Review of Economic Analysis 10 (2018) 371-407 1973-3909/2010371

# Liquidity Provision, Ambiguous Asset Returns and the Financial Crisis

**WILLEM SPANJERS***

*Department of Economics*
*School of Law, Social and Behavioral Sciences*
*Kingston University*
and
*Rimini Center for Economic Analysis*
*Department of Economics Wilfrid Laurier University*

For an economy with dysfunctional intertemporal financial markets the financial sector is modelled as a competitive banking sector offering deposit contracts. In a setting related to Allen and Gale (1998) properties of the optimal liquidity provision are analyzed for illiquid assets with ambiguous returns. In the context of our model, ambiguity — i.e. incalculable risk — leads to dynamically inconsistent investor behaviour. If the financial sector fails to recognize the presence of ambiguity, unanticipated fundamental crises may occur, which are incorrectly blamed on investors 'losing their nerves' and 'panicking'. The basic mechanism of the Financial

* w.spanjers@kingston.ac.uk Earlier versions of the paper were presented at the INFER Annual Conference 2008 in Évora, Portugal; at the Bundes- bank/FRIAS Workshop on 'Liquidity and Trust in Incomplete Markets' in Eltville, Germany, 15th - 17th March 2010; at the FMM Conference 'Stabilizing and Unequal Economy? Public Debt, Financial Regulation and Income Distribution' at the Hans B¨ockler Stiftung, Berlin, Germany, 29th - 30th October 2010; the INFINITI Conferences 2011 and 2013 in Dublin, Ireland, and Aix-en-Provence, France, and the European Economic Association's 28th Annual Conference in Gothenburg, Sweden, 26th - 30th August 2013, as well as research seminars at the University of Freiburg, Germany, the University of Bologna at Rimini, Italy, and at Kingston University London, Kingston-upon-Thames, United Kingdom.
The paper benefited from helpful comments from and discussions with Paul Auerbach, Jürgen Eichberger, Thomas Gehrig and participants in the conferences, the workshop and the research seminars. Financial support of the Freiburg Institute of Advanced Studies (FRIAS) is gratefully acknowledged.


* The mathematical structure of this paper, and the resulting Theorems, Propositions, and Corollaries, closely follows Spanjers (1999/2008, Chapter 5), where the proofs can be found. The contribution of this paper is the application of this mathematical structure and the interpretation of its results and policy implications in the context of the Financial Crisis.

Crisis resembles the liquidation of illiquid assets during a banking panic. The combination of providing additional liquidity and supporting distressed financial institutions implements the regulatory policy suggested by the model. A credible commitment to such 'bail-out policy' does not create a moral hazard problem. Rather, it implements the second best efficient outcome by discouraging excessive caution. Reducing ambiguity by increasing stability, transparency and predictability — as suggested by ordo-liberalism and the 'Freiburger Schule' — enhances ex-ante welfare.

**Keywords**: *Financial Intermediation, Liquidity, Ambiguity, Choquet Expected Utility, Financial Crisis*

**JEL Classification:** *D8, G1, G2*

# 1 Introduction

The Financial Crisis re-fueled the debate on the causes of — and regulatory responses to — asset price bubbles and the crises they may cause. Despite unresolved issues on how to define and identify bubbles, a range of measures to prevent them was discussed. Under the impression of the losses, the desirability of avoiding inflated asset prices seemed to be taken for granted. Still, it is not obvious that bubbles are bad and should be prevented. In particular, the question should be addressed if it may be preferable to refrain from taking additional preventive measures and, instead, to resolve crises when they occur.

As an example, consider the choice between a high technology growth strategy and one that relies on investments in low technology. Clearly, the former strategy is more prone to bubbles and crises than the latter. But if the difference in growth rates is high, crises do not occur too often, and the costs of a crisis are not excessive, the high technology strategy will be preferred. This is an example of 'optimal financial crises', as e.g. in Allen and Gale (1998), Spanjers (2008a) and Spanjers (2009).

The above trade-off may seem more relevant for developing countries than for the industrialized states from which the Financial Crisis originated. Still, even for developed economies bubbles may be the price for progress. The seeds for life-changing innovations tend to be laid in times of financial bubbles. Who would doubt that the benefits of the information and communication technology by far outweigh the costs of the dot.com bubble? Given the key role of technological progress in the long term improvement of living standards, such arguments in the spirit of Hayek (1935) deserve careful consideration. For more recent studies on the effects of bubbles on growth see e.g. Olivier (2000) and Caballero et al. (2006); for an empirical analysis of systemic crises and growth see e.g. Rancière et al. (2008).

Rather than on preventing bubbles and crises at (almost) any cost, the discussion should focus on how to cushion the impact of crises if and when they occur. When comparing the Great Depression, the Japan crisis in the 1990s, the Financial Crisis, and the Euro Crisis, it is clear that the effectiveness of dealing with crises has increased dramatically. For a discussion of the Financial Crisis see e.g. Hellwig (2008).

Hellwig (2008) provides compelling arguments for the incompleteness and imperfection of the financial markets having played a key role both in triggering the crisis and in amplifying its effects. It seems reasonable to assume that in normal times the intertemporal liquidity allocation of the financial sector of the economy can be reasonably approximated by the invisible hand as in the Arrow-Debreu model. We argue that this is not the case in times of financial crises, when the aggregate liquidity demand differs strongly from that anticipated by the financial sector. In the latter respect, the approach taken is fundamentally different from that in Holmström and Tirole (1997).

In situations where the aggregate liquidity demand differs strongly from that anticipated by the financial sector, the basic interactions in the financial system are not about the intra-temporal re-allocation of liquid and illiquid assets. Rather, they are about the decision whether to continue ongoing projects or to liquidate assets at a significant loss. For economies characterized by dysfunctional intertemporal markets, the aggregate institutional framework is better represented by the extreme case of an unregulated competitive banking sector, offering deposit contracts in the tradition of Diamond and Dybvig (1983). To capture the driving mechanism of the Financial Crisis, we extend the liquidity provision model of Eichberger and Harper (1997, Chapter 7) and Spanjers (1999/2008, Chapter 3) by introducing ambiguity (i.e. incalculable risk) regarding the payout structure of the illiquid asset.

Fundamental aspects of the model – e.g. the presence of ambiguity, deposit taking institutions, and insufficient levels of loss absorbing equity – also relate to specific instances of the Financial Crisis. In particular, they play a role in the drying up of liquidity in the inter-bank money market during the sub-prime mortgage crisis, in the run on Northern Rock in the United Kingdom, in the bail-out of the mortgage giants Freddie Mac and Fanny Mae, and in the collapse of Lehman Brothers. But we maintain that the relevance of the aggregate problems of intertemporal liquidity allocation surpass these specific instances.

Our point of departure is that each of these events was caused by an increase in incalculable risk or, as it is referred to in the relevant literature, ambiguity. Knight (1921) and Keynes (1937) provide an intuition for differentiating between (calculable) risk and (incalculable) ambiguity. We use a simple representation of ambiguity by Eichberger and Kelsey (1999) in the tradition of Ellsberg (1961). It is integrated in a linear model of liquidity provision with risky assets related to Jacklin and Bhattacharya (1988) and Allen and Gale (1998).

The effects of ambiguity in a financial and monetary setting have been analyzed in a number of papers. For the effects on financial markets see e.g. Dow and Werlang (1992), Epstein and Wang (1994) and Agliardi et al. (2015, and 2016); for the effects on financial institutions see e.g. Spanjers (1999/2008,), Eichberger and Spanjers (2008) and Spanjers (2008a). The impact of ambiguity on monetary policy has amongst others been addressed in Hansen and Sargent (2001), Wagner (2007), Ghatak and Spanjers (2007) and Spanjers (2008). Brach and Spanjers (2012) consider the impact of ambiguity in the form of incalculable political risk on development strategies and growth in the Middle East and Northern Africa.

In our model financial crises — which are modelled as banking panics in the financial sector of the economy — can be triggered by an increase in the level of ambiguity experienced by investors. Such a loss of confidence can be caused by events that are exogenous to the model. More interestingly, the arrival of new information on the prospects of the asset returns can lead to an endogenous loss of confidence through the updating of ambiguous beliefs. If one is not aware of the presence of ambiguity, this endogenous loss of confidence can easily be mistaken for an irrational overreaction by investors. Spanjers (1999/2008a) shows that this mechanism can also be used to explain the 1997 East-Asian crisis.

As some observers noted — one of them being the former President of the Bundesbank, Axel Weber — the problems with mortgage and asset backed securities had strong similarities with the bank runs in the theoretical models, as did the problems of hedge funds.

> ‘The current turmoil in the financial markets has all the characteristics of a classic banking crisis, but one that is taking place outside the traditional banking sector, Axel Weber, president of the Bundesbank, said at the weekend.[...]
>
> Some Federal Reserve policymakers also privately see comparisons between the current distress in credit markets and the bank runs of the 19th century, in which savers lost confidence in banks and demanded their money back, creating a spiralling liquidity crisis for institutions that had invested this money in longer-term assets.[...]
>
> “What we are seeing at the moment is a total overreaction,” he said. “There is no overall problem in terms of solvency – it is one of liquidity.” He said the challenge for central banks – which cannot

> supply liquidity directly to the non-bank sector – was to help banks absorb the influx of assets onto their balance sheets.'
>
> (Financial Times, 2nd September 2007)

We use our theoretical model to propose and evaluate policy measures. One possibility is to prevent financial crises by requiring that 'prudential' low-risk investment strategies are pursued. An alternative measure is for the public sector to underwrite 'toxic assets' to counteract the loss of confidence.

The results of our analysis are revealing. The effectiveness of requiring the financial sector to 'prudentially' follow low-risk investment strategies is less clear than may be expected. Like the financial system, our model is driven by undiversifiable systematic risk. Any regulation that reduces this undiversifiable risk must by its very nature affect the aggregate state contingent returns on investment.

The investment decisions in our model involve a choice between a low yielding liquid asset and an illiquid asset, which provides a high expected yield when it matures, but which incurs a loss in the case of premature liquidation. The latter occurs in a financial crisis. If the anticipated return on the illiquid asset is sufficiently higher than that of the liquid asset; if the costs of liquidation are low; and/or the occurrence of a crisis is unlikely, the benefits of the higher returns outweigh the losses incurred in the occasional crisis. Under such circumstances, any regulation causing the financial sector to 'prudentially' follow less risky investment strategies damages the long run prospects of the economy.

Considering the underwriting of 'toxic assets' by the public sector, we find that this is an effective policy measure which removes the distortive effects of the presence of ambiguity. Furthermore, the direct cost of underwriting only reflects the amount of risk that is insured, whereas its main impact is through the costless (and priceless) insurance of the unfounded fear caused by the presence of ambiguity.

An often-mentioned concern regarding the underwriting of toxic assets is that it creates a moral hazard problem, reducing financial institutions' incentives to follow prudent investment strategies. If the policy is anticipated, it is claimed to encourage overly risky investment decisions in the anticipation of being bailed out in a crisis. This argument does not apply for the model of this paper.

On the contrary, for the range of parameter values we consider, it is the potential loss due to the premature liquidation of illiquid assets that distorts incentives. These liquidation losses make financial institutions follow overly cautious investment strategies. As a result, the overall investment strategy is not as profitable as it could have been given the information constraints, i.e. the outcomes fails to be second best efficient.

The anticipation and implementation of a bail-out policy when a crisis occurs corrects these distorted incentives, thus removing a moral hazard problem, rather than creating one.

The remainder of the paper is organized as follows. Section 2 introduces the basic model. The second best efficient liquidity allocation is determined in Section 3. Section 4 analyzes the liquidity provision by an unregulated financial sector. In Section 5 it is established that the liquidity provision by an unregulated financial sector is not second best efficient and regulatory measures are discussed. Section 6 addresses the impact of a failure to recognize the presence of ambiguity and concluding remarks are made in Section 7. These remarks relate to dynamic inconsistency in updating ambiguous beliefs and provide policy recommendations.

## 2. The Basic Model

We consider a simplified linear model in the spirit of Diamond and Dybvig (1983) and Jacklin and Battacharya (1988). Investors contribute initial wealth to pooled investment vehicles ('banks') that choose between investing in a riskless liquid low yield asset and a risky illiquid asset with a high expected return. After one period, investors obtain two signals: a public signal regarding the perspectives of the risky asset and a private signal regarding their own immediate liquidity needs. On the basis of this information they decide whether or not to withdraw their initial contributions. If too many investors withdraw their contributions, illiquid assets will be liquidated and a 'banking panic' (i.e. a financial crisis) occurs.

The model differs from the usual setting in an important way. In our formulation investors' immediate liquidity needs are the basis for the withdrawal of funds, whereas in the setting of Diamond and Dybvig (1983) withdrawals are caused by an increase in the individual probability of not surviving to enjoy the future high returns of the illiquid assets. In terms of investors' preferences, this changed setting leads to a qualitatively different ex-ante expected utility function. 1

[1] In the context of the current paper the two approaches lead to very different results. Under the assumption of differences in immediate liquidity preference, the optimal reserve holdings for the social planner and for the representative bank are both determined by the incentive constraints for the ad interim patient investors. Under the assumption of differences in preference for late consumption, however, the optimal reserve holdings for the social planner are determined by the incentive constraint of the ad interim impatient investors, whereas the optimal reserve holdings for the representative bank remain determined by the incentive constraint of the patient investors. The result of Theorem 6 that the anticipation and implementation of a bail-out policy leads to the second best efficient outcome, however, carries over.

Although the original Diamond-Dybvig setting is excellently suited to address issues surrounding the design and analysis of pension funds and different systems of pension provision, we maintain it is less suited for the analysis of the recent Financial Crisis, which was driven by more immediate liquidity needs, rather than long term considerations.

### 2.1. Investment opportunities

The three period economy of our model consists of a continuum $\mathrm{I} \coloneqq [0,1]$ of ex-ante identical investors who at $\mathrm{t} = 0$ each have one unit of wealth. The economy offers two investment opportunities: a riskless liquid zero-yield asset called money and an illiquid asset which provides a high return when it matures and is successful, but leads to a severe loss when it matures and fails to be successful. In the period after investment decisions have been made, investors obtain a public signal regarding the prospects of the risky illiquid asset and a private signal regarding their individual liquidity needs. If they decide to prematurely liquidate their illiquid asset after receiving the signal, the severe loss can be prevented, but the initial investment will not be recovered in full.

The following table states the pay-outs of the investments.

| Investment | Period 0 | Period 1 | Period 2 Success | Period 2 Failure |
|---|---|---|---|---|
| Money 0 to 1 | $-1$ | 0 | 0 | 0 |
| Money 1 to 2 | 0 | $-1$ | 1 | 1 |
| Investment matured | $-1$ | 0 | $\alpha_h$ | $\alpha_\ell$ |
| Investment liquidated | $-1$ | $\alpha_1$ | 0 | 0 |

The probabilities associated with the public signal regarding the prospects of the illiquid asset are as in the table below. It is assumed that the investments carry no idiosyncratic risk. All risk related to their payouts is non-diversifiable systematic risk. The signal $\sigma \in \{\mathrm{b}, \mathrm{g}\}$ at $\mathrm{t} = 1$ can be interpreted as a forecast of the economy's prospects, the return $\varrho \in \{\mathrm{h}, \ell\}$ as the actual economic development.

| Signal $\sigma$ | Payout $\varrho$ | Probability $\pi_{\sigma\varrho}$ |
|---|---|---|
| $b$ | $\alpha_h$ | $\delta$ |
| $b$ | $\alpha_\ell$ | $\varepsilon$ |
| $g$ | $\alpha_h$ | $1-(\delta+\varepsilon)$ |
| $g$ | $\alpha_\ell$ | 0 |

The timing of the decisions of the investors is as follows:

| | |
|---|---|
| Period 0 | Investment decisions |
| Period 1 | Individual liquidity preference $\theta \in \{H, L\}$ becomes privately known |
| | Signal $\sigma \in \{b, g\}$ becomes publicly known |
| | Possibility for liquidation of assets |
| | Consumption in Period 1 |
| Period 2: | Return $\varrho \in \{h, \ell\}$ occurs |
| | Consumption in Period 2 |

**2.2. Beliefs**

The investors face uncertainty over their individual liquidity preference and over the combinations of signals and asset returns. This uncertainty is partly in the form of calculable risk and partly in the form of incalculable ambiguity.

In particular, the investors face ex-ante risk with respect to their individual liquidity preference, which is either high (H) or low (L) and which is represented by $\theta \in \{\mathrm{H}, \mathrm{L}\}$. The liquidity type of the investors is assumed to be independent of the asset state $(\sigma, \varrho) \in \{\mathrm{b}, \mathrm{g}\} \times \{\mathrm{h}, \ell\}$. In addition to risk about their liquidity type, investors face ambiguity over the combinations of signals – good (g) or bad (b) – and asset returns – high (h) and low ($\ell$) – that may arise.

Investors' uncertainty is over their individual state space $\{\mathrm{H}, \mathrm{L}\} \times \{\mathrm{b}, \mathrm{g}\} \times \{\mathrm{h}, \ell\}$. In the face of this uncertainty, their decisions will be guided by their beliefs over these combinations of potential outcomes. In particular, we assume investors' beliefs over $\{\mathrm{H}, \mathrm{L}\} \times \{\mathrm{b}, \mathrm{g}\} \times \{\mathrm{h}, \ell\}$ to be represented by an E(llsberg)-capacity $(\pi, \{\mathrm{F}_\theta\}_{\theta \in \{\mathrm{H}, \mathrm{L}\}}, \gamma)$ which consist of:

- an additive probability distribution $\pi$;
- a level of confidence $\gamma \in [0,1]$ in $\pi$; and
- an additive partition of the state space with components

$$F_H := \{(H, b, h), (H, b, \ell), (H, g, h), (H, g, \ell)\} \text{ and}$$
$$F_L := \{(L, b, h), (L, b, \ell), (L, g, h), (L, g, \ell)\}$$

The interpretation of these beliefs is as follows. Each investor has a conventional additive probability estimate of what may happen. This probability estimate is represented by the probability distribution $\pi$. The presence of ambiguity, however, causes the investors to have restricted confidence in the validity of this probability estimate. The level of confidence in the probability estimate is denoted by $\gamma$, where

$\gamma = 1$ denotes full confidence. In the case of $\gamma = 0$ the investor has no confidence in his probability estimate whatsoever; he thinks that anything may happen.

The final element of the investor's beliefs is represented by the additive partition of the state space. This partition captures the notion that investors may consider some aspects of the uncertainty they face to be best represented by calculable risk, while they may experience ambiguity with respect to other aspects. The components $F_H$ and $F_L$ being 'additive' as indicated above means that the investor has full confidence in the probabilities he assigns to the events $F_H$ and $F_L$, but fails to have full confidence in their sub-events and appropriate combinations thereof. The interpretation is that the investor faces risk with respect to his prospective individual liquidity needs, but faces ambiguity with respect to the signal and the return of the asset. Thus, investors' beliefs are represented by E-capacities as in Eichberger and Kelsey (1999). See also Chateauneuf et al. (2007).

### 2.3. Updating ambiguous beliefs

The updating of ambiguous beliefs differs from the updating of additive ones. In particular, there is a number of competing natural generalizations of Bayes' rule to the context with ambiguity. The first generalization that comes to mind when ambiguity is represented by belief functions is Full Bayesian updating as described in Jaffray (1992). This rule is best understood by considering the set of probability distributions that provides an equivalent representation of the belief function. In this setting, the Full Bayesian update is obtained by updating each of these probability distributions separately. Since the set of (multiple prior representations of) belief functions is closed under Full Bayesian updating, the updated set of priors once again represents a belief function.

The problem with this Full Bayesian updating is that in the context of our application the initial belief functions are arrived at using the axiomatic approach to Choquet expected utility as in Schmeidler (1982/89) and Gilboa (1987). In this approach the beliefs are represented by a capacity, which not only represents the ambiguity faced by the decision maker, but at the same time includes the decision maker's ambiguity attitude. Therefore, in using Full Bayesian updating, one would simultaneously update the ambiguous beliefs and the ambiguity attitude. Since the ambiguity attitude is best considered to be an individual characteristic of the decision maker – rather than a components of their beliefs – the implied updating of the ambiguity attitude would be inappropriate.

In the context of dynamic preferences that result from updating, Gilboa and Schmeidler (1993) address this problem by providing an axiomatization of updating rules. Their analysis is based on the same fundamental concepts — i.e. preference relations — that were used to derive the Choquet expected utility representation in the first place. They formulate 'reasonable' properties which a combination of dynamic preferences should satisfy. In the context of cautious/pessimistic decision makers they arrive at the Dempster-Shafer rule (see Dempster, 1968, and Shafer, 1976) as a plausible rule for updating. In the context of exuberant/optimistic decision makers, they arrive at Bayes' rule for updating capacities, which is not to be confused with Full Bayesian updating.

When applied to the multiple priors representation of belief functions, the Dempster-Shafer rule is a 'maximum likelihood' rule. It restricts attention to those probability distributions in the initial set of priors for which the received signal had the highest probability. The set of updated priors is now obtained as the set of Bayesian updates of these 'maximum likelihood' additive priors. The updating results in another belief function.

In our setting, the Full Bayesian update of the E-capacity and the Dempster-Shafer update of the E-capacity (which are also E-capacities) are identical. Therefore both interpretations of the capacity — describing ambiguity per se or describing a combination of ambiguity and ambiguity aversion — are consistent with our model.

Applying the result of Eichberger and Kelsey (1999) regarding the updating of E-capacities to our setting, we find that updating the ambiguous beliefs leads to:

- Bayesian updating of the probability distribution $\pi$; and
- an endogenous decrease in the level of confidence: $\gamma^{\sigma} < \gamma$.

Updating the E-capacity $(\pi, \{F_\theta\}_{\theta \in \{H,L\}}, \gamma)$ after receiving the information $\sigma$ for $\gamma \in (0,1)$ one obtains the E-capacity $\left(\pi^{\sigma}, \{F^{\sigma}_\theta\}_{\theta \in \{H,L\}}, \gamma^{\sigma}\right)$ with

| | $(\theta, b)$ | $(\theta, g)$ |
|---|---|---|
| $\pi^{\sigma}_{\theta}$ | $\frac{\delta}{\delta+\varepsilon}$ | 1 |
| $\gamma^{\sigma}$ | $\gamma \cdot \frac{\pi_b}{1-\gamma \cdot \pi_g} = \gamma \cdot \frac{\delta + \varepsilon}{1-\gamma \cdot (1-(\delta+\varepsilon))} \leq \gamma$ | $\gamma \cdot \frac{\pi_g}{1-\gamma \cdot \pi_b} = \gamma \cdot \frac{1-(\delta+\varepsilon)}{1-\gamma \cdot (\delta+\varepsilon)} \leq \gamma$ |
| $F^{\sigma}_{\theta}$ | $\{\{(\theta, b, h), (\theta, b, \ell)\}\}$ | $\{\{(\theta, g, h), (\theta, g, \ell)\}\}$ |

## 2.4. Preferences

The basic rational for the subjective expected utility approach is to describe beliefs separately from the evaluation of outcomes if and when they are attained. The beliefs typically relate to the likelihood with which certain outcomes or states of nature are expected to occur. The evaluation of outcomes normally takes place by a von

Neumann-Morgenstern utility index or — e.g. in the case of prospect theory — by a value function. In order to represent standard preference relations, the beliefs and the evaluation of outcomes are combined through an evaluation functional, e.g. by taking the expected value.

In the Choquet expected utility approach, a von Neumann-Morgenstern utility index is applied for the evaluation of outcomes if and when they occur. The combination of the ambiguous beliefs and the ambiguity attitude of the decision maker are described by a non-additive probability distribution called a capacity. The Choquet integral is used to evaluate the von Neumann-Morgenstern utility index over such capacity. It is our purpose to analyse the impact of the level of ambiguity, rather than the impact of the ambiguity attitude, so we assume throughout the paper that in the face of ambiguity decision makers have a constant ambiguity attitude of full pessimism.

To obtain a clear separation of the effects of risk from the effects of ambiguity, we assume investors are risk neutral. For an investor of (ad interim) liquidity type $\theta \in \{H, L\}$ we have the von Neumann- Morgenstern utility index

$$u(x_1(\sigma), x_2(\sigma, \varrho); \theta) = \beta_\theta \cdot x_1(\sigma) + x_2(\sigma, \varrho)$$

where $\beta_H > \beta_L > 1$ reflects the intensity of the investor's preference for liquidity at $t = 1$.

An investor derives ex-ante utility from a state contingent income bundle

$$x := \left(x_1^\theta(\sigma), x_2^\theta(\sigma, \varrho)\right)_{(\theta,\sigma,\varrho)\in\{H,L\}\times\{b,g\}\times\{h,\ell\}}.$$

The minimal utility obtained for liquidity preference $\theta$ within the additive component $F_\theta$, given the type contingent income $x^\theta(\sigma, \varrho) := \left(x_1^\theta(\sigma), x_2^\theta(\sigma, \varrho)\right)$, is denoted by

$$m_\theta(x) := \min_{(\theta,\sigma,\varrho)\in F_\theta} u\left(x^\theta(\sigma, \varrho); \theta\right) = \min_{(\sigma,\varrho)\in\{b,g\}\times\{h,el\}} u\left(x^\theta(\sigma, \varrho); \theta\right).$$

The investor's ex-ante Choquet expected utility function over a state contingent income bundle $x$ for beliefs $\left(\pi, \{F_\theta\}_{\theta\in\{H,L\}}, \gamma\right)$ is now obtained as

$$U(x) := \gamma \cdot \boldsymbol{E}_{(\theta,\sigma,\varrho)}\left\{u(x^\theta(\sigma, \varrho); \theta\right\} + (1 - \gamma) \cdot \boldsymbol{E}_\theta\{m_\theta(x)\}.$$

### 2.5. Indirect Utility Representations

When looking for graphical illustrations of ex-ante decision problems in a two dimensional diagram, one runs into obvious problems depicting investors' indifference curves. We circumvent these problems by using an indirect utility representation. This representation is based on type dependent income in terms of ex-ante money holdings and ex-ante investments.

In particular, for given fractions of money holdings $\mu$ and investments $1-\mu$, we consider $y_H := \frac{\mu}{\tau_H}$ and $y_L := \frac{1-\mu}{\tau_L}$, where $\tau_H := \pi_H$ and $\tau_L := \pi_L$ denote the population fractions of $H-$type investors and $L-$type investors, respectively. This enables us to depict the feasible combinations as a 'budget line' in a $(y_H, y_L)-$diagram, independent of the institutional framework under consideration. To complete the illustration in the two dimensional diagram all we need is the appropriate counterpart of indifference curves.

For this purpose we consider indifference curves of the indirect utility functions, which evaluate the outcomes obtained under the different institutional framework for various combinations $(\mathrm{y_H}, \mathrm{y_L})$. For some of the institutional settings, we may find that there are different ways to define the associated indirect utility function V for out-of-equilibrium combinations $(\mathrm{y_H}, \mathrm{y_L})$. Therefore, the indirect utility functions require the specification of assumed out-of-equilibrium reactions.

Once these indirect utility functions are arrived at, the decision problems for different institutional settings can be depicted within a single diagram. In particular, such diagram reveals whether different institutional settings lead to the same aggregate equilibrium money holdings and investments. But even if the equilibrium money holdings and investments are identical for two institutional settings, they may lead to different ex-ante utility levels for the investors.

### 2.6. Assumptions on Parameter Values

In the remainder of this paper, the following simplifying assumptions regarding the parameter values apply:

1. $\delta, \varepsilon > 0$ and $\delta + \varepsilon < 1$

   The assessed probability of each of the asset states $(b, h)$, $(b, \ell)$ and $(g, h)$ exceeds zero.

2. $\alpha_\ell = 0$.

   The asset return in the case of failure is zero.

3. $\beta_H > \alpha_h > \beta_L$.

From the ad interim perspective and assuming the asset is successful, $H-$types would have preferred holding their initial wealth as money, but $L-$types would have preferred investing in the illiquid asset.

4. $\alpha_1 \cdot \beta_H < \boldsymbol{CE}_\varrho\{\alpha_\varrho | b\}$.
   Ad interim, $H$-type investors prefer the asset to mature, even after receiving a bad signal regarding the assets' prospects.

5. $\bar{\alpha}^b = \boldsymbol{E}_\varrho\{\alpha_\varrho | b\} < \beta_L$.
   This ensures that the aggregate fractional money holdings $\mu_b(\gamma)$ are not second best efficient (see Theorem 1 below). It follows that for each level of confidence $\gamma$ we have $\beta_L > \tilde{\alpha}^b \coloneqq \boldsymbol{CE}_\varrho\{\alpha_\varrho | b\}$.

## 3. Second-Best Efficiency

As a point of reference we consider second best efficiency. The results of any financial system based on voluntary participation cannot be worse than autarky. Nor can it outperform the second best efficient outcome. In particular, we analyze the symmetric ex-ante second best efficient liquidity allocation. This can be visualized as a social planner who maximizes the ex-ante utility of the ex-ante (at $t = 0$) identical investors. Since the presence of ambiguity leads to dynamic inconsistency in the updating of investors' beliefs, one must distinguish between the ex-ante beliefs used to evaluate the outcomes at $t = 0$ and the updated beliefs which determine the investors' ad interim behavior at $t = 1$.

### 3.1. The Planer's Decision Problem

To determine a second-best efficient outcome, a social planner offers an incentive compatible contract which maximizes the ex-ante utility of the (ex-ante identical) investors. At $t = 0$, the planner decides which fraction $1 - \mu$ of investors' wealth is invested in the illiquid assets. For each liquidity type $\theta \in \{H, L\}$ at $t = 1$, the contract specifies payouts $x_1^\theta(\sigma)$ at $t = 1$ which are contingent on the public signal $\sigma \in \{b, g\}$. Furthermore, payouts $x_2^\theta(\sigma, \varrho)$ for $t = 2$ are specified, contingent on both the public signal $\sigma \in \{b, g\}$ at $t = 1$ and the return status $\varrho \in \{h, \ell\}$ of the illiquid asset at $t = 2$. Since the liquidity type of the individual investors is private knowledge, the ex-ante utility is maximized subject to feasibility and incentive constraints.

At $t = 0$, the planner decides for each signal $\sigma \in \{b, g\}$ the fraction $\lambda(\sigma)$ of assets that will be liquidated at $t = 1$ and the fraction $\rho(\sigma)$ of the money holdings that will be transferred to $t = 2$. 2 The planner's objective function is:

$$\boldsymbol{CE}_{(\theta,\sigma,\varrho)}\left\{\beta_\theta \cdot x_1^\theta(\sigma) + x_2^\theta(\sigma,\varrho)\right\} =$$

$$\sum_{\theta\in\{H,L\}} \pi_\theta \cdot \left[\sum\nolimits_{(\sigma,\varrho)\in\{b,g\}\times\{h,\ell\}} \gamma \cdot \left[\beta_\theta \cdot x_1^\theta(\sigma') + x_2^\theta(\sigma',\varrho')\right]\right] +$$
$$\sum_{\theta\in\{H,L\}} \pi_\theta \cdot (1-\gamma) \cdot \left[\min_{(\sigma',\varrho')\in\{b,g\}\times\{h,\ell\}} \left[\beta_\theta \cdot x_1^\theta(\sigma') + x_2^\theta(\sigma',\varrho')\right]\right]$$

The planner now faces the following decision problem:

$$\max_{(x_1^\theta(\sigma),\, x_2^\theta(\sigma,\varrho))_{(\theta,\sigma,\varrho)\in\{H,L\}\times\{b,g\}\times\{h,\ell\}},\mu,(\lambda(\sigma),\rho(\sigma))_{\sigma\in\{b,g\}}} \boldsymbol{CE}_{(\theta,\sigma,\varrho)}\left\{\beta_\theta \cdot x_1^\theta(\sigma) + x_2^\theta(\sigma,\varrho)\right\}$$

such that

$$\pi_H \cdot x_1^H(\sigma) + \pi_L \cdot x_1^L(\sigma) = \big(1-\rho(\sigma)\big) \cdot [\alpha_1 \cdot \lambda(\sigma) \cdot (1-\mu) + \mu] \qquad (F_1)$$

$$\pi_H \cdot x_2^H(\sigma,h) + \pi_L \cdot x_2^L(\sigma,h) =$$
$$\rho(\sigma) \cdot [\alpha_1 \cdot \lambda(\sigma) \cdot (1-\mu) + \mu] + \alpha_h \cdot (1-\mu) \cdot (1-\lambda(\sigma)) \qquad (F_{2h})$$

$$\pi_H \cdot x_2^H(\sigma,\ell) + \pi_L \cdot x_2^L(\sigma,\ell) = \rho(\sigma) \cdot [\alpha_1 \cdot \lambda(\sigma) \cdot (1-\mu) + \mu] \qquad (F_{2\ell})$$

$$\beta_H \cdot x_1^H(\sigma) + \boldsymbol{CE}_\varrho\{x_2^H(\sigma,\varrho)|\sigma\} \geq \beta_H \cdot x_1^L(\sigma) + \boldsymbol{CE}_\varrho\{x_2^L(\sigma,\varrho)|\sigma\} \qquad (IC_H^\sigma)$$

$$\beta_L \cdot x_1^L(\sigma) + \boldsymbol{CE}_\varrho\{x_2^L(\sigma,\varrho)|\sigma\} \geq \beta_L \cdot x_1^H(\sigma) + \boldsymbol{CE}_\varrho\{x_2^H(\sigma,\varrho)|\sigma\} \qquad (IC_L^\sigma)$$

In Section 3.3 this problem is solved, but first we take a closer look at the incentive constraints at $t = 1$.

### 3.2. The Ad Interim Constrains

The constraints at $t = 1$ refer to the situation after the planner made the investment decision at $t = 0$, investors learnt their individual liquidity preference type $\theta \in \{H, L\}$, and the economy received the signal $\sigma \in \{b, g\}$. Because of $\beta_H > \beta_L > 1$, it is optimal to pay out the entire money holdings at $t = 1$, so $\rho^E(b) = \rho^E(g) = 0$.

[2] For the assumed parameter values, the optimal solution has $\lambda(\sigma) \cdot \rho(\sigma) = 0$.

Due to the linear structure of the economy and since $\beta_H > \alpha_h$, the first best efficient solution would be to hold the entire wealth of the economy as money at $t = 0$ and transfer the money at $t = 1$ to the investors with a high liquidity preference. So one would intuitively expect that the second best efficient outcome is characterized by the largest redistribution from $L$-types to $H$-types that still satisfies incentive compatibility – an intuition that will be confirmed in Section 3.3.

For the aggregate fractional money holdings $\mu$, denote $y_H(\mu) \coloneqq \frac{\mu}{\pi_H}$ and $y_L(\mu) \coloneqq \frac{1-\mu}{\pi_L}$. Let

$$\tilde{\alpha}^\sigma \coloneqq \boldsymbol{CE}_\varrho\{\alpha_\varrho | \sigma\} = \gamma^\sigma \cdot \overline{\alpha}^\sigma = \gamma^\sigma \cdot \pi_h^\sigma \cdot \alpha_h.$$

For $\sigma \in \{b, g\}$ denote by $\mu_L^\sigma(\gamma^\sigma)$ the fractional money holdings such that after signal $\sigma$ is received, the interim incentive constraint ($IC_L^\sigma$) holds with equality if:

- the entire money holdings are paid out to $H$-types at $t = 1$ and
- the entire returns of the illiquid investments are paid out to $L$-types at $t = 2$.

That is, for the updated level of confidence $\gamma^\sigma$ and the type-contingent payouts $x_1^H(\sigma) = y_H(\mu_L^\sigma(\gamma^\sigma))$, $x_2^H(\sigma, \varrho) = 0$, $x_1^L(\sigma) = 0$, $x_2^L(\sigma, \varrho) = \alpha_\varrho \cdot y_L(\mu_L^\sigma(\gamma^\sigma))$, we have

$$\beta_L \cdot x_1^H(\sigma) = \boldsymbol{CE}_\varrho\{x_2^L(\sigma, \varrho) | \sigma\}.$$

After expanding the Choquet expected value, substituting out $x_1^H(\sigma)$ and $\left(x_2^L(\sigma, h), x_2^L(\sigma, \ell)\right)$, and rearranging terms, we obtain

$$\mu_\sigma(\gamma) \coloneqq \mu_L^\sigma(\gamma^\sigma) = \frac{\pi_H \cdot \tilde{\alpha}^\sigma}{\pi_L \cdot \beta_L + \pi_H \cdot \tilde{\alpha}^\sigma}$$

where $\gamma$ denotes the ex-ante level of confidence.

Since $\tilde{\alpha}^b < \tilde{\alpha}^g$, it follows that $\mu_g(\gamma) > \mu_b(\gamma)$.

If for money holdings $\mu_\sigma(\gamma)$ signal $\sigma$ is received, the type-contingent payouts $x_1^H(\sigma) = y_H(\mu_\sigma(\gamma))$, $x_2^H(\sigma, \varrho) = 0$, $x_1^L(\sigma) = 0$, $x_2^L(\sigma, \varrho) = \alpha_\varrho \cdot y_L(\mu_\sigma(\gamma))$, are optimal. Next we determine the optimal type-contingent payouts for the case where money holdings and signal b fail to match.

In case signal b is received for money holdings $\mu_g(\gamma)$, paying out the entire money holdings to H-types violates the incentive compatibility of investors with a low liquidity preference. To restore incentive compatibility, L-types will need to receive some payout at $t = 1$ over and above the entire returns of the illiquid investments at $\mathrm{t}=2$. Denoting the fraction of the liquidity reserves at that are paid out to L-types at $\mathrm{t}=1$ by $\widehat{\rho}$ we find:

$$\hat{\rho} = \pi_L \cdot \left(1 - \frac{\tilde{\alpha}^b}{\tilde{\alpha}^g}\right) = \pi_L \cdot \left(1 - \frac{\gamma^b}{\gamma^g} \cdot \pi_h^b\right).$$

Next consider the opposite case with fractional money holdings $\mu_b(\gamma)$ and signal $g$. Since $\mu_g(\gamma) > \mu_b(\gamma)$, the incentive constraint for $L-$types, $(IC_L^g)$, is satisfied if the money holdings are paid out to $H-$types at $t=1$ and the returns of the illiquid assets are paid out to $L-$types at $t=2$. The incentive constraint will be satisfied, but the money holdings will be too low for this incentive constraint to hold with equality. Depending on the parameter values, the following three cases may potentially occur.

Firstly, the type-contingent payouts $\mathrm{x}_1^{\mathrm{H}}(\sigma) = \mathrm{y}_{\mathrm{H}}(\mu_\sigma(\gamma))$, $\mathrm{x}_2^{\mathrm{H}}(\sigma, \varrho) = 0$, $\mathrm{x}_1^{\mathrm{L}}(\sigma) = 0$, $\mathrm{x}_2^{\mathrm{L}}(\sigma, \varrho) = \alpha_\varrho \cdot \mathrm{y}_{\mathrm{L}}(\mu_\sigma(\gamma))$, satisfy the incentive constraint for H−types, $(\mathrm{IC}_{\mathrm{H}}^{\mathrm{g}})$, and are optimal. In the second case, $(\mathrm{IC}_{\mathrm{H}}^{\mathrm{g}})$ is violated but can best be restored by making payouts to H−types at $\mathrm{t}=2$, over and above giving them the entire money holdings at $\mathrm{t}=1$. In the third and final case, $(\mathrm{IC}_{\mathrm{H}}^{\mathrm{g}})$ is violated and is efficiently restored by liquidating some of the illiquid assets at $\mathrm{t}=1$ and paying the revenue to the H−types.

Denote the fraction of the asset holdings that is liquidated at $\mathrm{t}=1$ by $\lambda$ and the fraction of the illiquid assets whose revenue is paid out to H−types at $\mathrm{t}=2$ by $\widehat{\lambda}$. By the assumptions of Section 2.6 we have $\lambda=0$. For $(\mathrm{IC}_{\mathrm{H}}^{\mathrm{g}})$ to be binding we must have

$$\hat{\lambda} = \pi_H \cdot \left[1 - \frac{\beta_H}{\beta_L} \cdot \frac{\gamma^b}{\gamma^g} \cdot \pi_h^b\right].$$

### 3.3. The Ex-Ante Problem

After having analyzed the ad interim constraints we turn our attention to the planner's ex-ante decision problem. The following theorem confirms that the linear structure of the model allows us to focus on corner solutions.

**Theorem 1** *Under the assumptions of Section 2.6, there exists a level of confidence* $\hat{\gamma}^E \in (0,1)$ *such that for the second best efficient money holdings we have*

$$\mu^E(\gamma) := \begin{cases} \mu_g(\gamma) & for \quad \gamma \in (\hat{\gamma}^E, 1] \\ any\ value\ in\ [\mu_g(\gamma), 1] & for \quad \gamma = \hat{\gamma}^E \\ 1 & for \quad \gamma \in [0, \hat{\gamma}^E) \end{cases}$$

*where*

$$\mu_g(\gamma) = \frac{\pi_H \cdot \widetilde{\alpha}^g}{\pi_L \cdot \beta_L + \pi_H \cdot \widetilde{\alpha}^g} = \frac{\pi_H \cdot \gamma^\sigma \cdot \alpha_h}{\pi_L \cdot \beta_L + \pi_H \cdot \gamma^\sigma \cdot \alpha_h} < 1.$$

From these optimal reserve holdings we obtain the second-best efficient contract $(x^{HE}, x^{LE}, \mu^E(\gamma))$ with for all $(\sigma, \varrho) \in \{b, g\} \times \{h, \ell\}$:

$$x_1^{HE}(\sigma) := \begin{cases} 1 & for \quad \mu^E(\gamma) = 1 \\ \left(1 + \frac{\pi_L}{\pi_H} \cdot \frac{\gamma^b}{\gamma^g} \cdot \pi_h^b\right) \cdot \mu^E(\gamma) & for \quad \mu^E(\gamma) \neq 1 \; and \; \sigma = b \\ \frac{\mu^E(\gamma)}{\pi_H} & for \quad \mu^E(\gamma) \neq 1 \; and \; \sigma = g \end{cases}$$

and $x_2^{HE}(\sigma, \varrho) := 0$, as well as

$$x_1^{LE}(\sigma) := \begin{cases} 1 & for \quad \mu^E(\gamma) = 1 \\ \left(1 - \frac{\gamma^b}{\gamma^g} \cdot \pi_h^b\right) \cdot \mu^E(\gamma) & for \quad \mu^E(\gamma) \neq 1 \; and \; \sigma = b \\ 0 & for \quad \mu^E(\gamma) \neq 1 \; and \; \sigma = g \end{cases}$$

and

$$x_2^{LE}(\sigma, \varrho) := \begin{cases} 0 & for \quad \mu^E(\gamma) = 1 \\ \alpha_\varrho \cdot \frac{(1 - \mu^E(\gamma))}{\pi_L} & for \quad \mu^E(\gamma) \neq 1 \; and \; \sigma = b \\ \alpha_\varrho \cdot \frac{(1 - \mu^E(\gamma))}{\pi_L} & for \quad \mu^E(\gamma) \neq and \; \sigma = g. \end{cases}$$

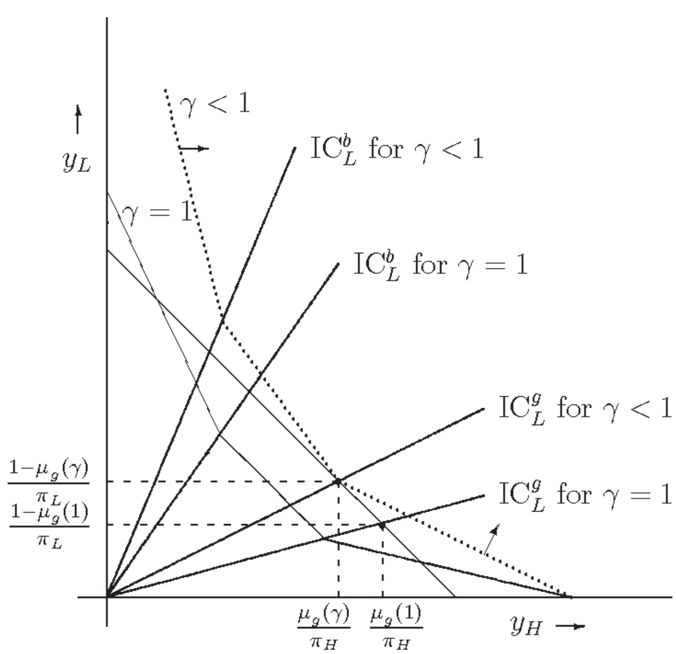
This follows directly from Theorem 1 and the fraction of liquidity reserves paid out to L−types, $\widehat{\rho}$, as indicated above. Other contracts which obtain the same ex-ante Choquet expected utility through an incentive compatible redistribution of payouts at $t = 2$ are also second-best efficient.

The planner's decision problem is illustrated in Figure 1. The horizontal axis depicts the money holdings per type$-H$ investor, $y_H := \frac{\mu}{\pi_H}$. The investment in the asset per $L-$type investor, $y_L := \frac{1-\mu}{\pi_L}$, is on the vertical axis. If all money holdings are paid out to $H-$types at $t = 1$, each of them receives $y_H$. Similarly, if the returns of the matured investment are paid out to $L-$type investors, they each receive $\alpha_\varrho \cdot y_L$ where $\varrho \in \{h, \ell\}$. Since $\alpha_\ell = 0$, we have $x_2^L(\ell) = 0$.

The feasibility line denotes the combinations $(y_H, y_L)$ the planner can obtain through his ex-ante choice of money holdings and investments. The indifference curves relate to the indirect ex-ante utility function, $V^E$, assuming that if incentive constraints are violated, incentive compatibility is efficiently restored through appropriate redistributions between H$-$types and L$-$types. The $(IC_L^b)$-curve denotes the combinations $(y_H, y_L)$ for which the incentive compatibility constraint of the type$-$L investors holds with equality after a bad signal is received. Similarly, the $(IC_L^g)$-curve denotes the combinations $(y_H, y_L)$ for which the incentive constraint of L$-$types holds with equality after a good signal.

The indifference curves of $V^E$ have two kinks, one at the $(IC_L^b)$-curve and one at the $(IC_L^g)$-curve. The $(IC_H^g)$-curves do not affect the indifference curves since the re-distribution needed to restore $(IC_H^g)$ only involves re-distribution of income at $t = 2$.

Figure 1: The Decision Problem of the Social Planner.



Income at $t = 2$ enters the vNM utility index of both types of investors in the same linear way, so such redistributions do not affect the ex-ante Choquet expected utility. We consider three cases.

Firstly, to the left of the $(IC_L^b)$-line, the payouts of $y_H$ to H−types at $t = 1$ and of $\alpha_\varrho \cdot y_L$ to L−types at $t = 2$ either are incentive compatible, or can be made incentive compatible by a redistribution of payouts at $t = 2$ from L−types to H−types that does not affect the ex-ante utility of the investors. The slope of the indifference curves is $\frac{\pi_H}{\pi_L} \cdot \frac{\gamma \cdot \beta_H}{\gamma \cdot \pi_h \cdot \alpha_h} = \frac{\pi_H}{\pi_L} \cdot \frac{\beta_H}{\pi_h \cdot \alpha_h}$, so in this area the indifference curves are steeper than the feasibility line whenever $\beta_H > \pi_h \cdot \alpha_h$.

Secondly, to the right of the $(IC_L^b)$-line but to the left of the $(IC_L^g)$-line, the payouts of $y_H$ to $H$−types and of $\alpha_\varrho \cdot y_L$ to $L$−types are incentive compatible after a good signal. But after a bad signal the incentive compatibility of type−$L$ investors is violated and a redistribution from $H$−types to $L$−types is needed. The efficient redistribution is to continue paying out all revenues from the illiquid asset at $t = 2$ to $L$−types, but to provide them with some of the money holdings at $t = 1$ too. As a consequence – as is confirmed by Theorem 1 – the indifference curves between the $(IC_L^b)$-line and the $(IC_L^g)$-line are steeper than the feasibility line, due to the assumptions in Section 2.6.

In the third and final case, to the right of the $(IC_L^g)$-line, the incentive compatibility of the $L$−type investors is violated for both signals. After each signal, a redistribution from $H$−types to $L$−types is needed. The efficient redistribution is to continue paying all revenues from the illiquid asset at $t = 2$ to $L$−types, but also providing them with some of the money holding at $t = 1$. As follows from Theorem 1, in this area the indifference curves are flatter than the feasibility line, unless $\gamma \leq \hat{\gamma}^E$.

In Figure 1, whenever the indifference curves are steeper than the feasibility line to the left of $(IC_L^g)$, but flatter to the right of it, $\mu_g(\gamma)$ is the second best efficient level of money holdings.

The expression for $\mu_g(\gamma)$ in Theorem 1 indicates that for $\gamma > \hat{\gamma}^E$ the second best efficient money holdings fall when the level of confidence decreases. So for sufficiently large levels of confidence, investment in the illiquid asset increases as the asset becomes more ambiguous.

The intuition for this is as follows. The ad interim Choquet expected utility that investors derive from the matured illiquid asset is $\gamma^\sigma \cdot \pi_h^\sigma \cdot \alpha_h \cdot y_L$, where a reduction in $\gamma$ causes a reduction in the updated level of confidence $\gamma^\sigma$. For the incentive constraint for $L-$type investors to be satisfied, this reduction in $\gamma^{\sigma}$ must be (partially) compensated for by an increase in $y_L := \frac{1-\mu}{\pi_L}$, the only variable in the expression for the ad interim Choquet expected utility. This requires an increase in $1-\mu$, the fractional investment in the illiquid asset, reducing its 'price' at $t=1$ and increasing the spread between the long term yield and the short term yield of wealth entrusted to the social planner at $t=0$.

**Corollary 2** *For any two levels of confidence $\gamma \in (\hat{\gamma}^E, 1]$ and $\gamma' \in (\hat{\gamma}^E, \gamma)$, we have $\mu^E(\gamma') < \mu^E(\gamma)$, i.e. a loss of confidence in the asset returns reduces the overall money holdings.*

## 4. The Unregulated Financial Sector

The standard form of financial intermediation to provide liquidity is by deposit taking institutions, typically referred to as 'banks'. The financial sector of an economy in which the liquidation of illiquid assets becomes an issue is modeled as a competitive banking sector with free entry, in which banks have no equity. In the context of linear asset returns and ex-ante identical investors as in our model, the competitive banking sector can be represented by a single representative bank that operates under state contingent zero profit constraints at $t=2$ and maximizes the ex-ante utility function of the investors who deposit their wealth.

### 4.1. Deposit contracts

A deposit contract $(r_1, r_2, m)$ is a combination of promised payouts of either $r_1$ at $t=1$, or $r_2$ at $t=2$, and fractional liquidity reserves $m$. The deposit contract states that requests for the payout of $r_1$ at $t=1$ have priority over requests for later payout at $t=2$. When the requested payouts at $t=1$ exceed the bank's capacity for repayment after the liquidation of its illiquid assets, the available resources are distributed proportional to these requests. The resulting effective payouts are denoted by $(x_1, x_2)$.

The effective payouts are a function of the deposit contract $(r_1, r_2, m)$, the aggregate fractional withdrawals $w$ at $t=1$ and, in the case of $x_2$, of the asset return status $\varrho \in \{h, \ell\}$. The zero state-contingent profit condition allows for the effective payouts to be represented by a combination $(w; r, m)$, where $r = r_1$ and where $(r, m)$ is a short hand for the deposit contract $\left(r, \alpha_h \cdot \frac{r \cdot (1-m)}{r-m}, m\right)$, using $r = \frac{m}{\pi_H}$. For fractional aggregate withdrawals $w \in [0,1]$ the implied effective payouts are

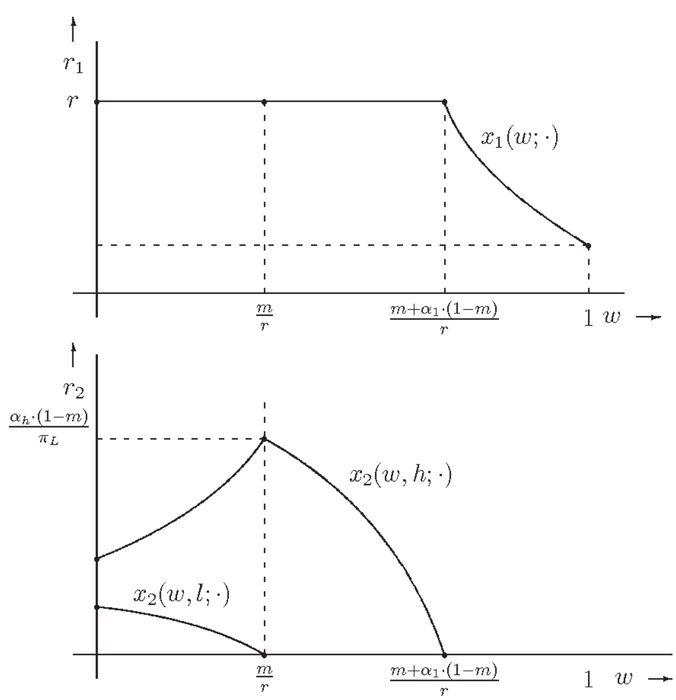
$$x_1(w;r,m) := \begin{cases} r & if \quad w \in \left[0, \frac{m+\alpha_1\cdot(1-m)}{r}\right] \\ \frac{1}{w}\cdot[\alpha_1+(1-\alpha_1)\cdot m] & if \quad w \in \left[\frac{m+\alpha_1\cdot(1-m)}{r}, 1\right] \end{cases}$$

and

$$x_2(w,\varrho;r,m) := \begin{cases} \frac{1}{1-w}\cdot\left[m-w\cdot r+\alpha_\varrho\cdot(1-m)\right] & if \quad w\in\left[0,\frac{m}{r}\right] \\ \frac{\alpha_\varrho}{1-w}\cdot\left[1+\left(\frac{1}{\alpha_1}-1\right)\cdot m-\frac{1}{\alpha_1}\cdot w\cdot r\right] & if \quad w\in\left[\frac{m}{r},\frac{m+\alpha_1\cdot(1-m)}{r}\right] \\ 0 & if \quad w\in\left[\frac{m+\alpha_1\cdot(1-m)}{r},1\right] \end{cases}$$

The effective payouts $x_1(w;\cdot)$, $x_2(w,h;\cdot)$ and $x_2(w,\ell;\cdot)$ are depicted in Figure 2.

Figure 2: Effective Payouts at $t=1$ and $t=2$.



### 4.2. The Representative Bank's Decision Problem

At $t = 0$ the representative bank offers a deposit contract that maximizes the ex-ante Choquet expected utility of the (ex-ante identical) individual investors. The investors are assumed to have perfect foresight regarding the aggregate withdrawal behavior at $t = 1$. As each individual investor is negligible, the aggregate withdrawal behavior is assumed to be independent of the investor's own behavior. In particular, for the individual investors, the dynamic inconsistency in the behavior of the *other* investors is known and anticipated, but the dynamic inconsistency of their *own* behavior is not. At $t = 0$, the ex-ante Choquet expected utility is maximized with respect to the beliefs $\left(\pi, \{F_\theta\}_{\theta\in\{H,L\}}, \gamma\right)$.

The decision problem of the representative bank is

$$\max_{(r,\mu),\varpi,\lambda(\varpi),\rho(\varpi)} \boldsymbol{CE}_{(\theta,\sigma,\varrho)}\{u(x(\varpi(\sigma;r,\mu),\varrho;r,\mu);\theta)\}$$

such that

$$\varpi \coloneqq \varpi(\sigma;r,\mu) \in \varpi^*(\sigma;r,\mu) \tag{$E$}$$

$$\varpi \cdot x_1(\varpi;\cdot) = \left(1-\rho(\varpi)\right)\cdot[\mu + \alpha_1 \cdot \lambda(\varpi)\cdot(1-\mu)] \tag{$F_1$}$$

$$(1-\varpi)\cdot x_2(\varpi,h;\cdot) =$$

$$\rho(\varpi)\cdot[\mu+\alpha_1\cdot\lambda(\varpi)\cdot(1-\mu)] + \alpha_h\cdot(1-\lambda(\varpi))\cdot(1-\mu) \tag{$F_{2h}$}$$

$$(1-\varpi)\cdot x_2(\varpi,\ell;\cdot) = \rho(\varpi)\cdot[\mu+\alpha_1\cdot\lambda(\varpi)\cdot(1-\mu)] \tag{$F_{2\ell}$}.$$

Here $\varpi^*(\sigma;r,\mu)$ denotes the set of withdrawal equilibria for the deposit contract $(r,\mu)$ when the signal $\sigma$ is obtained; $\rho$ denotes the fraction of reserve holdings transferred from $t = 1$ to $t = 2$; and $\lambda$ denotes the fraction of assets liquidated at $t = 1$. Constraint $(E)$ demands a withdrawal equilibrium; $(F_1)$ is the feasibility constraint for $t = 1$ and $(F_{2h})$ and $(F_{2\ell})$ are the feasibility constraints for $t = 2$ if the asset returns status is $h$ and $\ell$, respectively.

The formulation of the optimization problem assumes that at $t = 1$ the best possible withdrawal equilibrium is obtained. This implies that banking panics (i.e. simultaneous bank runs on all individual banks in the competitive banking sector) only occur if they are the unique withdrawal equilibrium. Bank runs and banking panics that result from coordination failures amongst the depositors are disregarded.

### 4.3. Ad Interim Withdrawal Equilibria

In solving the decision problem of the representative bank, we first consider the withdrawal equilibria $\varpi^*(\sigma; r, \mu)$ at $t = 1$. For this it is assumed that at $t = 0$ all investors deposited their entire wealth in the bank.

Firstly, consider a deposit contract $(r, \mu)$ such that after receiving signal $\sigma$ the incentive constraint for the $L$−type investors is violated if all $H$−types withdraw at $t = 1$. That is, if for the effective payouts $x_1(\pi_H; r, \mu)$ and $x_2(\pi_H, \varrho; r, \mu)$ and we have

$$\beta_L \cdot x_1(\pi_H; r, \mu) > \boldsymbol{CE}_\varrho\{x_2(\pi_H, \varrho; r, \mu)|\sigma\}.$$

All investors will withdraw their deposits at $t = 1$, which results in a run on the representative bank. This banking panic is inevitable given the effective payouts of the deposit contract. It is a fundamental run on the representative bank and, in the terminology of Freixas and Rochet (2008), reflects a fundamental banking panic.

Secondly, consider deposit contracts $(r, \mu)$ such that after receiving signal $\sigma$ and for assumed withdrawals $\varpi = \pi_H$ the incentive constrains $(IC_H^\sigma)$ and $(IC_L^\sigma)$ are satisfied for the effective payouts $x_1$ and $x_2$. In one of the withdrawal equilibria only type−$H$ investors withdraw their deposits; another withdrawal equilibrium is a banking panic, which is disregarded.

Finally, consider deposit contracts $(r, \mu)$ such that after receiving signal $\sigma$ the incentive compatibility constraint of $H$−type investors for the effective payouts is violated for $\varpi = \pi_H$ i.e.

$$\beta_H \cdot x_1(\pi_H; r, \mu) < \boldsymbol{CE}_\varrho\{x_2(\pi_H, \varrho; r, \mu)|\sigma\}.$$

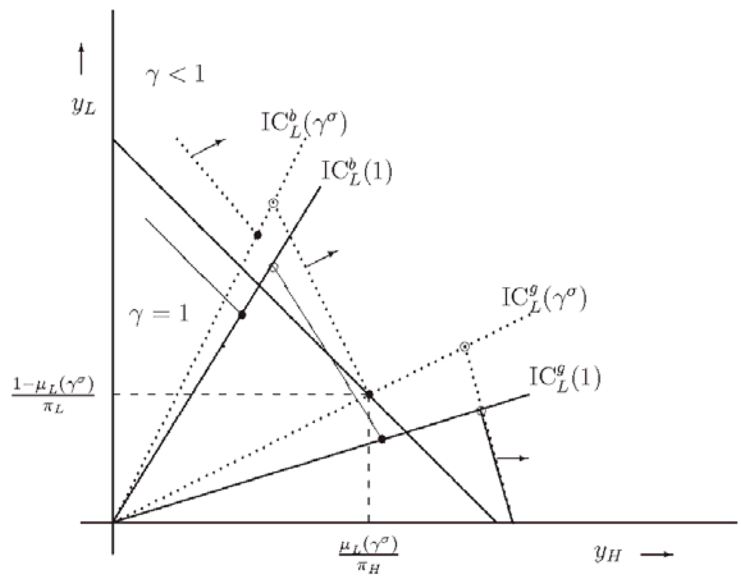
In equilibrium some of the affected investors will now defer withdrawing until $t = 2$. As a consequence, the representative bank must transfer some of its reserve holdings from $t = 1$ to $t = 2$, which reduces the expected effective payouts at $t = 2$ for $\varrho = h$ to a value below $x_2(\pi_H, h; r, \mu)$. As before, a second withdrawal equilibrium is a banking panic, which is disregarded unless $x_1(1; r, \mu) > \boldsymbol{CE}_\varrho\{x_2(1, \varrho; r, \mu)|\sigma\}$.

### 4.4. Ex-Ante Reserve Holdings

Now we have determined the relevant withdrawal equilibria, we consider the fractional reserve holdings that solve the decision problem of the representative bank. This requires a separate analysis of each of the above three possibilities for withdrawal equilibrium for the signals $b$ and $g$.

The decision problem of the representative bank is illustrated in Figure 3. The feasibility line and the ad interim incentive constraints are the same as in Figure 1.

Figure 3: The Decision Problem for the Representative Bank




In Figure 3, the indifference curves of the representative bank are those of the indirect ex-ante utility function $V^B$, which is obtained as follows. For an equilibrium in the competitive banking sector, the efficient payout condition must be satisfied, which requires $r_1 = \frac{\mu}{\pi_H}$ and $r_2 = \alpha_h \cdot \frac{1-\mu}{\pi_L}$. Now, in specifying the actions of the representative bank for 'non-equilibrium' reserve holdings, the promised payouts are assumed to remain $r_1 = \frac{\mu}{\pi_H}$ and $r_2 = \alpha_h \cdot \frac{1-\mu}{\pi_L}$, even if they lead to banking panics that might have been prevented by different promised repayments. The alternative assumption that the banks adapt their promised repayment would lead to a different indirect utility function for this institutional setting, based on different out-of-equilibrium behavior.

The contracts for non-equilibrium reserve holdings as assumed here lead to indifference curves of $V^B$ as in Figure 3. To the right of the $(IC_L^g)$-curve, a fundamental banking panic occurs after each signal $\sigma \in \{b, g\}$, which leads to steep indifference curves. On the $(IC_L^g)$-curve there will be no fundamental banking panic after a good signal, which causes the indifference curve to 'jump inward'.

Between the $(IC_L^g)$-curve and the $(IC_L^b)$-curve, there will be a banking panic after a bad signal but not after a good signal. Therefore, the indifference curves in this area are less steep than the ones to the right of the $(IC_L^g)$-curve. On and to the left of the $(IC_L^b)$-curve there will be no banking panic after either signal. Therefore, on the $(IC_L^b)$-curve the indifference curves once again ‘jump inward’, with indifference curves being flatter to the left of the $(IC_L^b)$-curve.

In Figure 3, the case is depicted in which the $(IC_H^g)$-curve is to the left of the $(IC_L^b)$-curve, without the $(IC_H^g)$-curve being included in the figure. For the solution of the decision problem, the area to the left of the $(IC_H^g)$ is not relevant.

The case in which the $(IC_H^g)$-curve is to the right of the $(IC_L^b)$-curve is not depicted. The precise shape of the indifference curves in the area between the $(IC_H^g)$-curve and the $(IC_L^b)$-curve is not relevant for the solution of the representative bank’s decision problem.

The following proposition results.

**Proposition 3** *For the equilibrium fractional reserve holdings we have*

$$\mu^B(\gamma) \in \{\mu_b(\gamma), \mu_g(\gamma), 1\}.$$

As can be seen in Figure 3, a reduction in the level of confidence affects the indifference curves for the representative bank in two ways. Firstly, the ad interim incentive constraints are rotated counter-clockwise. Secondly, the indifference curves become steeper. A reduction in the level of confidence makes investment in the illiquid asset less attractive, so a larger amount of the illiquid asset is needed to satisfy the relevant incentive constraint, leading to a reduction in reserve holdings.

Similar as in Corollary 2 for second best efficient money holdings, a ‘small’ reduction in the level of confidence leads to a decrease in the representative bank’s equilibrium reserve holdings and to an increase in investments in the illiquid asset. This is formalized in the following corollary.

**Corollary 4** *Let* $\gamma', \gamma'' \in (\hat{\gamma}^E, 1]$ *be such that for some* $\sigma \in \{b, g\}$ *we have both* $\mu^B(\gamma') = \mu_\sigma(\gamma')$ *and* $\mu^B(\gamma'') = \mu_\sigma(\gamma'')$. *Now*

$$\gamma' < \gamma'' \Leftrightarrow \mu^B(\gamma') < \mu^B(\gamma'').$$

## 5. Comparison

The analysis in Section 3 and Section 4 enables us to address the question whether the outcome of the financial sector, as obtained by the representative bank, is second best efficient. In Section 5.1 we find that the outcome fails to be second best efficient, which implies that the counter part of the ‘invisible hand’ fails to hold in the setting of financial crises, even when making allowances for the information constraints. In Section 5.2 we consider the question which policy measures which may restore second best efficiency. We find that minimum reserve requirements are ineffective, but that the ex-ante commitment to an appropriate revenue neutral tax-and-subsidize policy results in a second best efficient outcome.

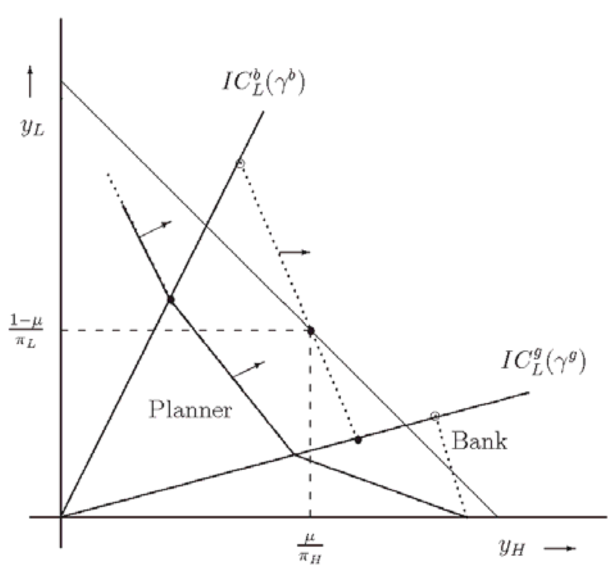
### 5.1. Inefficient Outcomes

When comparing the solutions for the decision problems of the social planner and the representative bank, the graphical analysis outlined in Section 2.5 proves useful. It enables the comparison of the two ex-ante decision problems of Figure 1 and Figure 3 in one single diagram, facilitating the analysis.

Figure 4 compares the decision problem of the representative bank with that of the social planner. To the right of the $(IC_L^b)$-curve, the ex-ante utility derived from aggregate fractional reserve holdings $\mu$ through the competitive banking sector is less than the ex-ante utility achieved by the social planner for the same fractional money holdings. The reason is that the competitive banking sector has less possibilities for ad interim redistribution between the two types of investors.

In his choice of contracts, the planner is only restricted by feasibility and incentive compatibility constraints. When the incentive constraint for $L-$type investors is violated after signal $\sigma$, the more extensive possibilities for redistribution available to the planner

Figure 4: Comparison of Bank and Social Planner for $\gamma < 1$.



come to bear. They result in a better liquidity allocation than the competitive banking sector can provide. The deposit contracts offered in the competitive banking sector do not give banks the opportunity to restore incentive compatibility by allowing $L-$type investors to withdraw some of their deposits at $t = 1$ without reducing their effective payouts at $t = 2$. As a consequence, a fundamental banking panic and the resulting loss of ex-ante utility are inevitable if $(IC_L^g)$ is violated.

Things are different if $(IC_H^g)$ is violated. In the case of 'small' violations of the incentive compatibility constraint of $H-$type investors, incentive compatibility can be restored by some of them deferring their withdrawal to $t = 2$. Some of the money holdings at $t = 1$ must be transferred to $t = 2$. This leads to a reduction in the expected effective payouts at $t = 2$, which may restore incentive compatibility. In Figure 4, the indifference curves of the indirect ex-ante utility function $V^B$ would have a strictly convex shape where this applies.

The violation of the incentive constraint for investors with a high liquidity preference may, however, be such that the incentive constraint remains violated even if all type$-H$ investors withdraw at $t = 2$. In the area of Figure 4 in which this arises, the indifference curves for the representative bank are linear.

As is indicated in Theorem 1, under the assumptions of Section 2.6, the second best efficient aggregate fractional money holding is either $\mu_g(\gamma)$ or 1. Under these assumptions equilibria in the competitive banking sector with a fractional reserve holding of $\mu_b(\gamma)$ fail to be second best efficient. But even if the fractional reserve holdings in the banking sector are $\mu_g(\gamma)$, the second best efficient liquidity allocation fails to be obtained. When a bad signal regarding the prospects of the asset is received, a fundamental banking panic occurs and the investments in the illiquid asset will be liquidated. By contrast, a social planner would restore incentive compatibility by allocating part of the money holdings at $t = 1$ to type$-L$ investors without reducing their payouts at $t = 2$.

This leads to the following result.

**Theorem 5** *Under the assumptions of Section 2.6, no deposit contract* $(r, \mu)$ *with* $\mu < 1$ *implements the second best efficient liquidity allocation.*

### 5.2. Restoring Second Best Efficiency

In a competitive banking sector representing the financial sector as above with $\gamma \in (\hat{\gamma}^E, 1]$, as in Theorem 5, there are two possible reasons why the liquidity allocation may fail to be second-best efficient. Firstly, the decision problem of the representative bank may be solved for the inefficient fractional reserve holdings $\mu_b(\gamma)$, rather than for the second best efficient reserve holdings $\mu_g(\gamma)$. But, secondly, even if the reserve holdings are at the second best efficient level $\mu_g(\gamma)$, there will be a fundamental banking panic — i.e. an unavoidable financial crisis — after the bad signal is received. This financial crisis leads to the inefficient liquidation of illiquid assets.

A regulator can solve this second problem by restoring the incentive compatibility constraint of type$-L$ investors after a bad signal. One way of doing this is by a revenue neutral combination of taxes and subsidies at $t = 1$. Investors who withdraw their deposits at $t = 1$ pay a tax, which is used at $t = 1$ to subsidize investors who do not withdraw. For an appropriate combination of tax and subsidies, the second-best liquidity allocation is obtained.

The next question is how a regulator can address the potential problem of inefficient money holdings $\mu_b(\gamma)$. In the absence of regulation dealing with financial crises, minimum reserve requirements fail to improve the liquidity allocation, as the equilibrium reserve holdings $\mu_b(\gamma)$ are lower than $\mu_g(\gamma)$, but they lead to a higher ex-ante utility; if this would not be the case, the representative bank would have chosen $\mu_g(\gamma)$ rather than $\mu_b(\gamma)$.

There is, however, a more subtle solution to the problem. The regulator can credibly announce to implement a policy of taxes and subsidies that ensures second-best efficient contingent payouts whenever a bad signal is received for money holdings $\mu_g(\gamma)$. This policy is now anticipated by the financial sector and by the investors, whose decisions are based on the suitably adapted effective payouts. As a result, competition within the financial sector leads to the deposit contract $\left(\frac{\mu_g(\gamma)}{\pi_H}, \mu_g(\gamma)\right)$ which now implements the second best efficient liquidity allocation.

This leads to the following result.

**Theorem 6** *Under the assumptions of Section 2.6, a second best efficient outcome can be implemented by the regulator's commitment to a revenue neutral tax-and-subsidize policy which ensures ad interim efficient payouts given the financial sector's money holdings.*

At first glance, it may seem that the credible announcement of the 'bail out policy' by the regulator inappropriately creates a moral hazard problem. The policy 'tempts' the financial sector to 'carelessly' risk a financial crisis after a bad signal, rather than 'prudently' ensuring no fundamental crisis can occur. 'Greedy' investors are but too willing to play along, as they can rely on the regulator to 'bail them out' if and when a fundamental financial crisis occurs.

This impression, however, is misleading. In an unregulated financial sector, crises lead to the liquidation of assets, which can be very costly. Both the financial sector and investors may consider inefficiently low reserve holdings that prevent crises to be the 'lesser evil'.

But in the presence of a credibly announced 'bail out policy' a (potential) financial crisis no longer leads to the costly and inefficient liquidation of assets. Therefore, the financial sector no longer needs to choose the 'lesser evil' of inefficiently low reserve holdings. It can increase reserve holdings to the level that is required for the ex-ante second best efficient redistribution at of wealth from investors with a low preference for liquidity to investors with a high liquidity preference.

The 'bail out policy' of the regulator does not inappropriately create a moral hazard problem, but rather provides an antidote to the inefficiencies created by the deposit contract's prioritizing of withdrawals at $t = 1$ over those at $t = 2$.

## 6. Failure to Recognize Ambiguity

The failure of the financial sector to recognize the presence of ambiguity in investors' beliefs can have severe consequences. Analyzing these consequences is the purpose of this section. Strictly speaking, the question is outside the framework of the model and a thorough analysis would require a more general formal framework.

The construction of the imagined representative bank is based on the combination of competition in the banking sector and perfect foresight by the investors. Under these assumptions, the failure of the representative bank to recognize the presence of ambiguity implies the failure of investors themselves to recognize presence of ambiguity. As long as investors have perfect foresight and recognize the presence of ambiguity in the economy, competition forces the banks to act accordingly, even if they do not agree. Therefore, the failure of the financial sector to recognize the presence of ambiguity to some extent violates the internal logic of the model.

This being said, the issue itself seems to be of practical relevance and deserves consideration in the context of this paper. The assumptions of the model — including the representation of the financial sector by a competitive banking sector – are abstractions from reality. So even if some relevant issues are not fully covered by a model's internal logic, the model can still shed some light on them. In particular, by offering a logically consistent benchmark, it can provide a starting point for economic thinking that surpasses the model's boundaries. This approach to understanding economic issues is the basis for the following analysis.

We start by assuming that the effective payouts $(x_1, x_2)$ implied by the deposit contract obtained in the competitive banking sector are exogenously guaranteed. Thus, the beliefs of the investors do not correct the mistaken beliefs of the banks through the competitive process.

We consider two ways in which the financial sector may fail to recognize the presence of ambiguity. The first way basically assumes that the financial sector is ambiguity neutral. It ignores the presence of ambiguity and treats the probability estimate $\pi$ as if it was a (subjective) probability distribution held with full confidence. Alternatively, the financial sector may treat the weighting of the states of nature in the investors' ex-ante utility function as a subjective probability estimate. Due to the presence of ambiguity, this implied weighting differs from the probability estimate $\pi$.

When the financial sector shares the investors' probability assessment but wrongly assumes a level of confidence $\gamma = 1$, it offers deposit contracts $(r, m) = \left(\frac{\mu_\sigma(1)}{\pi_H}, \mu_\sigma(1)\right)$. Since $\mu_\sigma(1) > \mu_\sigma(\gamma)$ the actual incentive constraint of type$-L$ investors will unexpectedly be violated when the signal $\sigma$ occurs. The financial sector will be confronted with a (fundamental!) crisis it is unable to explain.

If a regulator is committed to the policy of taxes and subsidies as in Theorem 6, the aggregate money holdings would be $\mu_g(1) > \mu_g(\gamma)$. After either signal, a crisis would be imminent and regulatory intervention would be required.

The second, more sophisticated, way in which the financial sector may fail to recognize the presence of ambiguity is by mis-interpreting the weights of the states of nature in the investors' ex-ante utility function as a (subjective) probability distribution $\hat{\pi}$. In general settings with ambiguity, this will not lead to a unique probability distribution. The weights change when different payout profiles are considered that fail to be co-monotonic, e.g. profiles that obtain their worst case in different states of nature. But for the specific payouts that occur in a competitive banking sector, the weights *are* consistent with a unique (subjective) probability distribution $\hat{\pi}$.

Consider a representative bank with reserve holdings $\hat{\mu}_g(1)$ based on $\hat{\pi}$. For the implied effective payouts we have $u(x_H(g, \varrho); H) > u(x_H(b, \varrho); H)$ and $u(x_L(g, h); L) > u(x_L(b, \varrho); L) > u(x_L(g, \ell); L)$. In the ex-ante utility function of the investors, this leads to the following weighting of the states $(\theta, \sigma, \varrho)$:

<table>
<tr><td></td><td>$(b, h)$</td><td>$(b, \ell)$</td><td>$(g, h)$</td><td>$(g, \ell)$</td></tr>
<tr><td>$H$</td><td colspan="2">$\pi_H \cdot (1 - \gamma + \gamma \cdot (\delta + \varepsilon))$</td><td colspan="2">$\pi_H \cdot \gamma \cdot (1 - (\delta + \varepsilon))$</td></tr>
<tr><td>$L$</td><td>$\pi_L \cdot \gamma \cdot \delta$</td><td>$\pi_L \cdot \gamma \cdot \varepsilon$</td><td>$\pi_L \cdot \gamma \cdot (1 - (\delta + \varepsilon))$</td><td>$\pi_L \cdot (1 - \gamma)$</td></tr>
</table>

These weights are consistent with one of the weightings obtained for reserve holdings $\hat{\mu}_b(1)$ and $\mu = 1$. So the financial sector's assumption that it is dealing with investors whose preferences can be represented by a subjective expected utility function is not obviously proven wrong.

Still, when interpreted as a probability distribution, these weights imply that the probability of receiving a bad signal conditional on having a high liquidity preference, $\theta = H$, differs from the probability of receiving a bad signal conditional on having a low liquidity preference, $\theta = L$. This 'pessimistic superstition' of the investors may raise suspicion in the financial sector, since it is not justified by the basic structure of liquidity preference and asset returns.

This superstition turns into a more serious problem when investors' ex-ante beliefs are updated at $t = 1$. For this updating the Dempster-Shafer rule is used, which leads to the following ad interim weights of the states of nature for an investor of type $L$, taking into account the effective payouts as above:

| | $(b,h)\|b$ | $(b,\ell)\|b$ | $(g,h)\|g$ | $(g,\ell)\|g$ |
|---|---|---|---|---|
| $L$ | $\frac{\gamma\cdot\delta}{\gamma\cdot(\delta+\varepsilon)+(1-\gamma)}$ | $\frac{\gamma\cdot\varepsilon+(1-\gamma)}{\gamma\cdot(\delta+\varepsilon)+(1-\gamma)}$ | $\frac{\gamma\cdot(1-(\delta+\varepsilon))}{1-\gamma\cdot(\delta+\varepsilon)}$ | $\frac{1-\gamma}{1-\gamma\cdot(\delta+\varepsilon)}$ |

But being unaware of the presence of ambiguity, the financial sector updates the probability distribution $\hat{\pi}$ by using Bayes' rule. Regarding a type$-L$ investor this leads to the updated probability distribution

| | $(b,h)\|b$ | $(b,\ell)\|b$ | $(g,h)\|g$ | $(g,\ell)\|g$ |
|---|---|---|---|---|
| $L$ | $\frac{\delta}{\delta+\varepsilon}$ | $\frac{\varepsilon}{\delta+\varepsilon}$ | $\frac{\gamma\cdot(1-(\delta+\varepsilon))}{1-\gamma\cdot(\delta+\varepsilon)}$ | $\frac{1-\gamma}{1-\gamma\cdot(\delta+\varepsilon)}$ |

For aggregate fractional reserve holdings $\hat{\mu}_g(1)$ based on $\hat{\pi}$, the financial sector anticipates a fundamental crisis after a bad signal. After a bad signal, this fundamental crisis occurs on the basis of the investors' actual beliefs. After a good signal, the weights assigned by the update of $\hat{\pi}$ coincide with the weights that result for the ambiguous beliefs. The actual incentive constraint of $L-$types after a good signal equals the incentive constraint as anticipated by the financial sector. We have $\hat{\mu}_g(1) = \mu_g(\gamma)$ and it remains unnoticed that the financial sector is unaware of the presence of ambiguity.

For money holdings $\hat{\mu}_b(1)$ however, things are different. For the effective payouts of the associated deposit contract we now have $x_L(g,h) = x_L(b,h) > x_L(g,\ell) = x_L(b,\ell)$, leading to the ex-ante weights

| | $\{(b,h),(g,h)\}$ | $\{(b,\ell),(g,\ell)\}$ |
|---|---|---|
| $H$ | $\pi_H$ | |
| $L$ | $\pi_L \cdot \gamma \cdot (1-\varepsilon)$ | $\pi_L \cdot (1-\gamma) \cdot (1-\varepsilon)$ |

which are consistent with the weights obtained for $\hat{\mu}_g(1)$.

After the bad signal, the financial sector finds that type$-L$ investors put an unexpectedly high weight on the state $(b,\ell)$. The investors apply the Dempster-Shafer rule to update their beliefs, resulting in an implied weight of $\frac{\gamma\cdot\varepsilon+(1-\gamma)}{\gamma\cdot(\delta+\varepsilon)+(1-\gamma)}$, whereas the financial sector anticipated the lower weight of $\frac{\varepsilon}{\delta+\varepsilon}$. Therefore, the ad interim incentive constraint of type$-L$ investors is violated and a fundamental crisis occurs.

This fundamental crisis takes the financial sector by surprise. It assumed that for its reserve holdings the incentive compatibility constraint of $L-$types would be satisfied even after a bad signal. From the perspective of the financial sector, the crisis is caused by an 'irrational overreaction' of the investors with a low liquidity preference, who 'lose their nerves' in the face of bad news. In reality, the financial sector is facing a fundamental crisis, caused by its own failure to recognize the presence of ambiguity in investors' beliefs.

In an economy where the financial sector fails to recognize ambiguity in investors' beliefs, two types of regulatory policy come to mind. The first is the policy announcement suggested in Theorem 6 with levels of taxes and subsidies that reflect the presence of ambiguity. This leads to the second best efficient liquidity allocation even if the financial sector misinterprets the weightings in the ex-ante utility functions as probabilities.

The second type of regulatory policy surpasses the financial sector. It aims at creating an institutional framework which enhances the level of confidence investors have in their probability estimates. This requires measures that increase the stability and transparency of the economy as a whole, including the predictability of (competent) economic policy.

## 7. Concluding Remarks

### 7.1. Dynamic Inconsistency

The presence of ambiguity leads to dynamic inconsistency in investors' beliefs. After receiving non- conclusive information, decision makers tend to deviate from the contingent course of action they initially planned.

This raises the question whether dynamic inconsistency should be regarded as an inherent property of decision making under ambiguity or, alternatively, as an undesirable artefact of the mathematical model used for its representation.

In the context of decisions made in the presence of ambiguity, Keynes (1937, pp. 114) states:

> 'How do we manage in such circumstances to behave in a manner which saves our faces as rational, economic man? We have devised for the purpose a variety of techniques, of which the most important are the three following:
>
> (1) We assume that the present is a much more servicable guide to the future than a candid examination of past experience would show it to have been hitherto. In other words, we largely ignore the prospect of future changes about the actual character of which we know nothing.
>
> (2) We assume that the existing state of opinion as expressed in prices and the character of existing output is based on a correct summing up of future prospects, so that we can accept it as such unless and until something new and relevant comes into the picture.
>
> (3) Knowing that our own individual judgment is worthless, we endeavor to fall back on the judgment of the rest of the world which is perhaps better informed. That is, we endeavor to conform with the behavior of the majority or the average. The psychology of a society of individuals each of whom is endeavoring to copy others leads to what we may strictly term a conventional judgment.'

This quotation — in particular the combination of points (1) and (2) — can be taken as an indication that Keynes considers dynamic inconsistency to be an inherent behavioral aspect of decision making under ambiguity. Keynes continues (Keynes, 1937, pp. 114 - 115):

> 'Now a practical theory of the future [...] has certain marked characteristics. In particular, being based on so flimsy a foundation, it is subject to sudden and violent changes. The practice of calmness and immobility, of certainty and security, suddenly breaks down. New fears and hopes will, without warning, take charge of human conduct. The forces of disillusion may suddenly impose a new conventional basis of valuation. All the pretty, polite techniques, made for a well-pannelled board room and a nicely regulated market, are liable to collapse. At all times vague panic fears and equally vague and unreasoned hopes are not really lulled and lie but a little way below the surface.'

which provides further support for this view. From this perspective, approaches that find conditions under which dynamic inconsistency in updating beliefs under ambiguity fails to occur, would be interpreted as identifying exceptional cases in which the basic mechanism fails to apply. For a discussion of these issues see e.g. Epstein and Le Breton (1993), Hanany and Klibanoff (2009) and Siniscalchi (2011).

### 7.2. Policy Implications

The model suggests that the second best efficient liquidity allocation will be obtained by a financial sector in which a bail out policy is credibly announced and implemented as in Theorem 6. In the case of an imminent crisis, where the violation of the relevant ad interim incentive constraint may have created the equivalent of a 'Minsky moment', the policy taxes withdrawals and uses the proceeds to subsidize those who do not withdraw their deposits at $t = 1$. Being a revenue neutral redistribution of resources at $t = 1$, it has the advantage of not requiring the provision of additional funding or liquidity.

In reality, however, implementing such a policy on short notice may fail to be feasible. It may be more promising to combat an imminent crisis by use of more flexible instruments. These could include the provision of additional liquidity by central banks through lowering interest rates and by quantitative easing. They could also include the underwriting of 'toxic assets' and the bailing out of distressed banks by governments.

The underwriting of 'toxic assets' and bailing out distressed banks reduces the ambiguity regarding banks' ability to honor future withdrawals, increasing the (Choquet) expected payout of these withdrawals. In the context of our model, this can be interpreted as a subsidy to investors who leave their money in the bank. It helps to restore the incentive compatibility for the investors with a low liquidity preference.

Similarly, flooding the financial markets with liquidity allows more investors to take advantage of idiosyncratic favorable opportunities, potentially reducing the value of such opportunities to the investors with a low liquidity preference and those with a high liquidity preference alike. In the model, this is would be reflected by either a decrease of $\beta_L$ and $\beta_H$, or by a tax on withdrawals at $t = 1$. By making withdrawals at $t = 1$ less attractive, flooding the market with liquidity helps restoring the incentive compatibility for type$-L$ investors.

Therefore, the combination of flooding the market with liquidity on the one hand and underwriting toxic assets and bailing out distressed banks on the other, effectively implements the proposed policy of taxes and subsidies by using the instruments readily available to policy makers.

A second type of regulatory policy would focus on reducing the level of ambiguity faced by the investors, i.e. by increasing the level of confidence. This requires a broader based approach which emphasizes stability, transparency and predictability of economic policy measures. It resembles the kind of institutional framework advocated by the 'ordo-liberalism' of the 'Freiburger Schule'. This approach is at the heart of the 'soziale Marktwirtschaft', the basis of the 'Wirtschaftswunder' — the German economic miracle in the 1950s and 1960s — and much of EU economic policy thereafter. For a more general discussion on this and related issues see Spanjers and Agliardi (2016).

## References

Allen, F. and D. Gale (1998), Optimal Financial Crises, *Journal of Finance*, 53, 1245-1284.

Agliardi, E., R. Agliardi and W. Spanjers (2015), 'Convertible Debt: Financing Decisions and Voluntary Conversion under Ambiguity', *International Review of Finance*, 15, 599-611.

Agliardi, E., R. Agliardi and W. Spanjers (2016), Corporate Financing Decisions under Ambiguity: Pecking Order and Liquidity Policy Implications, *Journal of Business Research*, 69, 6012-6020.

Brach, J. and W. Spanjers (2012), Political Ambiguity and Economic Development: The Cases of the MENA Countries, Working Paper 66-12, Rimini Centre for Economic Analysis.

Caballero, R., E. Farhi and M. Hammar (2006), Speculative Growth: Hints from the US Economy, *American Economic Review*, 94, 1159-1192.

Chateauneuf, A., J. Eichberger and S. Grant (2007), Choice under Uncertainty with the Best and the Worst in Mind: Neo-Additive Capacities, *Journal of Economic Theory*, 137, 538-567.

Dempster, A. (1968), A Generalization of Bayesian Inference, *Journal of the Royal Statistical Society, Series B*, 30, 205-247.

Dow, J. and S. Werlang (1992), Uncertainy Aversion, Risk Aversion and the Optimal Choice of Portfolio, *Econometrica*, 60, 197-205.

Diamond, D. and P. Dybvig (1983), Bank Runs, Deposit Insurance and Liquidity, *Journal of Political Economy*, 91, 401-419.

Eichberger, J. and I. Harper (1997), *Financial Economics*. Oxford: Oxford University Press.

Eichberger, J. and D. Kelsey (1999), E-Capacities and the Ellsberg Paradox, *Theory and Decision*, 46, 107-138.

Eichberger, J. and W. Spanjers (2007), Liquidity under Uncertainty: Banks or Asset Markets?, Discussion Paper 07-18, SFB504, University of Mannheim.

Epstein, L. and M. Le Breton (1993), Dynamically Consistent Beliefs must be Bayesian, *Journal of Economic Theory*, 61, 1-22.

Epstein, L. and T. Wang (1994), Intertemporal Asset Pricing under Knightian Uncertainty, *Econometrica*, 62, 283-322.

Ellsberg, D. (1961), Risk, Ambiguity and the Savage Axioms, *Quarterly Journal of Economics*, 75, 643-669.

Freixas, X. and J. Rochet (2008), *Microeconomics of Banking*, Second Edition. Cambridge MA: MIT Press.

Gilboa, I. (1987), Expected Utility with Purely Subjective Non-additive Probabilities, *Journal of Mathematical Economics*, 16, 65-88.

Gilboa, I. and D. Schmeidler (1993), Updating Ambiguous Beliefs, *Journal of Economic Theory*, 59, 33-49.

Ghatak, S. and W. Spanjers (2007), Monetary Policy Rules in Transition Economies: The Impact of Ambiguity, *International Journal of Development Issues*, 6, 26-37.

Hanany, E. and P. Klibanoff (2009), Updating Ambiguity Averse Preferences, *BE Journal of Theoretical Economics*, 9, Article 37.

Hansen, L. and T. Sargent (2001), Robust Control and Model Uncertainty, *American Economic Review, Proceedings*, 91, 60-66.

Hayek, F. (1935), *Production and Prices, 2nd Edition, in J. Salerno (ed.), Prices and Production and Other Works*. Auburn AL: Ludwig von Mises Institute.

Hellwig, M. (2009), Systemic Risk in the Financial Sector: An Analysis of the Subprime-Mortgage Financial Crisis, *De Economist*, 157, 129-207.

Holmström, B. and J. Tirole (1997), Financial Intermediation, Loanable Funds, and the Real Sector, *Quarterly Journal of Economics*, 112, 663-691.

Jacklin, C. and S. Bhattacharya (1988), Distinguishing Panics and Information-based Bank Runs: Welfare and Policy Implications, *Journal of Political Economy*, 96, 568-592.

Jaffray, J. (1992), Bayesian Updating and Belief Functions, *IEEE Transactions on Systems, Man and Cybernetics*, 22, 1144-1152.

Keynes, J. (1937), *The General Theory of Employment, from: Quarterly Journal of Economics*, Collected Writings, vol. 14. London: McMillan, 109-123.

Knight, F. (1921), *Risk, Uncertainty and Profit*. London: Houghton Mifflin.

Olivier, J. (2000), Growth Enhancing Bubbles, *International Economic Review*, 41, 133- 151.

Rancière, R., A. Tornell, and F. Westermann (2008), Systemic Crises and Growth, *Quarterly Journal of Economics*, 123, 359-406.

Schmeidler D. (1982/89), Subjective Probability and Expected Utility without Additivity, *Econometrica*, 57, 571-587.

Shafer, G. (1976), *A Mathematical Theory of Evidence*. Princeton NJ: Princeton University Press.

Siniscalchi, M. (2011), Dynamic Choice under Ambiguity, *Theoretical Economics*, 6, 379-421.

Spanjers, W. (1999/2008), Liquiditätversorgung bei Knight'scher Unsicherheit: Ein theoretischer Vergleich von Banken und Märkten, (Liquidity Provision under Ambiguity: A Theoretical Comparison of Banks and Markets), Habilitation Thesis, Saarland University. Leipzig: Engelsdorfer Verlag.

Spanjers (2008a), Loss of Confidence and Currency Crises, *International Journal of Economic Research*, 5, 219-237.

Spanjers, W. (2008b), Central Banks and Ambiguity, *International Review of Economics and Finance*, 17, 85-102.

Spanjers, W. (2009), The Asian Crisis and Macroeconomic Development: The Impact of Ambiguity, in S. Ghatak and P. Levine (eds), *Development Macroeconomics: Essay in Memory of Anita Ghatak*, Oxford: Routledge, 68-94.

Spanjers, W. and E. Agliardi (2016), Rethinking the Social Market Economy: A Basic Outline, Professional Report 16-01, Rimini Centre for Economic Analysis.

Wagner, H. (2007), Information and Uncertainty in the Theory of Monetary Policy, SUERF Studies 2007/1, SUERF.