precise results. Laffont and Tirole (1993) and Laffont (1994) apply the principal-agent methodology to contractual relationships between regulators and regulated firms. Applications to the financial sector and banks in particular are reviewed in section 3.3.4 below.

4 A POSITIVE THEORY OF REGULATION

4.1 The case for government regulation

According to the Second Theorem of Welfare Economics (section 2.2.2), any Pareto efficient allocation can be achieved through the price mechanism if initial endowments are redistributed accordingly. The allocation that can be achieved if lump-sum taxes and transfers are freely variable is referred to as the first-best solution. In general, however, the information that is required for this redistribution is unobservable or too costly to obtain. Hence, there may be binding constraints on the use of lump-sum redistributions. In that case the solution is called second-best. The theory of second-best policy seeks to specify the best choice of instruments for some policy objective (e.g., Atkinson and Stiglitz, 1980; Laffont, 1993). Up to this section, it is assumed that there is a case for government regulation. In other words, the theory of second-best is a normative theory. Another approach is followed by explaining the existence of regulation through the assumption of economic behaviour by selfish agents or interest groups.

4.2 Rent-seeking through regulation

The ‘economic theory of regulation’, and the theory of industrial organization in general, are based on the Chicago tradition, of which Stigler and Posner are the most prominent authors in this area. This school of thought emphasizes the need for rigorous theoretical analysis and empirical identification of competing theories. Following this methodology, Stigler analyzes regulation from a positive point of view. His case study on regulation of electric utilities (Stigler and Friedland, 1962) and later, more theoretical papers (Stigler, 1971; Posner, 1971) apply concepts from public choice theory to regulation. Peltzman (1976) took up the task of formulating Stigler’s model in a more rigorous mathematical manner. Becker (1983), focussing on pressure group formation, took further steps in developing the Stigler-Peltzman approach to regulation (see below). In brief, this approach stresses the rent-creating powers of regulators and the rent-seeking efforts of those regulated. It explains how regulation itself may lead to market failures with extensive welfare losses, instead of correcting those failures. An important characteristic of regulation that emerges from this literature is that well-organized interest groups tend to benefit more from regulation than broad, diffuse groups, which probably creates a bias in favor of producers, who are usually well-organized relative to consumers (Peltzman, 1989).

Posner (1974) gave the Stigler-Peltzman approach to regulation the name ‘economic theory of regulation’. The Stigler-Peltzman approach is also called ‘regulatory capture theory’, and as shown by Mueller (1989) it can be considered as an application of the rent seeking concept from public choice theory. According to Stigler and Peltzman, rent seeking explains how regulation is supplied by self-interested politicians, who try to achieve political support from voters, who may benefit from regulation. Suppose the number of votes received by a politician, *V*, is a function of the utility of regulated producers, U_R, and the utility of consumers of the regulated product, U_C:

$$V = V(U_R, U_C), \text{ with } \frac{\partial V}{\partial U_R} > 0, \ \frac{\partial V}{\partial U_C} > 0$$

Assume for simplicity that utilities are linear, and depend on consumer surplus captured by producers (rents, *R*) and consumer surplus lost (*L*); see also Figure 1: $U_R,=R$ and $U_C=K\text{-}R\text{-}L$, where *K* is a constant. Assuming an interior maximum exists, the vote maximizing regulator sets the price of the regulated product such that:

$$\frac{\partial V}{\partial P} = \frac{\partial V}{\partial U_R}\frac{dR}{dP} - \frac{\partial V}{\partial U_C}\frac{dR}{dP} - \frac{\partial V}{\partial U_C}\frac{dL}{dP} = 0$$

which is equivalent to

$$\frac{\partial V}{\partial U_R}\frac{dR}{dP} = \frac{\partial V}{\partial U_C}\left(\frac{dR}{dP} + \frac{dL}{dP}\right) \tag{6}$$

According to (6), the marginal gain in support from producers for higher monopoly rents is just offset by the loss in consumer votes resulting from the loss in consumer surplus. Many authors argue that the gain in votes resulting from producers is larger than the loss in votes from consumers, because of the much larger number of consumers, which makes it difficult to them to organize themselves and reduces the individual consumer rents. Note, however, that the regulator does not set the price at the rent-maximizing level preferred by producers, where $dR/dP=0$. Instead, he brings the price to a level somewhere between the pure monopoly and perfect competition prices. If oligopoly prices lie between monopoly and competitive levels, oligopolists and their consumers have less to gain from regulation than do the consumers of a natural monopoly product, or the producers of a competitive product. This could explain the widespread regulation of agriculture and other seemingly competitive industries.

There are many empirical applications of the economic theory of regulation. Most of the applied studies are restricted to a specific area of regulation. Empirical evidence for rent seeking through regulation is found in several cases, including the regulation of stock exchanges (Schwert, 1977). In a more general empirical analysis, Winston and Crandall (1994) try to find empirical support for the economic theory of regulation, explaining federal regulatory policy in the United States during 1900-1992. They find that regulatory policy influences the share of the popular vote in presidential elections. The results can be interpreted as voters responding to changes in regulatory policy in a manner that reflects consumer welfare. However, it seems that presidents also pursue their own ideology and do not always respond to voters' preferences regarding regulation. This makes it difficult to predict the evolution of regulatory policy.

4.3 Costs and benefits of regulation

Another approach to the economic theory of regulation is concerned with cost-benefit analyses of regulation. Many studies in this area were done in 'the age of deregulation' (Chang, 1997), which started around 1980. In calculating the overall economic costs and benefits of regulation it is important to distinguish between transfer payments and net changes in economic efficiency. Hahn and Hird (1991) find large differences in cost and benefit estimates for particular regulatory policies. Although they do not consider financial industries, their analysis makes clear that a cost-benefit analysis of regulation in the financial sector would encounter large methodological difficulties. Estimation of benefits, especially, suffers from lack of data, and is as much an art as a science. Viscusi (1997) stresses the importance for cost-benefit analysis of converting non-monetary consequences of regulation to a monetary value. If some type of regulation is analysed by comparing costs to benefits, one should try to attach a monetary value to all possible outcomes of the regulation. Instead of a monetary cost-benefit analysis, which is not feasible in all cases, a more limited test might be considered, comparing the risks generated by the regulation with the risks reduced by the regulation. Moral hazard involves an important type of trade-off between different risks that should be included in risk-risk analysis (Viscusi, 1997).

4.4 Concluding remarks on methodology

Regulation can be analyzed by a normative (section 3) or a positive approach (section 4). The purpose of regulation, according to the first approach, is to reduce welfare losses due to market failures. The second approach, in this section, gives an explanation for regulation that takes economic behaviour of the regulators into account. Also, it can be viewed as a warning that regulation might have negative side-effects, because regulators could be interested in creating welfare-reducing monopoly rents. Instead of the general treatment of regulation up to this section, the following sections are focussed on the financial sector. The analysis so far does not give an indication whether one of the two approaches, normative or positive, might be more fruitful in analyzing regulation of the financial sector. Obviously, the two approaches do not exclude each other, so both are interesting point of departure to analyze the regulation of banks. In the next section it is argued, however, that market failures are especially important in the financial sector. Therefore, most of the theoretical literature on regulation of banks takes a rather normative approach.

5 REGULATION OF FINANCIAL INTERMEDIARIES

5.1 Introduction

Financial intermediaries constitute a heavily regulated sector of the economy. In microeconomic theory there is an ongoing debate about the reasons for regulation of banks and other financial institutions, and about the form regulation should take. Macroeconomic theory stresses that the sound functioning of the banking system is vital for macroeconomic performance. The recent experience in Asia and those, in the 1980s, in Latin America, illustrate this. Especially those working in the field of monetary economics are aware of the danger of systemic risk and of the role of not only monetary policy, but also banking regulation in preventing or mitigating this risk. Finally, economic growth is shown to be correlated significantly with the degree of financial development. Although one can argue about the direction of causality here – some argue, with Joan Robinson, that “where enterprise leads, finance follows” – the correlation itself makes this area of research important for macroeconomists too.

Most studies analyzing regulation of the financial sector depart from the notion that the existence of asymmetric information plays a key role. Although Gurley and Shaw (1955) stressed the importance of financial intermediation in improving the efficiency of intertemporal trade, the attention paid in economic theory to financial variables vanished away after Modigliani and Miller (1958) in their famous theorem had shown that, in a perfect market, real variables are not affected by financial ones. It was not until 1970 that the importance of financial intermediation reappeared prominently in economic theory (Kroes, 1996). This was mainly due to the attention for the concept of asymmetric information, following Akerlof’s (1970) paper (see section 3.6). From that time on, the micro-economic approach to financial markets blossomed. Moreover, the importance of the financial structure for macroeconomics was emphasized by Bernanke’s (1983) study of the Great Depression. And in the model developed by Stiglitz (1988), credit (and its rationing) plays an important role in the transmission of monetary policy.

Although regulation may help to stabilize the financial system, the degree to which the banking sector should be regulated and the best way to do it is subject of discussion. Thus, the safety net provided by deposit insurance may induce banks to take too much risk. Some authors (e.g., Mailath and Mester, 1994) argue, that regulation may not always be credible. This section gives an overview of the theory of financial regulation. In subsection 5.2 an overview is given of the arguments in favour of and against banking regulation. Subsection 5.3 introduces various theoretical models of banking regulation.

5.2 The justification of bank regulation

Banks and other financial intermediaries (e.g., insurance companies) are much more regulated than other sectors of the economy. As pointed out in the previous sections, government regulation is justified by market failures due to market power, the existence of externalities or public goods, and asymmetric information between buyers and sellers. For the financial sector, the two latter aspects are especially important. The failure of any type of firm creates externalities. However, in the case of financial intermediaries this problem is more acute for a number of reasons (Fase, 1995; Freixas and Rochet, 1997). The banking sector is crucial for the economy's payments system. If this system is harmed, the economy as a whole suffers. The provision of a means of payment is to be regarded as a public good, and therefore the government has the responsibility to protect the financial system. It is even more important to prevent bank panics, since they tend to be contagious. This so-called systemic risk is defined by the BIS as 'the risk that the failure of a participant to meet its contractual obligations may in turn cause other participants to default, with the chain reaction leading to broader financial difficulties' (BIS, 1994, p. 177). Moreover, the debt of banks is held by a large number of small agents. As a result, an important characteristic of banks is, that their creditors are also their customers, as they make use of payment services. In addition, compared to other firms, banks have a high debt-to-asset ratio (high leverage) and low cash-to-assets ratio (fractional reserve banking). In the case of banks this may imply, as will be shown below, that bank managers tend to choose investment policies that are too risky from the point of view of depositors. Because small depositors are unorganized and underinformed, the delegation of monitoring to a supervisor or regulator improves efficiency by creating economies of scale. Apart from the economic inefficiency, there would be a serious free-rider problem in the absence of a specialized institution: since any monitoring individual incurs costs, whereas the benefits go to all depositors. As a consequence, there would be insufficient monitoring of banks.

The characteristics enumerated above make the financial system fragile and subject to panics. According to Goodhart *et al.* (1998), the systemic dimension is important in the financial sector because the social costs of financial distress exceed the private costs to those involved in the failing institution (shareholders, managers, employees). However, regulation itself has a cost that should be weighted against its benefits. Also, the provision of a safety net for depositors may induce moral hazard on the part of banks, promoting instead of reducing bank failures, as was the case with the savings and loans crisis in the United States. Besides this, there are opponents to bank regulation who argue that, first, the danger of bank runs is extremely small and, second, that in the few cases observed in reality they have had no larger impact on economic performance than failures of nonfinancial firms (Benston and Kaufman, 1996; Kaufman, 1996). Hence, in their view, there is nothing special about banks that justifies any special regulation. Moreover, Kaufman (1996) argues that prudential regulation is frequently inefficient and counterproductive, increasing rather than decreasing both the probability and the cost of bank failures. Nevertheless, he admits that contagion in banking may be faster and spread more widely, though in his view depositors' losses are quite small. Also, he contends that large-scale banking failures may exacerbate economic downturns, but do not start them. It should be stressed that Kaufman does not explicitly pay attention to the possibility of inefficient bank runs, also because in his view depositors are able to differentiate between solvent and insolvent banks. Others object to the view of Kaufman *et al.* by pointing out that even if the probability of failure is extremely small, the consequences may be so large as to justify regulation, which is then comparable to an insurance premium against a low-probability risk (Goodhart *et al.*, 1998). Moreover, a systemic crisis is very difficult if not impossible to control.[10] It should be stressed, that the lender-of-last-resort function of the central bank is primarily directed at preventing systemic crises. Some (e.g., Bagehot, 1873) argue that the central bank should lend only to illiquid solvent financial institutions. Others (e.g., Goodhart, 1987 and 1995) state, that the distinction between illiquidity and insolvency is hard to establish in practice, because banks that have entered a situation of illiquidity will in time become insolvent as well.

5.3 Regulatory policies

5.3.1 Introduction

Goodhart *et al.* (1998) distinguish between regulation (establishing rules of behaviour), monitoring (observing whether rules are obeyed) and supervision (the oversight of the behaviour of the financial sector). This section is limited to the establishment and monitoring of regulatory

[10] In this respect, the role of central banks also deserves attention. Evidence by Miron (1986) shows that after the creation of the FED the frequency of bank panics decreased. However, this may have been the result of interest rate policy rather than regulation. Still, it shows how public intervention may lead to outcomes in the banking sector that are preferable to those resulting from the market mechanism.

rules for banks. According to many authors, bank regulation is much more important than regulation of non-bank financial intermediaries (pension funds, insurance and life insurance funds). Some even argue, that it is typically the combination of deposits that serve as means of payment – and hence can be withdrawn any moment – and loans that cannot be sold in the market is the only justification for regulation (Freixas and Rochet, 1997, ch.1). This is because banks are the only source of finance for a large number of borrowers, and because banks manage the payments system (Bernanke, 1983). In addition to this, the interconnections between different banks, e.g., through interbank deposits, is greater than that in other sectors, implying that one bank failure may lead to another. This is why Rochet and Tirole (1996) argue that not the size of a bank, but its financial connections with other banks may provide the criterion for a bailout. Another reason for more regulation of banks than of non-bank financial intermediaries is, that bank loans cannot be easily traded in a secondary market, whereas the assets of security firms can. However, this does not imply that there is no systemic risk at all in securities markets and insurance companies. The concentration in the securities industry and, especially, the links between the securities market and the banking system may constitute a danger of destabilisation of the financial system (Goodhart *et al.*, 1998; OECD, 1993).

Regulatory policies typical for the banking sector may include (Freixas and Rochet, 1997):[11]

- Deposit interest rate ceilings
- Portfolio restrictions
- Deposit insurance
- Capital requirements
- Regulatory monitoring
- Closure policy.

The justification of deposit rate ceilings is that these reduce the rates charged by banks to their borrowers, and that this is one way of controlling the volume of loans. The underlying mechanism is that banks use mark-up pricing. Regulation theory has not paid much attention to this subject, perhaps because it has more to do with macroeconomic policy than with regulation of the financial system. Still, one could argue that ceilings on interest rates may be used to limit the competition, by offering high interest rates, for deposits in case of a deposit insured banking system (see below).[12]

Portfolio restrictions limit the activities of banks in the sense that a distinction is made between commercial banks (allowed to hold demand deposits) and investment banks (allowed to hold

[11] Financial intermediaries may of course be subject to entry and merger restrictions, but these are not typical for this sector and will not be discussed here.
[12] Freixas and Rochet (1997) show that a ceiling on interest rates may not affect lending rates.

corporate equity). Arguments of those favouring separation are that the holding of equity may increase the bank's riskiness, and that there may be a conflict of interest because banks that have made loans to firms may be willing, when these firms get in trouble, to underwrite the firms (bad) securities. On the other hand, universal banks may have economies of scope and information advantages. This type of regulation does not exist in Europe, where banks are universal and allowed to employ both types of activities. Therefore, no further attention is paid to this type of intervention in the banking sector. The remainder of this subsection focuses on regulatory policies in the form of deposit insurance (5.3.2), capital requirements (5.3.3), capital requirements and regulatory monitoring (5.3.4) and monitoring and bank closure policy (5.3.5).

5.3.2 Deposit insurance

The purpose of a deposit insurance scheme is to avoid bank panics and the social costs that come with them. The mechanism is that banks pay a premium to a deposit insurance company.[13] In return for this, in case of a bank failure, depositors receive their deposits up to a fixed maximum amount. Usually, the deposit insurance fund is a public institution. However, some economists argue that private insurance companies could provide this service equally well or even better (Mishkin, 1992). The advantage of private schemes, they argue, is that competition leads to accurate pricing and optimal information extraction. Still, an important disadvantage of a private insurance scheme is that the insurance company itself is of course subject to the systemic risk. Hence credible insurance implies a backing by the government. Recent experiences in the Unites States (e.g., Ohio) have indeed shown that private deposit insurance companies failed along with the banks whose deposits they insured. In the end, the government had to intervene anyway. In the past, there have also been implicit deposit insurance schemes in some European countries. These entailed that after a bank failure, the government directly intervened to pay depositors. Kaufman (1996) even observes, that where there is no explicit deposit insurance scheme, the government offers in practice a 100 percent insurance.[14]

Deposit insurance and bank runs

Diamond and Dybvig (1983) use a three-period model with consumers who dispose of one good at t=0 and whose utility depends on consumption. The good can be freely stored or invested in an illiquid technology which gives return R>1 at t=2, but only L<1 if the investment (I) must be liquidated at t=1. Liquidation at t=1 is necessary for 'type 1' consumers, that is consumers who

[13] Alternatively, the deposit insurance fund may be financed by the taxpayer.
[14] This is why Kaufman is in favour of explicit, but limited deposit insurance.

are subject to a liquidity shock. Ex ante, every consumer knows his probability of becoming a type 1, π_1. In order to protect themselves against this risk, consumers will invest a suboptimally low amount if there is no trade or banking system. The consumption by type-1 and type-2 consumers will be:

$C_1 = 1 - I + LI \leq 1,$
$C_2 = 1 - I + RI \leq R$

with a maximum for type-1 consumers of 1 if zero is invested, and a maximum of R for type-2 consumers if they have invested the whole good. The exact amount invested by the consumers depends on their (expected) utility function:

$U = \pi_1 u(C_1^{\,1}) + \pi_2 \rho u(C_2^{\,2})$, with ρ<1 as discount factor.

Introduction of a bond market at t=1 leads to a Pareto-improvement, because, as at that time everyone knows his type, no liquidation is necessary. Type-1 consumers will sell bonds to type 2 consumers in exchange for liquidity, and the equilibrium, where the price of bonds equals p=1/R, implies that C_1=1, C_2=R and I=π_2. However, this is not the optimal allocation, because that would have to satisfy $u'(C_1^*) = \rho R u'(C_2^*)$ (equality of marginal utility of consumption). The intuition is, that optimal insurance against liquidity shocks enables type-1 consumers to consume more. A bank can take care for this Pareto-optimum. In exchange for the deposit of the good, the bank offers every consumer the possibility to consume at t=1, if necessary the Pareto-optimal amount. This guarantee can be given because the bank stores a fraction of the goods corresponding to the probability of liquidity shocks, π_1, times the optimal first-period consumption, C_1^*, and invests the remainder. This gives every consumer ex ante his optimal insurance against a liquidity shock.[15] A crucial assumption for this result is, that type-2 consumers do not withdraw their deposits prematurely, that is at t=1. This assumption is reasonable as long as the extra return on the investment (postponing of consumption) more than compensates for the consumer's impatience, i.e., $R\rho$>1. However, if for some reason some consumers may want to withdraw at an early stage even if they have *not* undergone a liquidity shock, the bank has insufficient resources. The most likely reason for this type of withdrawal is a lack of confidence in the bank. In the model, this can be represented by an increase in the degree of impatience of type-2 consumers, as a reflection of their doubt whether the return on the investment will indeed be R. If some type-2 depositors withdraw, it becomes attractive for others to do the same. Hence,

[15] Obviously, as with any insurance system, ex post this is not an optimum for those who turn out not to suffer from a shock.

the belief that others may decide to withdraw leads to behaviour that generates a self-fulfilling prophecy (Wit, 1998). Coordination among depositors might prevent a bank run, but is difficult to realize. Deposit insurance, which guarantees that depositors receive the full value of their deposits, is a solution to this problem. The mere existence of this insurance can prevent bank runs, in which case the insurance system does not have to intervene. The problem is, that the existence of deposit insurance may adversely affect the behaviour of banks, creating a moral hazard problem.

Deposit insurance and moral hazard

Although deposit insurance may prevent bank runs, its drawback is that it creates a moral hazard problem. Banks that have their deposits insured are inclined to invest in projects that are too risky in the sense of having a low probability of success. This is because if the projects fail, the cost is for the insurance company, whereas if it succeeds the high payoff is for the bank. This is also optimal behaviour in a competitive environment, as the banks can in this way attract more customers by paying a higher interest rate. The depositors, for their part, do not have an incentive to monitor because their deposits are insured anyway. A simple two-period model can illustrate this (Freixas and Rochet, 1997).

At t=1 the bank receives deposits D, issues equity F_1, makes loans L_1 and pays an insurance premium P. At t=2 the bank is liquidated and its stockholders receive the liquidation value V_2:

$$V_2 = L_2 - D + S_2 \tag{7}$$

where L_2 is the repayment on the loans and S_2 is the payment by the deposit insurance company. Hence S_2=0 in case of success and S_2=D–L_2 in case of failure. Using the first-period balance sheet, equation (7) can be rewritten to find the increase in the bank’s value:

$$V_2 = F + (L_2 - L_1) + (\max\,[0, D_1 - L_2] - P)$$

Here the first term on the right hand side is the expected increase in the loans value and the second right hand side term is the net receipt from the deposit insurance. If it is for simplicity assumed that the loan repayment is X in case of success and 0 in case of failure, the deposit insurance payment is max (0,D_1). Hence if the probability of success is θ, the expected insurance payment is $(1-\theta)D_1$. The expected profit for the bank’s stockholders is the expected liquidation value minus the equity capital. Given the assumptions made, this can be written

$$E(V_2) - F_1 = (\theta X - L_1) + ((1-\theta)D - P) \quad (8)$$

where the second term is the expected net receipt from the deposit insurance.[16] Obviously, if the stockholders/owners are able to choose the projects they finance, and if bankruptcy does not bring about any additional cost to them, they will select the combination (θ, X) that maximizes the expected profit given by equation (8). If it is assumed that they can choose among a range of projects with a *given* net expected value $\theta X–L$, they will choose the project with the highest risk, as θ appears with a minus sign in the second term of (8). This is regarded as the moral hazard problem of deposit insurance. However, it is the limited liability that generates the incentive for the bank to choose the riskier project (Mailath and Mester, 1994). If there is unlimited liability for the stockholders, net expected profit in the absence of deposit insurance would be $\theta X–L$ and the bank would be indifferent between the projects. However, in practice there is limited liability and stockholders can at worst have a liquidation value of zero. In the absence of deposit insurance, moral hazard would be tempered through the depositors' monitoring of bank behaviour. The crucial difference with deposit insurance is that it takes away the depositors' incentive to temper the bank's moral hazard. Note, that the moral hazard issue arises even though the bank owners have their own capital F at stake. It should be stressed, that a crucial assumption for the incentive to take more risk if deposits are insured is, that the insurance premium rate P/D is unrelated to the risk taken by the bank. This is why a shift to risk-related insurance premiums is advocated as one of the solutions to the moral hazard problem aggravated by deposit insurance.[17]

Merton (1977) shows that the deposit insurance payment S_2 can be regarded as a put option on the repayment of the bank loans L_2 with exercise price D_1. Therefore, the arbitrage pricing method is applicable to deposit insurance as well – provided, that the condition of perfect and complete financial markets is met. In line with the Black-Scholes formula the optimal deposit insurance premium is an increasing function of the deposit-to-asset ratio and of the volatility of the loan repayments L_2: $(P/D)^* = f(D/L, \sigma_L)$. The problem may be, that the deposit insurance fund is likely to have an information disadvantage about the volatility of the bank's assets, σ_L. As market imperfections constitute the main justification of regulatory policy, one can doubt the relevance of intervention regimes based on the assumption of complete information (de Lange, 1992). Fairly priced deposit insurance would be possible if θ were observable, since $P^*=(1–\theta)D$. However, Chan, Greenbaum and Thakor (1992) show, that if θ is private information of the

[16] If insurance would be fairly priced the premium would equal the expected probability of failure times the value of the deposits (P=(1–T)D) and then the second term in equation (1) would be zero.
[17] Others include market value accounting and additional capital regulation (see below).

bank, an incentive-compatible premium, $P[D^*(\theta)]$, with D^* at a profit maximizing level, does not exist. In this case there is asymmetric information, which leads to adverse selection.

It can be concluded that deposit insurance takes away the incentive for depositors to monitor the bank, thus making risky behaviour easier. If interest rates on deposits are allowed for, banks are likely to engage even more in risky activities in order to attract more depositors by giving a higher rate of interest.

5.3.3 Capital requirements

In the previous subsection, attention was paid to regulatory policy aimed at influencing the behaviour of depositors. By ensuring repayment of their deposits, regulatory authorities may prevent a liquidity crisis of a solvent bank. However, regulatory authorities are also interested in maintaining a sound banking system, thus preventing the probability of bank failures due to insolvency. Hence whereas deposit insurance schemes, by providing a safety net to depositors, are aimed at preventing bank panics and the bank failures resulting from them, capital requirements are directed at influencing the behaviour of bank managers in such a way as to prevent bank failures through insolvency. Obviously, this may indirectly influence depositors' trust in the banking system. The main objective of studies in this area of research is to find the optimal capital/asset ratio. Capital requirements are a good method to limit the risk of insolvency. If there is more at stake, for the bank, it is likely to take less risk. However, higher capital requirements may not always lead to more safety. There may be circumstances in which it even has 'perverse' effects, increasing rather than decreasing risk. This is because high leverage is not the only factor influencing insolvency.

The so-called portfolio approach analyses how the capital/asset ratio should be related to risk in order to induce banks to choose the desired portfolio strategy. Analogous to the theory of optimal deposit insurance pricing, this approach concludes that the optimal capital/asset ratio is an increasing function of the portfolio risk. Many authors state, that in order to model bank behaviour appropriately, one should take into account that the banking sector is an incomplete market, in the sense that one cannot trade away all risk.

Kim and Santomero (1988) model this characteristic by introducing risk aversion in the bank's objective function. Deposits are insured against a fixed-rate. As we have seen in the previous section, this type of insurance pricing is likely to create risky behaviour on the part of the bank. The regulator has the objective of reducing the probability of bank insolvency, and has capital requirements as his instrument. Kim and Santomero show that under these assumptions solvency regulation through simple capital ratios has the effect of increasing instead of decreasing the risk of the portfolio chosen by the bank. This is because although the *size* of the risky part of the portfolio decreases in reaction to the capital requirements, its *riskiness* increases. The intuition is the following. Each bank's strategy consists of choosing the proportion of each type of asset relative to the equity capital. The regulator – faced with a trade-off between economic efficiency on the one hand and safety on the other – chooses a capital adequacy ratio with the objective of setting an upper bound to the probability of bank insolvency. As argued in the previous section, because of deposit insurance, insufficient monitoring by depositors and/or limited liability, banks themselves may not take enough care for safety and for preventing insolvency.

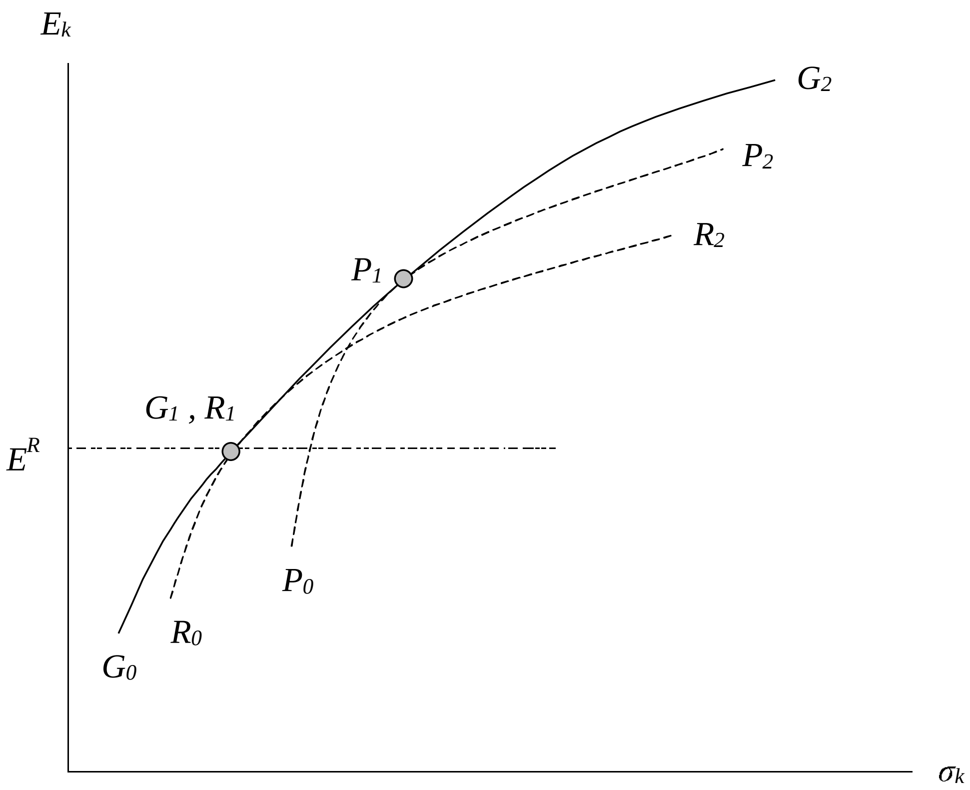

Figure 2 Capital regulation and the probability of insolvency

In the present analysis, it is assumed that the upper bound to the probability of insolvency is given, i.e., chosen by the regulator. The minimum capital requirement limits the feasible set of portfolios and the regulator hopes that all banks will make an adjustment toward a low-risk portfolio. However, banks with a relatively low rate of risk aversion will adjust their portfolios in such a way that the impact of forced *lower financial risk* (through lower leverage) is offset by *higher business risk* (riskier assets). This is illustrated in Figure 2, which has expected portfolio return per unit of equity capital, E_k, on the vertical axis, and the standard deviation of portfolio return per unit of equity capital, σ_k on the horizontal axis. For any leverage ratio k, there is an efficient portfolio frontier in terms of E_k, σ_k. An example of such a frontier, namely for $k=k_p$, is given by $P_0P_1P_2$ in Figure 2. Lower leverage, for example $k_r>k_p$, implies a frontier down to the left, e.g., $R_0R_1R_2$ in Figure 2. $G_0G_1G_2$ is the global frontier, that is an envelope of all efficient frontiers for varying k. Hence $G_0G_1G_2$ gives the efficient combinations of risk and return (measured per unit of equity capital) for different equity-to-asset ratio's k. Moving upwards and to the right along this frontier corresponds going to a portfolio with higher expected return and lower capital-to-asset ratio, and hence with more risk. Absent minimum capital requirements, banks are free to choose leverage and investment projects. A bank's actual choice depends on its rate of risk aversion. It is assumed, that the bank's objective function is strictly concave and is defined over mean and standard deviation of the return on equity capital. The regulator's purpose is to limit the probability of insolvency, which is defined here as the situation that $E\leq-1$, i.e., that the bank's equity capital is eliminated.If the regulator sets a certain minimum capital adequacy ratio, for example, k_r – limiting the allowable leverage – points on the G curve in between G_1 and G_2 are not feasible anymore. The aim of a minimum k_r is to make the banks move along the G curve to the left and downwards, choosing a portfolio with a lower standard deviation per unit of equity capital. However, despite the capital requirement k_r, banks are still allowed to choose a portfolio to the right of G_1 on the R-curve, as the R-curve is the curve connecting points representing different combinations of E_k,σ_k, for leverage ratio k_r. Whether a bank indeed chooses a portfolio between G_1 and R_2 depends on its utility function: banks with a low rate of risk aversion are more likely to do so (their indifference curves are more to the right in Figure 2). In that case they fulfill the capital requirement, but their insolvency risk is too high in the eyes of the regulator. As a solution to this problem Kim and Santomero derive optimal weights for risk-related capital adequacy requirements. These make sure, that the expected portfolio return does not exceed E^R. This is done in the following manner. For any asset that would bring the expected return of the portfolio per unit of equity capital above E^R, additional capital is required so that despite the introduction of this asset in the portfolio, the expected *portfolio return per unit of equity capital* is E^R. In other words, the capital requirements make sure that the expected return has an upward bound E^R. As banks are by assumption risk-averse, they will never choose a portfolio along the horizontal line to the right of E^R, because that would imply higher risk (measured by σ_k) but no higher expected return E^R. The appropriate weight for any asset i is shown to be:

$$a^*_i = (u_i - u_0)/(E^r - u_0) \quad \text{if} \quad u_i - u_0 > 0$$
$$a^*_i = 0 \quad \text{if} \quad u_i - u_0 \leq 0$$

where a_i is the number of units of equity capital that a bank should hold for one unit of the i-th asset in its portfolio (i.e., against one unit of asset i the bank can hold a maximum of $1-a_i$

deposits); E^r is expected portfolio return per unit of equity capital given by point G_1 in Figure 2 (the maximum allowed by the regulator); u_0 is cost of deposits; u_i are returns on assets i=1,…,n.

For high-risk, high return assets so much additional capital is required that the expected return *per unit of equity capital* does not exceed E^R. Assets with a zero or negative expected net return ($u_i \leq u_0$) have a zero weight because these assets would never be used to increase the expected return on the portfolio in response to capital requirements.[18] Campbell *et al.* argue, that as the optimal weights are independent of the bank's risk aversion, they can be uniformly applied to the banking sector as a whole. Still, as detailed knowledge is required, it is doubtful whether this policy is easy to implement in practice. Obviously, u_i is not observable, or is based on subjective bank judgment. Indeed, Kaufman (1996) argues, that the necessary information for risk-based capital requirements is too difficult to obtain and that "The risk classifications and weights adopted by the regulators to date have been arbitrary, incomplete, insufficiently reflective of the riskiness of the bank as a whole … and modified to pursue political and social objectives."

5.3.4 Capital requirements and monitoring

Campbell, Chan and Marino (1992) assume that heavy reliance on capital requirements, at the expense of demand deposits, for financial intermediaries is socially sub-optimal. As the liquidity insurance model of Diamond and Dybvig (1983) has shown, demand deposits, by serving as the medium of exchange, provide liquidity for otherwise illiquid assets. This facilitates the allocation of savings and reduces the rate of unnecessary project liquidation, thereby promoting economic growth and generating social surplus.[19] Campbell *et al.* use this idea in their analysis by assuming that demand deposits have value for the depositors not only because they earn a deposit interest, but also because they provide a liquidity service. They show that it is optimal to partially substitute capital requirements with regulatory monitoring. Monitoring entails that the regulator closely follows the behaviour of banks and overrules bank decisions if it wishes to do so. The problem is more complicated, however, because Campbell *et al.* assume that monitors have a propensity to shirk. The purpose of the analysis is thus twofold: to determine the optimal combination of capital requirements and monitoring, and to design an appropriate incentive contract for the monitor to guarantee that monitoring is applied optimally.

---

[18] The bank might hold these types of assets from the point of view of diversification, but this does not interfere with the issue of capital and insolvency regulation.

[19] In the absence of demand deposits offered by financial intermediaries, households would have to insure themselves against liquidity shocks, thus investing less, and/or would have to liquidate investment projects in case of unexpected liquidity shocks. See Diamond and Dybvig (1983), Freixas and Rochet (1997), Wit (1998).

The model assumes risk neutrality for all market participants and a representative bank in a competitive environment with access to two alternative investment projects. Both projects are risky, because they have a positive probability of a payoff of zero. However, project 1 has a lower risk and a higher expected return than project 2 and is therefore socially preferable. There is asymmetric information about the additional risk of the second project. That is, the bank knows the probability distribution of the payoffs for the second project, whereas the public (depositors, monitor) does not. The first problem analyzed here is one of finding, from the point of view of society (representing the depositors and acting as a principal for the monitor) the optimal combination of capital requirements and monitoring. The timing is as follows. First, depositors choose c, that is the fraction of investment financed with capital (hence, $1–c$ is the fraction of deposits). Then, the payment to depositors M is determined by the break-even condition for the banking industry. Of course, $M=(1–c)r$, where r equals one plus the deposit interest rate. Next, the bank observes the probability distribution on both projects where the expected payoffs and their probabilities are given by:

State	*Probabilities*		*Payoff*
	Project 1	Project 2	
1	θ_1	$\theta_1 + h$	0
2	θ_2	$\theta_2 - 2h$	x_2
3	θ_3	$\theta_3 + h$	x_3

where $\theta_1+\theta_2+\theta_3=1$ and where h (>0), which is a measure of the extra risk on project 2, is observed only by the bank. It is assumed that $2x_2>x_3$, hence project 2 has a lower expected return than project 1.[20] As a benchmark, Campbell *et al.* analyze how, in the absence of monitoring, depositors can use their instrument, the rate of capitalization c, to influence the bank's project choice. Because higher capitalization implies less deposits, it reduces liquidity service. Depositors will choose a positive c only if this induces the bank to choose the less risky project. Two cases can be distinguished. If the repayment to the depositors, M – which decreases with c – exceeds x_2 but is smaller than x_3, the bank is solvent only in state 3. In that case, the bank chooses the risky project (project 2), no matter the rate of capitalization. This is because the extra payoff for the bank by choosing project 2 exceeds zero for any M:

$$B_1 - B_2 = \theta_3(x_3–M)–(\theta_3+h)(x_3–M) = h(x_3–M) > 0$$

However, for values of M below x_2, the project choice by the bank depends on the capital requirement. In that case, the additional payoff to the bank if it chooses the risky project is:

[20] The expected return for project 1 is $\theta_2 x_2 + \theta_3 x_3$ and that for project 2 is $\theta_2 x_2 + \theta_3 x_3 - h(2x_2 - x_3)$.

$B_2 - B_1 = (\theta_2 - 2h)(x_2 - M) + (\theta_3 + h)(x_3 - M) - \theta_2(x_2 - M) + \theta_3(x_3 - M) = h(M - 2x_2 + x_3)$

Obviously, the bank chooses the risky project if $M > 2x_2 - x_3$. Therefore, depositors should choose c to make sure that $M \leq 2x_2 - x_3$. They will choose the lowest capital requirement that guarantees that this condition is met, hence they choose $c = c^*$ such that $M = 2x_2 - x_3$. As in equilibrium the banks break even, $c^* = B_1 = \theta_2(x_2 - M) + \theta_3(x_3 - M)$. Therefore, $c^* = (\theta_2 + 2\theta_3)(x_3 - x_2)$. Obviously, this minimum will be chosen because higher capitalization reduces the liquidity benefit of deposits.

Having derived this benchmark solution without monitoring, Campbell *et al.* investigate whether the introduction of monitoring can reduce the optimal capital requirement c below the level c^*. This would be an improvement because lower c increases the liquidity service of deposits. If monitoring reduces the minimum capital requirement that makes the bank choose the safe project from c^* to, say, $c° < c^*$, the gain to depositors is $\lambda(c^* - c°)$, where λ represents the liquidity service of deposits. If this gain exceeds the monitoring cost, a combination of capital requirements and monitoring is preferable. Campbell *et al.* analyze both the case of a benevolent and of a self-interested monitor. The monitor's instrument is his effort to receive a signal about the bank's project choice. His cost of making this effort is $V(\beta)$, where β – the monitor's instrument – is the probability of observing a signal about the project choice and can be interpreted as the level of monitoring.[21] As before, the instrument of the depositors is the capital requirement c. The objective of the depositors is to maximize the gains from monitoring (a decrease in c^*) minus the cost of compensating the monitor, which is given by $V(\beta) + \underline{U}$, where $\underline{U}$ is the monitor's fixed cost. The timing is as follows. First, the monitoring effort and the rate of capitalization are chosen. Having observed these, the bank offers the payment M to depositors. Finally, the bank receives its private information about the difference in risk between the projects 1 and 2 and makes a choice between them. If the monitor discovers that the bank chooses the risky project 2, the bank is forced to switch to project 1 and receives a penalty $\phi(c)$. This penalty is assumed to be an increasing function of, but not higher than, the bank capital c. The depositors receive this punishment. The optimal project choice of the bank takes the depositor's choice of the rate of capitalization, the monitor's effort β and the penalty $\phi(\beta)$ into account. The risky project will be chosen if its gain over project 1 times the probability of not being detected outweighs the penalty times the probability of being detected:

$$(1-\beta)\{h[(M(c)-M(c^*)]\} - \beta\phi(c) > 0,$$

[21] $V(0)=0$, $V'(\beta)>0$, $V''(\beta)>0$, $V' \rightarrow 0$ as $\beta \rightarrow 0$, $V' \rightarrow \infty$ as $\beta \rightarrow 1$.

where $M(c^*)=2x_2-x_3$ is the payment corresponding to the optimal level of capital in the absence of monitoring. Thus, whereas in the absence of monitoring the bank chooses the risky project if the degree of capitalization is too low (and the payment he must make to the depositors too high), after introduction of monitoring both the degree of capitalization and the monitoring effort can affect the project choice.

In addition to the liquidity gain through a lower capital requirement, the depositors benefit from monitoring through the receipts from the punishment in case the bank's choice of project 2 is *detected*. However, in addition to the direct monitoring cost $V(\beta) + \underline{U}$, there is an indirect cost because the monitoring solution does not exclude that project 2 is chosen and that this remains *undetected*. An equilibrium solution taking these benefits and costs into account does exist. This solution can be regarded as the optimal combination of the degree of monitoring and the capital structure of banks. As the monitor is benevolent, the choice of the optimal and actual monitoring is made by the depositors and applied by the monitor. If the monitor is self-interested and the monitoring effort is not observable by society/the depositors (and hence not contractable), depositors must combine optimal capital requirements with an incentive contract for the monitor. This incentive contract entails both rewards and a punishment. The monitor's compensation cannot depend on the project's payoff, which is unobservable, hence it should depend on whether or not the bank is solvent. The depositors have three instruments at their disposal, namely the payment to the monitor in case of solvency, the payment in case of insolvency and the degree of capitalization. They choose these variables, taking into account the way the monitor and the bank react to them. Campbell *et al.* assume, that the monitor is constrained in the penalty that can be imposed upon him. They show that only if the monitor's constraint – his ability to pay a penalty – is not binding, an incentive compatible contract resulting in the optimal combination of monitoring and capitalization (namely that from the solution with the benevolent monitor) is possible. If the constraint is binding, that is if the incentive contract must take the monitor's limited resources into account, the best incentive contract implies less monitoring and higher capital requirements.[22] Therefore, a 'richer' monitor is preferable. The implication for regulatory policy is that the choice for a private or public monitor should depend on which of these two has the largest amount of resources. Campbell *et al.* argue that a private regulator is likely to have larger resources so that his ability to pay a punishment is not constrained. One could object to this by saying that a public regulator can always levy taxes to pay the punishment. In that case, however, depositors themselves pay the punishment, something which is not reckoned with in the model.

---

[22] As this problem is too complex to solve analytically, Campbell *et al.* present simulation results for various parameter values.

5.3.5 Bank closure policy and monitoring

The studies reviewed in the previous sections assumed that a bank is owned and managed by one and the same person, and hence models the interaction between depositors on the one hand and ‘the bank’ on the other. A more realistic assumption is that bank managers and stockholders are separate categories, although management rewards in the form of options have become fashionable lately. One strand of the literature addresses the question of what kind of governance structure can provide the bank manager with incentives that increase the quality of loans. This is the well-known principal-agent problem (see section 3.6.3). Others investigate whether closure threats induce banks to behave safely. Finally, some authors approach this problem as one between taxpayers, who bear the burden of monitoring cost and deposit insurance, and the regulator. This section describes these three approaches to closure policy.

Closure policy and manager effort

Dewatripont and Tirole (1994) analyze how closure policy can be used to induce a bank manager to make an effort to improve the quality of loans given by the bank. They conclude that the financial structure of banks (their solvency) matters for optimal closure or reorganization policy. The optimal closure policy under asymmetric information is derived, and it turns out that the allocation of control is contingent on the bank’s performance in the first period. Shareholders should be given the power to close the bank in case of good performance, whereas if performance is poor the depositors – or the regulator as their representative – should decide.

The timing is as follows. At the start, the bank balance is given and consists of deposits and equity, both used to finance loans. Moreover, a control right, contingent on the first-period repayment of the loans, is allocated. The loan quality can be improved, but this is at a fixed cost to the manager. This cost can be interpreted as reflecting his effort. If the bank is allowed to continue, the manager receives a fixed private benefit which can be thought of as his social status or ego. First, the allocation of control rights is set up. Next, the bank manager decides on whether or not to make an effort to improve the quality of the loans. Next, the loans generate their first repayment and a signal is received about their liquidation value in the future. At this moment, the regulator can decide to let the bank continue or to close/reorganize it. It will do so on the basis of the existing information at this time, of course. If the bank is allowed to continue, it is liquidated in the final period and its value is observed. This timing of events is thus:

Financial structure (return streams, control rights) is set up $\rightarrow$ Manager chooses effort (e) $\rightarrow$ (v,u) observed & action (R,C) taken $\rightarrow$ liquidation value realized

Obviously, the manager's effort is influenced by the policy he expects from the regulator. Hence closure policy should be chosen so as to give the manager the incentive to maximize the utility of those represented by the principal (in the case of a regulator: the depositors; in the case of the board of directors: the stockholders). That is, the regulator's strategy should take the incentive-compatibility constraint into account. The focus of the model is on the question how the right to control the bank – the power to decide over closure or continuation of the bank at the end of the first period – should be allocated.

The purpose is to avoid excessive intervention (reorganization when this is not optimal) or too much passiveness (continuation when it would be optimal to reorganize), while at the same time designing a rule affecting the manager's behaviour in the sense that he will make a high effort. The bank balance sheet at the start of the 'game' is given by: $L_1=D_1+E_1$, where L=loans, D=deposits and E=equity. After one period, the loans generate a repayment v related to the manager's effort level e. Moreover, a signal u is received about the future loan repayment (the liquidation value of the bank at the end of the game), η. This signal is also related to the manager's effort. After the first-period repayment v has been revealed, the optimal decision about reorganization or continuation depends only on u, as this is an indication of future profits. The expected net profit (P) from continuing (C) instead of stopping/reorganizing (R) after the first period is defined as: $P(u)=E(\eta|u, C)–E(\eta|u, R)$, where the first right hand side term indicates expected profit in case of continuing and the second in case of reorganizing (or stopping). Obviously, the manager's effort depends on the rule governing the regulator's decision making at t=1. If the distribution of η, given u, is denoted by H, the expected profit will be:

$$P(u) = {}_0\!\int^{\infty} \eta dH_C(\eta|u) - {}_0\!\int^{\infty} \eta dH_R(\eta|u),$$

Under the plausible assumption that P is increasing in u this implies that it is optimal to close or reorganize the bank if u is below a critical value $\hat{u}$, $P(\hat{u})=0$, and to allow it to continue if $u\geq\hat{u}$. Note, that this is the ex post efficient closure rule. Ex ante, the optimal closure rule intends to affect the manager's first-period behaviour in such a way that a high effort level is chosen.

There is asymmetric information about the manager's effort level e, which can take on one of two values, namely low ($\underline{e}$), or high (e^*). Both v and u give information about the effort made by the manager. The best ex ante decision rule under asymmetric information is the one that ensures that the manager's cost of shirking (he will loose his social status) is lower than the (fixed) cost of making an effort, K. The manager's expected cost of shirking (the loss of the benefit he could get from making an effort) is given by

$$B\int\int x(u,v)\{f(u|e^*)g(v|e^*) - f(u|\underline{e})g(v|\underline{e})\}dudv \quad (9)$$

Where $x(u,v)$ denotes the probability of continuation if u, v are observed (the regulator's instrument) and where f, g denote the conditional densities of u and v, respectively. The manager's cost of making an effort is fixed at K. Hence $x(u,v)$, which can be interpreted as the probability that the regulator allows the bank to continue – a function of both u and v, must be chosen in such a way, that it maximizes the regulator's expected profit, under the constraint that the manager's expected cost of shirking is higher than his cost of making an effort (hence, equation (9) $\geq K$). Dewatripont and Tirole show that it is optimal for the controlling authority to choose:

$$\begin{array}{lll} x(u,v)=1 \text{ (continuation)} & \text{if } P(u)+\mu B \geq \mu B(F(\underline{e})/F(e) & (10) \\ x(u,v)=0 \text{ (bank reorganization or closure)} & \text{otherwise,} & \end{array}$$

where $F(.)=f(u|.)g(v|.)$, and where μ is the Lagrange-multiplier associated with the incentive compatibility constraint. The strategy given by (10) implies that the control strategy may imply ex post inefficient passiveness (bank is allowed to continue although under complete contracting it would be reorganized) or ex post inefficient intervention (bank is reorganized although it would have been efficient not to interfere). This is because the control strategy is designed to influence the manager's behaviour at t=0. At t=1, the manager has chosen whether or not to make an effort, and the optimal strategy depends only on u. This is illustrated in Figure 3, where $\hat{u}$ is the ex post efficient rule and $u^*(v)$ gives the efficient ex ante closure rule; only $u^*(\hat{v})$ is both ex ante and ex post efficient. In fact, there is a time-inconsistency problem. Dewatripont and Tirole suggest to overcome this problem by an appropriate allocation of control. Although it may be impossible to draw up a contract implying the rule $u^*(v)$, or even to verify u, the analysis still gives scope for recommendations regarding this allocation. Stockholders tend to favour risky projects and depositors tend to be on the safe side. Therefore, the analysis suggests that the right to control the bank should be allocated according to a rule that gives power ex post to stockholders in case of good performance in the first period (v above a threshold $\hat{v}$), and to the regulator (representing the depositors) if first-period performance is bad (v below $\hat{v}$). The optimal policy within each region can be implemented by net worth adjustments, composite claims or voluntary recapitalization.

If there are only the two standard types of claimholders, shareholders versus depositors, with the latter represented by the regulatory authority, optimal decisions require that net debt – that is, deposits minus v – must increase with v. This is because a higher net debt (D–v in the model) induces both depositors and shareholders to take more risk, that is they become more passive, and that is exactly what the optimal $u^*(v)$ rule indicates. In other words, for optimal closure policy the bank should be allowed to hold more deposits, the higher its verifiable performance. As Dewatripont and Tirole (1994) point out, the mere evolution of the bank's balance sheet gives rise to 'perverse incentives': given D, lower values of v (worse performance) give rise to more passivity. When D–v is very large, and equity holders are in control, they have an incentive to 'gamble for resurrection', choosing to continue although stopping is far superior.

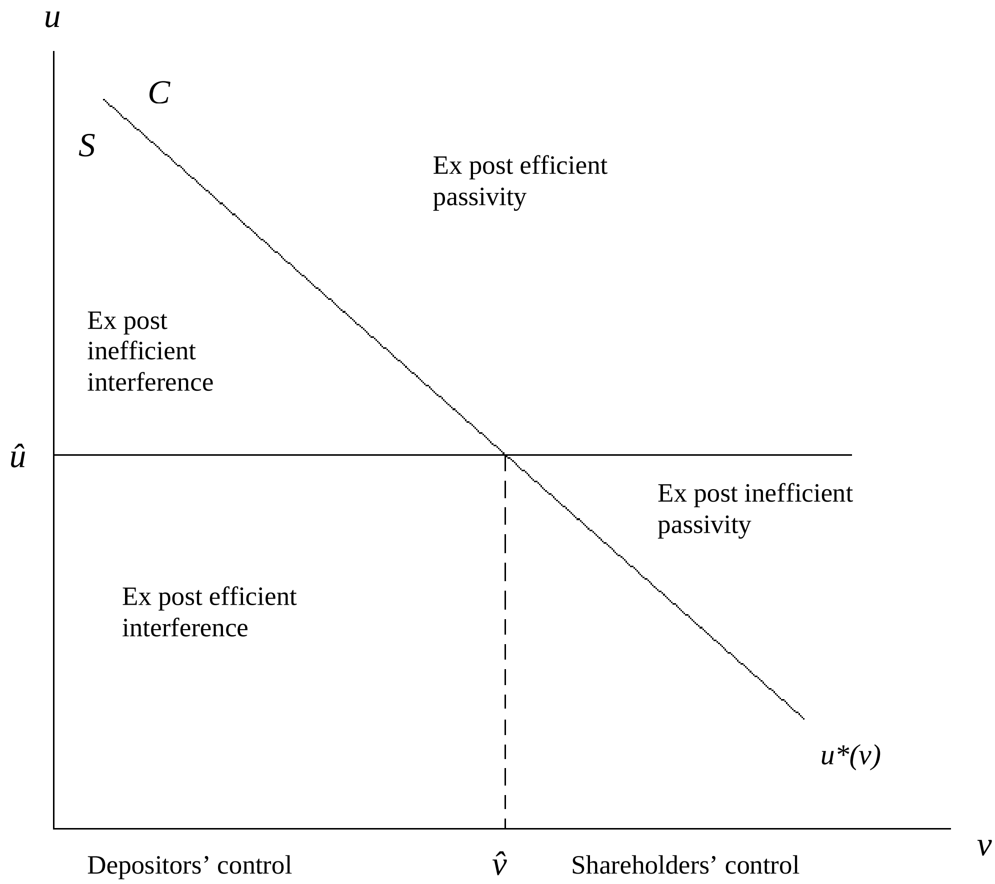

Figure 3 The optimal managerial incentive scheme
Source: Dewatripont, M. and J. Tirole (1997), Figure 7.1, p. 140

Composite claims are a combination of shares and debt. They can guarantee a continuous *decrease* in interference as first-period performance *increases,* according to $u^*(v)$. Assume a range of composite claims made up of $\beta(v)$ units of debt (deposits) and $1-\beta(v)$ shares, where $\beta(.)$ is a decreasing function. For every performance v there is then a composite-claimholder that should be given the power over closure or continuation, as, with that value of v, his decision implies $u^*(v)$. Instead of allocating power over closure to share- or debtholders depending on first period-performance v, the regulator could choose a policy that does always give the power to the shareholders. The condition is, that for low values of v – when otherwise control would go to depositors – shareholders inject new capital into the bank, so that they have more to loose.

Dewatripont and Tirole show, that this may lead to the implementation, by shareholders, of exactly the optimal closure policy $u^*(v)$.

The analysis by Dewatripont and Tirole shows, that financial structure matters. However, their proposals seem to be quite difficult to apply in practice. Rather than focusing on the ownership structure, it might be useful to pay attention to the necessity, for a regulator, to pay attention to both solvability and rentability of banks.

Are closure threats credible?

Mailath and Mester (1994) analyze credible closure threats under perfect information. The question is, how a regulator can use their power over closure to induce banks to behave in such a way that too excessive risk-taking is prevented, in which case closure is not necessary. They use a two-period model, in which at the end of the first period the regulator can decide whether to close the bank, in which case the regulator will incur a fixed cost C. At the start (t=1), the bank receives a deposit and must decide whether to invest it in a risky or a safe asset. At t=2, the regulatory body can decide whether or not to close the bank, and if the bank is not closed it must again choose between a risky or a safe investment. At t=3 both assets mature. Hence the return for the first asset can be realized only if the bank does not fail in the second period, because it is assumed that the return to a single asset is not enough to cover the loss on the other asset. Projects are bank specific, hence if a given bank does not finance a project it will not be undertaken at all. For simplicity the interest rate for depositors is set to zero, and there is no initial capital. Hence the expected profit for the bank is the return on the asset net of the principal. The net return on the safe asset is r (r>0) with probability 1. The net risky rate of return is ρ with probability p and zero with probability $1–p$. It is furthermore assumed that $\rho>r$, hence if the risky investment succeeds it has a higher return than the safe one. However, $p(\rho+1)<r+1$, hence the return on the safe asset is higher than the expected return on the risky one. These two assumptions create a tension between the incentives of society (the regulator) and those of the bank. It is assumed that if the bank could choose only once, it would choose the risky asset, this as a result of its limited liability and the absence of risk-rated deposit insurance premiums (the moral hazard problem discussed in section 5.3.2). The regulator may either be welfare-maximizing, i.e., acting in the interests of society, or cost-minimizing. In the first case, expected return minus the cost of closing the bank is maximized, whereas in the second case the regulator minimizes its costs (namely the returns to depositors and the cost of closing the bank). Although Mailath and Mester do not mention this, one could argue that the first type of regulator is necessarily a public institution under control of parliament, whereas the second one might be a separate body, possibly in the private sector. If banks prefer the safe project anyway, there would be

no reason to regulate. Therefore, we focus here on those cases where, absent regulation, banks would choose at least once the risky investment. This can be shown to be the case if the probability of success is not too low: $p>2r/(r+\rho)$. We analyze here the case where the regulator is of the cost-minimizing type.

At the start of the 'game', the expected return to the bank is $p(\rho+r)$ if it chooses once the risky and once the safe asset (S,R or R,S). The expected return is $p^2(2\rho)$ in case the risky strategy is chosen twice. Let us first turn to the situation in which the expected profits to the bank are highest if it chooses once the safe and once the risky investment. Hence it strictly prefers (S,R) or equivalently (R,S) to (R,R).[23] This is the case if the safe return r is high enough in relation to the risky return ρ and the probability of success p: $r>(2p-1)\rho$. Call this case 1. The regulator would like to induce the bank to choose the safe strategy twice. If he observes that the bank has chosen S in the first period, he knows that the second period investment will be the risky one R, hence if the closure cost is not too high he will close the bank at $t=1$. However, the bank knows in advance that by choosing S in period 1 it will be punished with closure and therefore the closure threat will have no effect but to induce him to choose R, S instead of S,R. Thus, there is no improvement by the closure threat. Let us now turn to the case where $r\leq(2p-1)\rho$, so the bank prefers R,R to R,S. Call this case 2. In this case, whatever strategy is observed in period 1, the regulator takes a second-period choice of R into account. If the cost of closure is not too high, the bank will be closed no matter what strategy is has chosen in the first period. If the closure cost is too high, the bank will remain open despite its risky investment choice. In case 1 (low probability of success, hence S,R preferred to R,R), the regulator will never close the bank, because the closure threat itself induces the bank to choose the risky investment first, and as the second-period strategy will certainly be the safe one it is then optimal to leave the bank open.

The analysis implies, that there is a range of parameters in which the closure policy is credible, but that for other values the regulator is 'forbearing' and regulation is useless. This result also applies to the case in which the regulator maximizes social welfare. More importantly, if it is assumed that the first-period risky asset is riskier than the second-period ($p_1< p_2$), the existence of a regulator may reduce social welfare. This is because the equilibrium implies that the bank will not be closed if it chooses S in the second period. Hence, for certain parameter values, regulation induces the bank to choose R,S instead of S,R. The analysis rests on a number of crucial and sometimes implausible assumptions. Thus, one- and two-period assets are allowed

---

[23] These are equivalent because both assets pay off only if the bank survives and because the first and second period risky assets have the same characterictics (probability of success, return). The latter is a simplification as compared to Mailath and Mester's original model, as they assume that $p_1 \neq p_2$.

for, where the first-period assets only pay off at the end of the second-period. Moreover, the assumption that all projects are bank-specific is far from realistic.

Rules versus discretion in regulatory policy

Boot and Thakor (1993) analyze the issue of bank regulation as a principal-agent problem with the tax-payer as principal and the regulator as agent, rather than as one between the regulator as principal and the bank as agent. The upshot is, that taxpayers bear the burden of deposit insurance, and that it is therefore in their interest to make sure that the insurance funds' liabilities are limited. The analysis of Boot and Thakor leads to some recommendations for the institutional design of regulatory policy. Therefore, their study is relevant, for example because the changeover to Economic and Monetary Union and the creation of the European Central Bank have already led to discussions about how future regulatory policy should be designed (Lannoo, 1998). Boot and Thakor assume that both banks and 'the market' have an information disadvantage about the regulator's ability to monitor the asset choice of banks. The regulator, on the other hand, has incomplete information related to the bank's balance sheet. To this we shall turn later. The regulator wishes to establish the reputation of being a capable monitor; this is his self-interest. The pursuit of this goal affects his bank closure policy in the sense that banks are sometimes allowed to continue where it would be optimal to close them. This increases the deposit insurance funds' liability, at the expense of the taxpayers.

The analysis uses a two-period model. At the beginning of the first period, the bank has two assets, one for which the pay off distribution is given, another for which it can choose the payoff distribution. For the latter an investment of 1 must be made at t=0, which is financed with equity capital K_1 and with insured deposits $(1–K_1)$. This investment has an expected payoff at t=1 of $E(R_1)=\theta_1 R(\theta_1)+(1-\theta)\cdot 0$, where θ is the possibility of receiving a positive payoff and can be chosen by the bank.[24] The socially optimal first-period choice, that is the choice that an all-equity financed bank would make, $\theta_1{*}$, maximizes the expected payoff. The regulator monitors the bank's choice of θ_1 imperfectly at t=0. If he detects that the bank does not choose $\theta_1{*}$, he will force the bank to switch to $\theta_1{*}$. The regulator is one of two types. If he is of good quality he detects the bank's choice with probability ρ_g, if he is bad he does so with probability ρ_b, where $\frac{1}{2}\leq\rho_b<\rho_g\leq 1$. There is a known prior probability of γ ($\gamma\in(0,1)$) that the regulator is good. At the end of the first period (t=1), the bank realizes its payoff $y'+R_1$, where y' is the payoff of the asset which payoff distribution was given. The depositors get their money back and the regulator de-

[24] $R(\theta_1)>0$, $R'(\theta_1)<0$, $R''(\theta_1)\leq 0$.

cides whether or not to close the bank. If there is not enough to pay off the depositors but the bank is allowed to continue, second period deposits are also raised to pay off first-period depositors. The regulator's information disadvantage is that he observes the sum $y'+R_1$, but not y' and R_1 separately. He has the power to close the bank at the end of the first period. If the bank is not closed, it invests the difference between its payoff and the repayment to the depositors and raises second-period deposits in order to invest an amount of 1. If the difference is negative, second-period deposits are raised to a level sufficient to repay the depositors and to invest 1. As in the first-period, the bank can choose the payoff distribution for the second period. If the return is insufficient to pay off the depositors, the deposit insurance company pays the rest. The regulator is both self-interested and socially concerned. His objective is to use his power to close the bank in order to maximize a weighted sum of his reputation and of the social surplus from the second-period asset:

$$\lambda_1[\gamma_1{}^m+\delta\gamma_2{}^m]+\lambda_2[\theta_2 R(\theta_2)-1] \tag{11}$$

where γ_i (i=1,2) is the posterior belief that the regulator is of good quality. The bank's first-period behaviour can be affected by the closure policy rule. It must be kept in mind that an insolvent bank may be allowed to remain open and that a solvent bank may be closed (see below).

As a benchmark, the bank's socially optimal investment for the second-period can be derived. It maximizes the difference between the expected second-period payoff and the risk-free rate. As the latter is given, this implies maximization of $\theta_2 R(\theta_2)$, hence $\theta_2{}^*= -R(\theta_2{}^*)/R'(\theta_2{}^*)$. The bank would choose this strategy if its second-period investment were fully financed with equity capital (in that case, K_2 would equal 1). At the end of the first period, depositors receive their initial deposit plus the risk-free rate. Hence second-period capital is $K_2=y+R_1-(1-K_1)r_f$, where r_f is one plus the risk-free rate. By assumption, $K_2<1$, so that additional deposits are needed to make sure that the second-period asset equals 1. Introduction of deposits is the rationale for regulation. Given that the bank's assets are financed with both deposits and equity, the bank chooses θ_2 to maximize its rent M, which is defined as $M(\theta_2)=\theta_2\{R(\theta_2)-[1-K_2]r_f\}-K_2r_f$, where the term between brackets gives the payoff of the investment in the second period asset minus the repayment to depositors, and the second term gives the 'risk-free' payoff for the bank's shareholders. Boot and Thakor prove that θ_2 – the probability of success of the second-period investment chosen by the bank – increases with K_2. Hence it is lower than in the social optimum, because $K_2<1$. In other words, the bank takes too much risk in the second period. This does not imply, however, that it is always optimal to close the bank at the end of the first period. Rather, the socially optimal closure rule can be thought of as preventing a negative net present value asset portfolio in the second period. Boot and Thakor show that this rule depends on the bank's second-period capital K_2 : if this is lower than some critical value K_2, the bank should be closed, whereas it is optimal to let the bank continue if $K_2 \geq K°$. They also show, that this critical value is positive – in other words, that it may be optimal to close banks with positive second-period capital – as long as the riskless interest rate (the deposit funding cost) is not too high. This is because even with positive second-period capital the deposit insurance may result in a second-period portfolio choice by the bank with a negative net present value.

The bank's second-period decision is about choosing θ_2, given K_2 and we have seen that the higher K_2, the higher θ_2, hence the higher the possibility of success or, equivalently, the lower the risk taken by the bank in the second period. However, K_2 itself is of course affected by the first-period decisions of the bank. Hence it is important to investigate how the second-period rent M, given $\theta_2(K_2)$, varies with K_2. Boot and Thakor prove, that conditional on the bank being allowed to continue, its stockholders are better of with a lower K_2. The intuition behind this result is, that with lower K_2 the bank profits more from the deposit insurance, which – as we mentioned earlier – can, according to Miller (1977) be regarded as a put option. The implication is, that the first-period risk chosen by the bank is higher than it would be if no second period would ensue. The only way for the regulator to counterbalance this behaviour is by designing a rule that punishes a bank with closure if its level of K_2 is 'too low'.

The regulator decides whether or not to close the bank at the end of the first period, when he observes $y+R_1$. Assume that the bank is closed if $y+R_1$ is below some critical value, say z^*. As a benchmark, the socially optimal closure rule can be derived. This rule entails closure for values of K_2 lower than a critical value $K°_2$. As $K_2=y+R_1–(1–K_1)r_f$, the optimal closure rule implies $z^*=K°_2+(1–K_1)r_f$. However, by assumption the regulator does not choose the socially optimal closure rule, but weighs social utility against his reputation of being of good quality according to equation (12). Boot and Thakor choose the exogenous parameters such, that the bank is not closed if $R_1(\theta_1)>0$.[25] Whereas the regulator observes the bank's book capital at the start of each period, it is assumed that the market does so only with a lag. Hence closure of the bank informs the public (the market) that $R_1=0$. This is a crucial assumption for the result obtained, as will be made clear below. Boot and Thakor show, that under this assumption, maximization of (11) implies a critical value $z^*<K°_2+(1–K_1)r_f$. This is lower than the socially optimal one: the regulator is 'too lax'. This is because closure implies an insufficient bank capital, which makes it more likely that the bank's first period asset choice was too risky ($\theta_1<\theta_1^*$), which then has gone

[25] This is not a deliberate choice by the policymaker, as by assumption he does not observe R_1 separately.

undetected. Hence the market reduces its probability that the regulator is a good monitor. Boot and Thakor derive the following five policy implications from this result. First, responsibility for closure policy should be separated from asset-quality monitoring. Second, long-term appointment of the regulatory authority tends to acerbate the problem, as in this case, reputation-building is not in the interest of society. Third, regulatory discretion should be reduced in favour of rigid rules, for example by introducing minimal capital requirements, although the observability and measurement is in practice a problem here. Fourth, asset portfolio restrictions limit the range of possible risks chosen by the bank, thereby improving the regulator's monitoring possibilities. This makes closure policy less relevant as a signal of the regulator's ability and therefore closure policy will be less distorted. Finally, it follows that the public perception of the regulator's quality is extremely important, although it should be stressed that even a little uncertainty about his ability may lead to distortions. The analysis has several drawbacks. There is a different treatment of the first and second period. Thus, for example, it is assumed that θ_2 is not monitored by the regulator. The result depends crucially on the arbitrary choice of exogenous parameters. Also, introduction of interim dividend would probably crucially affect the results, given the importance of second-period capital K_2 for the results.

6 THE PRACTICE OF BANKING REGULATION

The previous section paid attention to the theory of banking regulation and discussed pros and cons of various regulatory instruments as well as the appropriate institutional design of regulatory policy. Some of the theoretical contributions have already had their impact on actual regulatory policy. In this section, some key features of actual bank regulation will be reviewed in the light of the theoretical contributions described here.

6.1 Deposit insurance

Most industrialized countries do have explicit deposit insurance schemes (de Lange, 1992). In the United States deposit insurance was introduced in 1934 in reaction to the massive bank failures during the Great Depression. Federal Reserve Member banks were obliged to insure their depositors with the Federal Deposit Insurance Corporation (FDIC), other banks could participate in this fund on a voluntary basis. At present, the deposit insurance covers losses due to bank failure up to $100,000. In the Netherlands, deposit insurance was introduced in 1965. In many other EU countries, depositor protection was introduced in the seventies. With the EU's deposit guarantee schemes directive it became obligatory in 1992. According to this directive, deposits must be protected up to a minimum level of 20.000 ECU. The insurance schemes in the US and the European Union are publicly run. According to Mishkin (1992), the introduction of the deposit insurance system in the United States sharply reduced the number of bank failures between 1934 and 1981. However, from 1982 the number of bank failures increased with the result that the FDIC was unable to pay off all depositors and therefore taxes had to be levied.[26] Moral hazard, i.e., overly risky behaviour by banks as described in section 3.3.2, is used as an explanation for this phenomenon. However, the deposit insurance scheme led to bank failures only fifty years after its introduction. The increase in bank failures may have been a result of the financial deregulation of the 1980 Depository Institutions Deregulation and Monetary Control Act, that gave more freedom of activities to thrift institutions. An additional feature influencing the failure of banks, possibly partly through incentives for moral hazard, was the monetary tightening applied by the Reagan Administration from 1979 until 1981, and resulting in sharp interest rate increases and a severe recession. As a reaction to the increase in bank failures the Federal Deposit Insurance Improvement Act (FDICIA) of 1991 took some measures to safeguard the banking system. Thus, in line with the recommendations of the theory, it instructed the FDIC to establish risk-based deposit insurance premiums

6.2 Capital requirements

Capital requirements are aimed at safeguarding banks' solvency. Indirectly, they serve to maintain confidence in the banking system and may therefore reduce the danger of bank runs and systemic risk. With minimum capital requirements, there is more at stake for the bank itself, and should the bank fail then there is more to distribute among claimholders. Therefore, the minimum capital can be viewed as a form of collateral. As we have seen in section 3.3.3, however, minimum capital requirements may have perverse effects on the behaviour of banks that have a low rate of risk aversion. Therefore, some authors argue that the requirements should be risk-related. They have shown, that optimal weights for risk-categories can be derived. However, others, notably Kaufman (1996) maintain, that there is insufficient knowledge to be able to actually apply the optimal weights. Still, in practice, regulating bodies have already acknowledged the fact that risk-related capital requirements are preferable. For example, capital requirements as agreed upon by the Basle Committee differ according to the degree of asset risk. Assets and off-balance sheet items are allocated into four categories depending on their risk as perceived by the regulator. The weighted items are added up to get the sum of risk-adjusted assets. Against these assets the bank must hold 4% of 'core capital' (Tier 1, stockholder equity), while total capital (core capital plus loan loss reserves and the like, i.e., Tier 1 + Tier 2 capital) must amount

[26] 1142 savings and loan associations and 1395 banks went bankrupt (Lannoo, 1998).

to 8% of the risk-adjusted assets.[27] In reaction to the recent financial crisis in Asia proposals have been made to increase the latter percentage, for example to 10%. In the light of the theory, however, one may wonder whether applying an undifferentiated increase is appropriate.

6.3 Monitoring and closure policy

According to the theory, capital requirements should be supplemented with monitoring because this reduces the degree of necessary capitalization for a given reduction of the probability of insolvency, thus limiting the loss in liquidity service. As depositors constitute a large group of small, uninformed and unorganized individuals, the monitoring task is better performed by a regulator. This is even more so in the presence of deposit insurance, which removes all monitoring incentives from depositors. In practice, monitoring is indeed applied. It takes the form of periodical bank examinations by the regulator, to monitor banks' compliance with capital requirements and asset restrictions. The monitoring enables the regulator to take the 'prompt corrective action', i.e., intervene if the bank is troubled. The objective is to prevent a further deterioration and the need for bank closure. In the US, provisions for such action are laid down in the FDICIA from 1991. Banks are classified according to their bank capital. This should eliminate the regulatory forbearance – resulting from the principal-agent problem – predicted by the theory (e.g., Boot and Thakor, 1993). In New Zealand the conventional monitoring practice has been replace, in 1996, by a system of disclosure requirements for the bank. The central bank monitors the bank's obligatory quarterly financial statement, that also must be made public. More importantly, bank directors face unlimited liability if they make false or misleading statements. This should make monitoring easier for depositors, and as there is no deposit insurance system the market should provide sufficient monitoring and discipline. Notice, however, that most banks in New Zealand are foreign and under control of the home country. The theory predicts, that closure policy is too lax. According to Mishkin (1997), regulatory forbearance is indeed observed in practice.

6.4 Institutional design of regulation

According to Goodhart and Schoenmaker (1995), the appropriate institutional design differs according to the banking structure. Table 1 shows, that the design of banking regulation indeed differs across countries. Lannoo (1998) observes, that this reflects the fact that there is no

---

[27] The weight is zero for assets without default risk, such as cash and government securities. The next risk category has a weight of 20% and includes low-default risk assets like interbank deposits. The third category has a weight of 50% and includes for example residential mortgages. Finally, there is a category with a weight of 100% including all remaining securities and loans.

theoretical consensus about the optimal institutional framework of regulatory policy. of a country. Nevertheless, there is a trend toward retreatment, by the central bank, from supervision. An example is the situation in the UK, where recently – partly in reaction to the failure of Barings – the Financial Services Authority has been established, a mega-financial supervisory authority. This is not so much a trend toward private supervision, but reflects according to Lannoo the growing acceptance that the government instead of the central bank should take responsibility for financial support. One reason is that bank rescues have become more expensive and central banks had insufficient resources to pay off depositors, as recent experiences in France, Sweden and Norway have shown. If the taxpayer pays the cost of rescuing banks, than regulation should be under parliamentary control. However, in the UK, the Bank of England remains responsible for monitoring. This separation of monitoring and other regulatory functions is in line with the recommendations by Boot and Thakor (1993) mentioned above.

7 CONCLUSIONS

In the previous subsections, the existence and desirability of economic regulation in general has been discussed, and a number of theoretical models of regulatory policy in the financial sector have been presented. In general, government regulation of firms or industries can be viewed either as a policy aimed against market failures or as the result of pressure from interest groups that try to capture rents from regulation. The first approach relies on the welfare-theoretic argument that market failures limit the possibilities to reach a Pareto efficient allocation of goods. Financial markets, and banks in particular, are especially vulnerable to market failures, like externalities and asymmetric information. The models on financial regulation all focus on banks. The theory of financial regulation has concentrated on the banking sector because the justification of regulation is based on the fact that deposits serve as a means of payment and can be withdrawn at any time, and that bank loans are not marketable. From the theoretical contributions reviewed here a number of policy conclusions can be drawn. Deposit insurance is important to maintain confidence in the banking system and to prevent inefficient bank runs. However, it creates an incentive for risky behaviour on the part of banks as it discourages monitoring on the part of depositors. Therefore, deposit insurance premiums should be related to risk. The portfolio approach to capital requirements has shown that capital requirements do not always result in reduced risk, and may even result in more risk-taking, and therefore, like deposit insurance premiums, capital ratios should be risk-related. The incentive approach stresses that monitoring is to be preferred over capital requirements, but that the optimal regulatory policy is a combination of both. A benevolent monitor is desirable; if the monitor is self-interested – which is perhaps more likely in the case of a private regulatory body – he should have more to loose if he does not behave in accordance with the depositors' interest. The incomplete contract approach to bank closure policy pays attention to the fact that stockholders and managers are different groups and may have competing interests. The conclusion is, that an appropriate ownership structure of the bank, instead of a regulatory agency, may induce desirable bank behaviour. Composite claims, consisting of debt *cum* shares, may provide the appropriate closure threat. If this is not feasible, the right to make decisions about whether or not to close or reorganize a bank should be made dependent on the actual performance of the bank. In this case, the ownership structure of the bank should have to be such, that the bank's net debt is lower, the worse its performance. Another conclusion from the theory on bank closure policy is that closure threats may not always be credible. This may explain why in practice regulators are criticized for being too lax in their closure decisions. Also, the existence of a regulator may, according to this approach, have perverse effects on the behaviour of banks. Some authors argue that decisions on bank closures should be institutionally separated from the task of monitoring, and that closure rules are to be preferred over discretionary closure policy.

Some of the theoretical recommendations have been (partly) introduced in the actual policy of bank regulation. Thus, deposit insurance premiums are sometimes risk-related, just as are capital requirements. Regulators combine monitoring and capital requirements. And in some cases, monitoring and closure policy are entrusted to separate institutions. However, as recent banking crises in Asia and Russia, and failure of the Long Term Capital Management hedge fund in the United States have shown, the banking sector throughout the world remains fragile. Moreover, as theory predicts, regulators are sometimes too forbearing or do not always take 'prompt corrective action'. On the other hand, as Goodhart *et al.* (1998, p. 54) point out, 'In contrast to the highly publicised major financial failures … cases where regulators have used their discretion … more successfully remain largely unreported'. Therefore, in their view a final answer to the question of whether regulators are or are not too forbearing can probably neven be given. New financial instruments create new regulatory problems. Therefore, it may be true, as Mishkin (1997) observes, that designing and applying financial regulation policy is like 'shooting on a moving target'.

REFERENCES

Atkinson, A.B., and J.E. Stiglitz, 1980, *Lectures on Public Economics*, McGraw-Hill, Maidenhead.

Akerlof, G.A., 1970, 'The Market for "Lemons": Qualitative Uncertainty and the Market Mechanism,' *Quarterly Journal of Economics*, 84, pp. 488-500.

Bagehot, W., 1873, *Lombard Street: A Description of the Money Market*, H.S.King, London.

Bank for International Settlements, 1994, *Annual Report 1993-1994*, BIS, Basle.

Becker, G.S., 1983, 'A Theory of Competition among Pressure Groups for Political Influence,' *Quarterly Journal of Economics*, 98, pp. 371-400.

Benston, G.J. and G.G. Kaufman, 1996, 'The Appropriate Role of Bank Regulation,' *Economic Journal*, 106, pp. 688-697.

Bernanke, B.S., 1983, 'Nonmonetary Effects of the Financial Collapse in the Propagation of the Great Depression,' *American Economic Review*, 73, pp. 257-276.

Boot, A. and A. Thakor, 1993, 'Self-interested Bank Regulation,' *American Economic Review*, 83, pp. 206-212.

Campbell, T.S., Y-S. Chan and A.M. Marino, 1992, 'An Incentive-based Theory of Bank Regulation,' *Journal of Financial Intermediation*, 2, pp. 255-276.

Chan, Y.S., S.I. Greenbaum and A.V. Thakor, 1992, 'Is fairly priced deposit insurance possible?' *Journal of Finance* 47, pp. 227-245.

Chang, H.-J., 1997, 'The Economics and Politics of Regulation,' *Cambridge Journal of Economics*, 21, pp. 703-728.

Coase, R., 1960, 'The Problem of Social Cost,' *Journal of Law and Economics*, 3, pp. 1-44.

Crouhy, M. and D. Galai, 1986, 'An Economic Assessment of Capital Requirements in the Banking Industry,' *Journal of Banking and Finance*, 10, pp. 231-241.

Dewatripont, M. and J. Tirole, 1994, *The Prudential Regulation of Banks*, MIT Press, Cambridge (Mass.).

Diamond, D. and P. Dybvig, 1986, 'Banking Theory, Deposit Insurance and Bank Regulation,' *Journal of Business*, 59, pp. 53-68.

Fase, M.M.G., 1995, 'Zoeken naar Monetaire Stabiliteit,' in: *Liber Amicorum voor Arnold Heertje*, pp. 39-44.

Freixas, X. and J.-Ch. Rochet, 1997, *Microeconomics of Banking*, MIT Press, Cambridge (Mass.).

Giammarino, R.M., T.R. Lewis and D. Sappington, 1993, 'An Incentive Approach to Banking Regulation,' *The Journal of Finance*, 48, pp. 1523-1542.

Goodhart, Ch., Ph. Hartmann, D. Llewellyn, L. Rojas-Suàres and S. Weisbrod, 1998, *Financial Regulation*, Routledge, London/New York.

Grossman, S. and O. Hart, 1983, 'An Analysis of the Principal Agent Problem,' *Econometrica*, 52, pp. 1-45.

Hägg, P.G.T., 1997, 'Theories on the Economics of Regulation: A Survey of the Literature from a European Perspective,' *European Journal of Law and Economics*, 4, pp. 337-370.

Hahn, R.W. and J.A. Hird, 1990, 'The Costs and Benefits of Regulation: Review and Synthesis,' *Yale Journal on Regulation*, Vol. 8, pag. 223-278.

Hirshleifer, J. and J.G. Riley, 1992, *The Analytics of Uncertainty and Information*, Cambridge University Press, Cambridge.

Holmström, B., 1979, 'Moral Hazard and Observability,' *Bell Journal of Economics*, 10, pp. 74-91.

Joskow, P.L. and R.G. Noll, 1981, 'Regulation in Theory and Practice: An Overview,' in: *Studies in Public Regulation*, G. Fromm (ed.), MIT Press, Cambridge (Mass.), pp. 1-65.

Kahane, Y., 1977, 'Capital Adequacy and the Regulation of Financial Intermediaries,' *Journal of Banking and Finance*, 18, pp. 307-323.

Kaufman, G.G., 1996, 'Bank Failures, Systemic Risk and Bank Regulation,' *Cato Journal*, 16.

Kim, D. and A.M. Santomero, 1988, 'Risk in Banking and Capital Regulation,' *The Journal of Finance*, 43, pp. 1219-1233.

Kroes, T., 1996, *Financial Intermediation and Monetary Transmission*, Tinbergen Institute Research Series, 120, Thesis Publishers, Amsterdam.

Laffont, J.J., 1989, *The Economics of Uncertainty and Information*, MIT Press, Cambridge (Mass.).

Laffont, J.J., 1994, 'The New Economics of Regulation Ten Years after,' *Econometrica*, 62, pp. 507-537.

Laffont, J.J. and J. Tirole, 1993, *A Theory of Incentives in Procurement and Regulation*, MIT Press, Cambridge (Mass.).

Lange, M. de, 1992, *Essays on the Theory of Financial Intermediation*, Thesis Publishers, Amsterdam.

Lannoo, K., 1998, 'EMU, the ECB and Prudential Supervision,' paper prepared for the 1998 SUERF conference, Frankfurt, October 14-17, 1998.

Leland, Hayne E. and David H. Pyle, 1977, 'Informational Asymmetries, Financial Structure, and Financial Intermediation,' *The Journal of Finance*, 32, pp. 371-387.

Levine, I.R., 1997, 'Financial Development and Economic Growth: Views and Agenda,' *Journal of Economic Literature*, June, pp. 688-726.

Lippman, S.A. and J.J. McCall, 1981, 'The Economics of Uncertainty: Selected Topics in Probabilistic Methods,' in: *Handbook of Mathematical Economics*, K.J. Arrow and M.D. Intriligator (eds.), North-Holland, Amsterdam, pp. 211-284.

Mailath, G. and L. Mester, 1994, 'A Positive Analysis of Bank Closure,' *Journal of Financial Intermediation*, 3, pp. 272-299.

Miron, J.A., 1986, 'Financial Panics, the Seasonality of the Nominal Interest Rate, and the Founding of the Fed,' *American Economic Review*, 76 (1), pp. 125-140.

Mishkin, F. S., 1992, *The Economics of Money, Banking and Financial Markets*, Fifth edition, Addison Wesley, New York.

Mueller, D.C., 1989, *Public Choice II*, Cambridge University Press, Cambridge.

OECD, 1993, *Systemic Risks in Securities Markets*, OECD, Paris.

Olson, M., 1965, *The Logic of Collective Action*, Harvard University Press, Cambridge (Mass.).

Peltzman, S., 1976, 'Toward a More General Theory of Regulation,' *The Journal of Law and Economics*, Vol. 19, Nr. 2, pp. 211-240.

Peltzman, S., 1989, 'The Economic Theory of Regulation after a Decade of Deregulation,' *Brookings Papers on Economic Activity, Microeconomics*, pp. 1-41.

Pigou, A.C., 1920 (revised 1932), *Economics of Welfare*, Macmillan, London.

Posner, R.A, 1971, 'Taxation by Regulation,' *Bell Journal of Economics and Management Science*, 2, Spring, pp. 22-50.

Posner, R.A, 1974, 'Theories of Economic Regulation,' *Bell Journal of Economics and Management Science*, 5, Autumn, pp. 335-358.

Rochet, J.C., 1992, 'Towards a Theory of Optimal Banking Regulation,' *Cahiers Economiques et Monétaires de la Banque de France*, 40, pp. 275-284.

Samuelson, P.A., 1954, 'The Pure Theory of Public Expenditure,' *Review of Economics and Statistics*, 36, pp. 386-389.

Schwert, G.W., 1977, 'Public Regulation of National Securities Exchanges: A Test of the Capture Hypothesis,' *Bell Journal of Economics and Management Science*, 8, pp. 128-150.

Stigler, G.J., 1971, 'The Theory of Economic Regulation,' *Bell Journal of Economics and Management Science*, 2, Spring, pp. 3-21.

Stigler, G.J. and C. Friedland, 1962, 'What Can Regulators Regulate? The Case of Electricity,' *Journal of Law and Economics*, 5, pp. 1-16.

Stiglitz, J.E., 1988, 'Money, Credit and Business Fluctuations,' Sir John Crawford Lecture delivered to the 1988 Australian Economics Congress, Canberra, 31-08-1988.

Stiglitz, J.E. and A. Weiss, 1981, 'Credit Rationing in Markets with Imperfect Information,' *American Economic Review*, 71, pp. 393-410.

Swann, D., 1989, 'The Regulatory Scene: An Overview,' in: K. Button and D. Swann, 'The Age of Regulatory Reform,' Clarendon Press, Oxford, pp.1-23.

Takayama, A., 1994, *Analytical Methods in Economics*, Harvester Wheatsheaf, New York.

Tirole, J., 1988, *The Theory of Industrial Organization*, MIT Press, Cambridge (Mass.).
Varian, H.R., 1984, *Microeconomic Analysis*, 2nd edition, Norton, New York/London.
Viscusi, W.K., 1996, 'Economic Foundations of the Current Regulatory Reform Efforts,' *Journal of Economic Perspectives*, 10, pp. 119-134.
Winston, C. and R.W. Crandall, 1994, 'Explaining Regulatory Policy,' *Brookings Papers on Economic Activity*, *Microeconomics*, pp. 1-49.
Wit, J.N.M., 1998, 'Bank Runs in het Diamond-Dybvig Model (Bank Runs in the Diamond-Dybvig Model),' *Onderzoeksrapport WO&E*, No. 561, De Nederlandsche Bank, Amsterdam.