Funding-Shortfall Risk and Asset Prices in General Equilibrium*

Majid Hasan†

October 26, 2016

*I would like to thank my PhD advisor, Raman Uppal, Jérôme Detemple (AsianFA discussant), David Feldman (FMA Asia discussant), Hassan Naqvi (NUS RMC discussant), and Jan Ericsson (NFA Discussant) for their comments. Helpful comments were also received from participants at EDHEC PhD Seminar, FMA Asia Conference, and NUS Risk Management Conference. All remaining errors are mine.

†EDHEC Infrastructure Institute, Singapore, and EDHEC Business School, Nice. Email: `majid.hasan@edhec.edu`. Tel.: `0065-6653 8585`. Address: `07-02, 01 George Street, Singapore, 049145`.

Abstract

Institutional investors, such as pensions and insurers, are typically constrained to hold enough wealth to be able to make their contractually promised payments to fund beneficiaries, and face a funding-shortfall risk. We seek to explore the optimal asset allocation strategies for institutions facing this risk, and its effects on asset prices. The constraint introduces two distinct regions in the economy, unconstrained and constrained, with the possibility of transitioning from the constrained to the unconstrained region, which leads to a two-factor asset pricing model. The funding-shortfall risk increases the conditional equity premium and Sharpe ratio, which evolve counter-cyclically, but decreases the conditional volatility of equity returns, which evolves cyclically. The constrained institution may optimally hold an under-diversified portfolio, and simultaneously increases its demand for the risk-free and higher-risk assets relative to medium-risk assets, inducing a bubble-like behaviour in the prices of higher-risk assets. The dynamics of required payouts can induce predictability in the dynamics of conditional moments of asset returns. The term structure of interest rates is predominantly upward sloping, but can change shape upon shocks to the growth rate of aggregate dividend relative to the growth rate of minimum payouts, providing a new channel through which the business cycle may affect the shape of the term structure of riskfree rates.

JEL Classification: G11, G12, G18, G21, G23, G28

Keywords: institutional investors, capital requirements, funding-shortfall risk, under-diversification, regime-switching.

1 Introduction

The primary goal of our paper is to develop an understanding of asset pricing and asset allocation implications of contractually promised payouts (constrained consumption), funding-shortfall risk, and tension between institutional preferences and the funding-ratio constraint in a general equilibrium setting. This paper makes contributions to the literature on consumption based asset pricing, portfolio insurance, and effects of institutional investors on asset prices. Its contribution to the literature on consumption based asset pricing is to study the improvement in the ability of Consumption Capital Asset Pricing Model (CCAPM) to explain asset pricing regularities, with the addition of a funding-ratio constraint. Its contribution to the portfolio insurance literature is to allow the insurance level to be set as the present value of a stream of future, and possibly time-varying, payouts (consumption). And its contribution to the literature on institutional investors' effects on asset-prices is to study the general equilibrium implications of institutions that have a contractually promised stream of payments, and face a tension between their optimal unconstrained decisions, and the demands of an externally imposed funding-ratio constraint.

Our focus on funding-ratio constraint is motivated by the growing importance of institutional investors, such as pensions, insurers, and sovereign wealth funds, which hold an increasingly larger share of traded assets (Friedman (1995), Gompers and Metrick (2001)). These financial institutions are corporation-like entities constrained by their contractual relations with different stakeholders including end-investors, regulators, employees etc. One of the most salient of these institutional constraints is their contractually promised payments to end-investors, which often constitute the largest share of institutional use of funds and rank senior to all other institutional payouts (uses of institutional funds), implying that a failure to meet these promised payments will leave an institution insolvent. This is most obvious in the case of defined-benefit pension funds and insurers, who are legally obliged to pay pension and insurance benefits. Endowments and sovereign wealth funds may not have ex-ante contractually promised payments, but nevertheless may have an ex-ante minimum-payout stream, which may be determined by their minimum

expenditure needs, that they may wish to maintain. Similarly, in the case of banks and mutual funds, expected withdrawals in each period can also be viewed as the minimum level of payouts that these institutions may wish to maintain. We refer to this minimum level of payouts (or withdrawal) as minimum-payouts and minimum-withdrawals interchangeably.

In order to insure that these minimum-payments can be made, these institutions often manage their funds such that the present value of their wealth exceeds the present value of these minimum-payouts. For instance, pension funds and insurers are often required by regulation to maintain a minimum amount of wealth that does not fall below a certain fraction, which typically ranges from 0.8 to 1.1, of the present value of all future payments promised to end-investors (Blome, Fachiner, Franzen, and Scheuenstuhl (2007)). Similarly, according to Basel III capital requirements (BIS (2011)), banks are required to invest equity capital of at least 10% of their total liabilities (deposits), which translates to a funding-ratio constraint that requires invested wealth to be no less than 110% of the total value of their deposits.

The recent financial crisis has left many institutions underfunded, and the funding-ratio requirements for such institutions and their effects on asset prices have been a topic of debate. A demand for higher capital requirements, in particular for banks, is often made on the grounds that raising capital requirements will increase banks' loss-bearing capacity, reducing the probability that the depositors will have to suffer any losses, and, hence, decreasing the need for bailouts. Carney (2016) points out that the capital requirements for banks in the UK have been considerably increased since the financial crises.

However, while higher capital requirements will increase institutions' loss-bearing capacity, these requirements will also have indirect effects on the economy through equilibrium prices of riskfree and risky assets, which may or may not be desirable. For instance, after the UK Pensions Act of 2004 stipulated that pension funds had to maintain a minimum ratio of assets to liabilities, the UK term structure went from being almost flat to downward sloping at the longer maturity end, due to higher demand by pension funds for longer maturity bonds triggered by the

newly introduced funding-ratio constraint (Gromb and Vayanos (2010)). Haldane (2014) highlights the need to understand the role of risk-based capital requirements in inducing pro-cyclical tendencies in asset allocation decisions of pensions and insurers, and amplifying market volatility. Our model allows us to explore these indirect effects that may be relevant for policy decisions.

The main results of our paper can be summarised as follows. The presence of institutional investors can affect financial markets through two distinct channels: (1) dynamics of minimum-payouts, and (2) funding-shortfall risk. The effect of funding-shortfall risk on asset prices is determined primarily by the probability of a shortfall, and not by the size of the institution, implying that even a small institution can have a non-negligible impact on asset prices if it is sufficiently constrained. The funding-ratio constraint introduces a constrained and an unconstrained region, with the possibility of transitioning from the constrained to the unconstrained region, thus creating a regime-switching behaviour in asset prices. The funding-shortfall risk gives rise to a new priced factor, which increases risk premia and Sharpe ratios, and can create differences in prices of risk for assets with identical consumption risk. The constrained institution simultaneously increases its demand for both riskfree and higher-risk assets, leading to an underdiversified holdings of risky assets, and a bubble-like behaviour in the prices of higher-risk assets. The dynamics of asset returns are influenced by the dynamics of minimum-payouts, leading to the possibility of predictability in asset returns. The demand for longer relative to shorter maturity bonds, as well as put options for maturities, is affected by the relative growth rate of minimum-payouts and aggregate dividend, creating a possible channel through which the business cycle may affect the term structures of risk free rates and implied volatility.

Based on these results, the main implications of our model can be summarised as follows:

- The more constrained institutions, even if they are smaller in size, may have a more significant impact on asset prices, compared to larger institutions that may be unconstrained;

- The shadow price of the funding-ratio constraint, which can be proxied by the funding-ratio, may serve as a factor in determining asset premia, and may be useful in constructing factor models of asset returns;
- Dynamics of constrained institutions' minimum-payouts, such as pensions' and insurers' liabilities, and factors affecting these dynamics, may be used to predict the dynamics of asset risks, returns, and risk-return ratios;
- Funding-ratio constrained institutions may be able to better meet their objectives through non-standard asset-allocation strategies, including under-diversified portfolios and short positions in risky assets;
- The funding-ratio constraint reduces the conditional volatility, especially during bad times (with lower aggregate consumption), and, hence, macro-prudential regulations, such as risk-based funding requirements may be effective in controlling market volatility;
- Higher funding-ratio requirements may increase the equilibrium price of risk, and, hence, the cost of capital for firms;
- Higher capital requirements for banks may increase their appetite for more-risky loans relative to the less-risky loans, and may discourage them from longer-term lending;
- In the presence of highly constrained institutions, a regulatory insistence on maintaining funding-ratio levels may inflate asset prices and lead to a bubble-like behaviour in the prices of more-risky assets.

The intuition for the above results is as follows. When the constrained institution holds a larger share of aggregate endowment at the initial date, its endowment in the future periods become less likely to fall below the level required by minimum-payouts. This creates a threshold level of initial endowment share above which the constrained institution can never fail to meet its minimum-payout requirements, and hence, the funding-ratio constraint never becomes binding. We refer to this as the unconstrained region. In the unconstrained region, the presence of a funding-ratio constraint has no effect, and the optimal decisions of the constrained

institution are characterized solely by its preferences and budget constraint. Since the preferences are assumed to be characterised by a power utility function, the standard consumption capital asset pricing model (CCAPM) holds in the unconstrained region. In contrast, in the constrained region, the constrained institution's share of aggregate endowment is small enough to yield a finite probability for the constraint to be violated under the constrained institution's true preferences. Thus, the funding-ratio constraint can become binding. This induces the constrained institution to hold more wealth than it would if it was unconstrained, in order to maintain its wealth above the required level.

Starting from the constrained region, the constrained institution can move out of the constrained region upon positive innovations in the aggregate dividend. A positive innovation in the aggregate dividend increases the aggregate wealth, and hence the constrained institution's wealth, while it decreases the present value of minimum-payouts, as the stochastic discount factors goes down while the promised institutional payouts remain unchanged. The possibility of transitioning from the constrained to the unconstrained region causes the stochastic discount factor to increase more sharply for states with lower aggregate dividend growth, where the constraint is more likely to be binding, yielding a non-linear relation between the log of stochastic discount factor and the growth rate of aggregate dividend. We show that this non-linear factor model can be expressed as a linear two-factor model, where the second factor is related to the shadow price of the funding-ratio constraint. Assets whose payoffs covary more with the funding-ratio factor require higher premia, and since the price of the funding-ratio factor varies with the state of the economy, risk premia for assets with constant consumption risk (beta) become time-varying.

Asset allocations also exhibit a cyclical behaviour, with allocations to the risky asset increasing upon positive innovations to aggregate dividend and decreasing upon negative innovations to aggregate dividend. Despite this cyclical behaviour, the volatility of equity returns does not increase, as the constrained institution is willing to pay more for future cash flows in bad states of the world. That is, the stochastic discount factor changes asymmetrically with innovations to aggregate dividend, causing the equity price to rise more in bad states of the world compared

to good states of the world, and thus lowering volatility. As a result, volatility goes down in earlier periods, where innovations in the aggregate dividend have a more pronounced effect on the stochastic discount factor, and it goes up to its unconstrained level near the expiration of the constraint, where innovations in the aggregate dividend no longer have any impact on the stochastic.

With time-varying minimum institutional payouts, the trend in the minimum-payouts affects the relative demand for consumption in different periods. Hence, asset return moments inherit a predictable trend, given the information about the trend in institutional payouts. The more constrained the institutions is, the larger the effect of prmoised payouts, and the more predictable asset returns are. In addition, the growth rate of minimum-payouts relative to the expected growth rate of aggregate dividend affects the relative demand of longer and shorter maturity bonds, and hence the shape of the term structure of riskfree rates. Thus, changes in the expected growth rate of aggregate dividend can affect the shape of the term structure. For a constant expected growth rate of aggregate dividend relative to the growth rate of minimum-payouts, the term structure is typically upward sloping. However, an increase (decrease) in the relative growth rate of aggregate dividend can make the term structure downward (upward) sloping.A similar argument holds for the term structure of implied volatility, which is predominantly upward sloping, but can be downward sloping when the expected growth rate of aggregate dividend sufficiently exceeds the growth rate of minimum-payouts, making the funding-ratio constraint more strongly binding in the shorter term than it is in the longer term.

These results have implications for setting optimal funding-ratio requirements. As discussed before, an increase in required funding-ratio levels decreases conditional volatility of equity returns, increases the price of aggregate risk, and also affects prices of risk in the cross-section of risky assets. The prices of risk for more-risky assets may be affected less that the prices of risk for less-risky assets. Moreover, in order to reduce the volatility of its funding-ratio, the constrained institution invests less in risky assets with longer duration, as the value of these assets is more sensitive to the cashflow shocks. In the case of banks, higher funding-ratio (capital) requirements may have an undesirable effect by making bank loans more

expensive, and increasing banks' appetite for more-risky loans relative to less-risky ones, while discouraging them from longer-term lending. However, higher funding-ratio requirements may decrease market volatility, by increasing asset prices in bad states of the world, even though the price of risk increases. Thus, the optimal capital requirement may have to trade-off the benefits of higher funding-ratio against some of its downsides.[1]

The rest of the paper is structured as follows. Section 2 discusses related literature. Section 3 presents our economic setting, constraints, and the optimization problem for the two institutions. Section 4 discusses the effects of the funding-ratio constraint on asset allocations, institutional payouts, asset prices, asset risk premia, return volatility, Sharpe ratios, and the term structures of risk free rates and implied volatility. Section 6 concludes.

2 Related Literature

Our paper is most closely related to the literature on portfolio insurance, which typically seeks to study the effects of portfolio insurers—investors who seek to maintain their wealth above a certain level—on asset prices. Examples of this approach can be found in Brennan and Schwartz (1989), Basak (1995), and Basak and Shapiro (2001). Brennan and Schwartz (1989), and Basak (1995) study portfolio insurance in a two-period, and a continuous time setting, respectively, and Basak and Shapiro (2001) study a value-at-risk (VaR) constraint by embedding the demand for insurance in the insurer's utility function. In the case of institutions, however, the constraint is typically imposed exogenously, for instance through regulation, and is not necessarily consistent with the institution's preferences, as discussed earlier. Therefore, our first major departure from the portfolio insurance literature is that we impose funding-ratio constraint alongside budget constraint, instead of modifying preferences such that the constraint is always

[1] Our results are derived on a pure-exchange economy, and may or may not hold in a production economy, which is more suitable for analysing the real effects of funding-ratio constraints. But a full analysis of the constraint in a production economy is beyond the scope of this paper, and is left for future work.

respected. This creates a tension between the institution's unconstrained preferences and the funding-ratio constraint, which gives rise to a two-factor asset pricing model, where the second factor is related to the funding-ratio, which does not arise in portfolio insurance models. We compare our results with the results of a portfolio insurance model in more detail in Appendix A.1.

Our modelling choice is closer to Grossman and Zhou (1996), and Cuoco, He, and Isaenko (2008), who impose a wealth and a VaR constraint, respectively, outside of preferences, like we do. However, they specify preferences over terminal wealth with no intermediate consumption, and set the wealth floor at a certain fraction of initial wealth. In the case of institutions, as discussed before, the minimum level of wealth at every date is determined by their promised future payments to fund beneficiaries, and is not necessarily constant across different states of the world. That is, the required level of wealth in states with lower aggregate endowment (bad states) can be higher than the required level of wealth in states with higher aggregate dividend (good states), because a given stream of future consumption becomes more valuable in bad states of the world. As a result, the constraint cannot be satisfied by buying a European put option with a fixed strike price and maturity, as is commonly assumed in the portfolio-insurance literature. Therefore, our second major departure from this portfolio insurance literature is that instead of specifying the required level of wealth directly, we set the level of required wealth as the present value of future minimum-payouts.

Also related is the literature on the effects of non-marketable labor income on asset prices. El Karoui and Jeanblanc-Picqué (1998), and Detemple and Serrat (2003) study the effects of constraints on borrowing against labor income. In this case, the traded assets are used to finance only surplus consumption (the consumption above labor income). The demand for surplus consumption goes down as labor income increases, thus reducing asset prices. In contrast, in our model as well as in other models with wealth (or consumption) constraints but no labor income, traded assets are used to finance the total consumption, thus increasing the demand for traded assets and inflating asset prices.

3 The Model

We consider a discrete-time T-period pure-exchange economy with a single perishable consumption good that serves as the numeraire, one riskfree and a single risky asset, populated by two representative institutions, one of which represents all the unconstrained institutions that either do not face a funding-ratio constraint or that are in the unconstrained region from $t = 0$, and the other represents all the funding-ratio constrained institutions in the economy. The traded assets are indexed by $n \in 0, 1$, where $n = 0$ denotes the riskfree asset and $n = 1$ denotes the risky asset, and the two institutions are indexed by $m \in 0, 1$, where $m = 0$ denotes the unconstrained institution and $m = 1$ denotes the constrained institution. The risky asset is in a positive net supply that is normalized to one, and is a claim to a stochastic dividend, $d_{1,t}$, which follows a binomial process, while the riskfree asset is a one-period riskfree claim to one unit of the consumption good, and is in zero net supply. The information generated by exogenous dividend realisations is denoted by a filtration, $\{\mathcal{F}_t; 0 \leq t \leq T\}$, and is available to all investors. We consider three modelling choices for the funding-ratio constraint: (1) an exogenously imposed constraint on the amount of invested wealth, (2) an exogenously imposed constraint on the payouts, and (3) a model where constrained institution's preferences are defined over surplus-payout (payout above the required payouts to end-investors). The equilibrium quantities are denoted by a ring, $\mathring{\cdot}$, a caret, $\widehat{\cdot}$, a tilde, $\tilde{\cdot}$, and a breve, $\breve{\cdot}$, for the case of unconstrained model, funding-ratio constraint model, minimum-payout constraint model, and surplus-payout utility model, respectively. In Appendix A, we provide a detailed discussion and comparison of these different modelling choices, and explain why an exogenously imposed constraint on invested wealth is selected as the more suitable modelling choice for institutions, which serves as the main focus of our analysis.

3.1 Aggregate Dividend Process

The risky asset's dividend is assumed to be following an exogenously given binomial process

$$d_{1,t} = d_{1,t-1} e^{\mu - \frac{1}{2}\sigma^2 + \sigma z} \tag{1}$$

where $d_{1,0}$ is the initial dividend, and μ and σ are the mean and volatility of the growth rate of the aggregate dividend, and z can take values of -1 and 1 with an equal probability of 1/2.

3.2 Preferences

In order to stay close to the canonical consumption CAPM, and in line with He and Krishnamurthy (2013), and Cuoco and Kaniel (2011), we assume that both the constrained and unconstrained institutions have constant relative risk aversion (CRRA) preferences with power utility. That is, the utility function for both institutions can be written as

$$u_m(c_{m,t}) = \frac{c_{m,t}^{1-\gamma}}{1-\gamma}, \tag{2}$$

where $c_{m,t}$ is the amount of fund-withdrawal (payout) at time t, and the risk aversion parameter, γ, for both institutions is set to be $\gamma = 2$. Because most institutional investors are generally allowed to withdraw funds in case the available funding exceeds the required level of funding, we interpret $c_{m,t}$ as the total withdrawal from the fund, which is used to make institutional payouts to all its stakeholders. A fraction of this total payout goes to the institution's end-investors, and is denoted by $c_{m,t}^{min}$. For instance, in the case of a defined-benefit pension plan, $c_{m,t}$ may consist of payments to pension holders ($c_{m,t}^{min}$), as well as payments to corporate plan sponsors, which may be used for corporate expenditures, such as investments and salary payments. This presence of multiple stakeholders who benefit from institutional payouts also provides a justification for the imposition of a funding-ratio constraint, which can then be seen as an attempt by end-investors

(recipients of $c_{m,t}^{min}$) to insure that their share of payouts is not diverted to other stakeholders.

3.3 Objective Function

We write the objective function of mth institution as

$$\begin{aligned} V_{m,t} &= \underset{c_{m,t},\theta_{m,t}^{n}}{max} \; E_{m,t}\left[\sum_{i=t}^{T} \beta_m^{(i-t)} u_m(c_{m,i})\right], \\ &= \underset{c_{m,t},\theta_{m,t}^{n}}{max} \left[u_m(c_{m,t}) + E_{m,t}\left(\sum_{i=t+1}^{T} \beta_m^{(i-t)} u_m(c_{m,i})\right)\right] \\ &= \underset{c_{m,t},\theta_{m,t}^{n}}{max} \left[u_m(c_{m,t}) + E_{m,t}\left(V_{m,t+1}(W_{m,t+1})\right)\right], \end{aligned} \tag{3}$$

where u_m denotes the utility function, $c_{m,t}$ denotes institutional payout (fund-withdrawal), β_m^t denotes subjective time discount rate, $\theta_{m,t}^n$ denote allocations to traded assets, and $W_{m,t}$ denotes the total wealth for m^{th} institution at date t.

Forming the Lagrangian for the unconstrained institution, we obtain

$$\begin{aligned} L_{0,t} &= u_0(c_{0,t}) + \beta E_t\left[V_{0,t+1}(W_{0,t+1})\right] \\ &\quad + \lambda_{0,t}^{bc}\left(\theta_{0,t-1}^0 + \theta_{0,t-1}^1(d_{1,t-1} + P_{1,t}) - (c_{0,t} + \theta_{0,t}^0 P_{0,t} + \theta_{0,t}^1 P_{1,t})\right), \end{aligned} \tag{4}$$

and forming the Lagrangian for the constrained institution, we obtain

$$\begin{aligned} \widehat{L}_{1,t} &= u_1(c_{1,t}) + \beta E_t\left[V_{1,t+1}(W_{1,t+1})\right] \\ &\quad + \lambda_{1,t}^{bc}\left(\theta_{1,t-1}^0 + \theta_{1,t-1}^1(d_{1,t-1} + P_{1,t}) - (c_{1,t} + \theta_{1,t}^0 P_{0,t} + \theta_{1,t}^1 P_{1,t})\right) \\ &\quad + \lambda_{1,t}^{fr}\left[\theta_{1,t}^0 P_{0,1} + \theta_{1,t}^1 P_{1,t} - \phi^{min} F_t^{min}\right], \end{aligned} \tag{5}$$

$W_{m,t}$ denotes entering wealth at time t, $F_{m,t}$ denotes exiting wealth at time t, $c_{m,t}^{min}$ denotes minimum-payout at date t, F_t^{min} denotes the present value of all future minimum-payouts at time t, $f_{m,t}$ denotes the ratio of total invested wealth, F_t, to minimum required wealth, F_t^{min}, at time t, ϕ_m^{min} denotes the minimum funding-

ratio requirement, and $\lambda^{bc}_{m,t}$, $\lambda^{fr}_{m,t}$, are the Lagrange multipliers for the budget and funding-ratio constraints, respectively.

3.4 The Constraints

Now we describe the constraints that are present in our model, starting with the budget constraint for both institutions:

$$\underbrace{c_{m,t} + \theta^0_{m,t} P_{0,t} + \theta^1_{m,t} P_{1,t}}_{\text{uses of funds}} = \underbrace{\theta^0_{m,t-1} + \theta^1_{m,t-1}(d_{1,t} + P_{1,t})}_{\text{available funds}}. \tag{6}$$

$\theta^0_{m,t}$ denotes allocation to the riskfree asset, $\theta^1_{m,t}$ denotes allocation to the risky asset for m^{th} institution at time t, $P_{0,t}$, $P_{1,t}$ denote the prices of riskfree and risky assets, respectively, and $d_{1,t}$ denotes the aggregate dividend at time t. The left hand side of Equation (6) shows the uses of funds at time t, and the right hand side of Equation (6) shows the sources of funds available at time t from investments made at time $t-1$.

The funding-ratio constraint is defined such that the invested wealth (present value of future consumption) of the constrained institution exceeds a fraction of the present value of future minimum-payouts ($c^{min}_{t:T} \geq 0$):

$$\underbrace{\sum_{i=t+1}^{T} E_t \left[\beta^{i-t} \frac{u_0'(c_{0,i})}{u_0'(c_{0,t})} \overbrace{c_{1,i}}^{\text{actual payout}} \right]}_{F_{1,t}} \geq \phi^{min}_t \underbrace{\sum_{i=t+1}^{T} E_t \left[\beta^{i-t} \frac{u_0'(c_{0,i})}{u_0'(c_{0,t})} \overbrace{c^{min}_{1,i}}^{\text{required payout}} \right]}_{F^{min}_{1,t}}, \tag{7}$$

where $c_{1,t:T}$ are the actual payouts, which, prior to the imposition of the funding-ratio constraint, are independent of the required payouts, $c^{min}_{1,t:T}$, $F_{1,t}$ is the amount of invested wealth that the constrained institution holds, $F^{min}_{1,t}$ is the present value of future minimum-payouts, ϕ^{min}_t is the fraction of $F^{min}_{1,t}$ that the constrained institution is required to hold, and $\beta^{i-t} \frac{u_0'(c_{0,i})}{u_0'(c_{0,t})}$ is the stochastic discount factor of the unconstrained institution, which in our setting is equal to the stochastic

discount factor of the constrained institution due to market completeness, and serves as the equilibrium discount factor.[2]

For $\phi_t^{min} = 0$, the funding-ratio constraint requires the constrained institution's invested wealth to be non-negative, and is subsumed by the budget constraint. For $\phi_t^{min} > 0$, the funding-ratio constraint requires the constrained institution to hold a positive amount of wealth at all times, and is not necessarily satisfied by satisfying the budget constraint, which allows the constrained institution's wealth to go to zero. Hence, for $\phi_t^{min} > 0$, the funding-ratio constraint will be binding in some states of the world.

The funding-ratio constrained is characterised by two exogenous variables: the level of required funding-ratio, ϕ_t^{min}, and minimum-payouts, $c_{t:T}^{min}$, from current time t to the terminal date T. In general, both of these variables can be time-varying. That is, both the required funding-ratio level, ϕ_t^{min}, and the stream of required payouts, $c_{t:T}^{min}$, can change with time. However, as can be seen from Equation (7), it is only their combined effect, $\phi_t^{min} c_{t:T}^{min}$, that matters. Therefore, we assume ϕ^{min} to be independent of time, because the effect of any time variation in ϕ^{min} can be modelled by adjusting $c_{t:T}^{min}$ accordingly.

$c_{t:T}^{min}$ is determined by an institution's required payouts to its end-investors. For instance, in the case of defined-benefit pension plans and insurers, $c_{t:T}^{min}$ can be interpreted as the contractually promised pension and insurance payments. In this case, ϕ^{min} is the level of funding-ratio that is required by the regulator (see Blome, Fachiner, Franzen, and Scheuenstuhl (2007) for the funding-ratio requirements for pensions and insurers in different parts of the world).

In the case of sovereign wealth funds and endowments, which do not have contractually promised payouts, $c_{t:T}^{min}$ can be interpreted as their expected payouts, which may be determined by their forecasted expenditure needs. Such institutions do

[2]Given that the aggregate dividend follows binomial process, and there are two traded assets, the two institutions equate their stochastic discount factors in equilibrium:

$$\beta \frac{u'(c_{0,t+1})}{u'(c_{0,t})} = \beta \frac{u'(c_{1,t+1})}{u'(c_{0,t}) - \lambda_{1,t}^{fr}}.$$

not necessarily face a regulatory constraint to preserve their funding-ratios above a threshold, but may wish to manage their funds according to such a constraint, to be able to meet their expenditure targets. For instance, endowments often manage their funds according to the 'capital preservation rule', which requires the spending rate to not exceed a constant expected return on assets under management (Dybvig and Qin (2016)). This rule can be seen as a special case of the funding-ratio constraint when expected returns and spending amounts are constant in time, and the fund has an infinitely long horizon. In this case, the funding-ratio constraint can be written as (see Appendix B)

$$\frac{c^{min}}{F_{1,t}} \leq IRR,$$

where IRR is the one-period constant discount rate used by the fund to discount its future spending needs, which may be chosen as the average expected return on assets under management, and the inequality states that the spending-ratio should be less than the IRR. Thus, the funding-ratio constraint provides a generalisation of the capital preservation rule that does not require making restrictive assumptions about the required spending, asset returns, and investment horizon.

Other institutions, such as banks and mutual funds, which do not have an ex-ante contractually promised payout, but are required to honour their end-investors' or depositors' withdrawal requests on a short notice. That is, these institutions face liquidity-ratio constraints, which require them to hold enough liquid wealth to be able to meet their end-investors' withdrawals over a given horizon (BIS (2011)). In this paper, all assets are liquid, and such a constraint would always be satisfied. But in a setting with illiquid assets, a liquidity-ratio constrained can be incorporated by interpreting $c^{min}_{t:T}$ as the expected liquidity needs, i.e. expected withdrawals from depositors and end-investors in the case of banks and mutual funds, respectively, over the given horizon, and F^{min}_t as the liquid wealth that the institution is required to hold.

In addition, banks often face capital requirements, which require them to hold equity capital that exceeds a certain fraction of their total liabilities (deposits). Such a requirement can be incorporated through a funding-ratio constraint by

interpreting F_t^{min} as the bank's total liabilities that are owed to its depositors, and $F_t - F_t^{min}$ as the bank's equity capital. The capital requirement thus requires the bank to maintain

$$\begin{aligned} F_t - F_t^{min} &\geq \phi^{capital} F_t^{min} \\ \Rightarrow F_t &\geq \left(1 + \phi^{capital}\right) F_t^{min}, \end{aligned} \tag{8}$$

where $\phi^{capital} \geq 0$, and can be obtained from the relevant capital requirements. For instance, in Basel III (BIS (2011)), banks are required to hold equity capital of about ten percent of their total risk-weighted assets, i.e. $\phi^{basel} \equiv \frac{F_t - F_t^{min}}{F_t} = 0.1$, which implies $\phi^{min} = 1 + \phi^{capital} = 1 + \frac{\phi^{basel}}{1-\phi^{basel}} \approx 1.11$.

In the case of defined-contribution pension plans and hedge funds, for which payouts depend on fund performance, and which do not have contractually promised payouts, c_t^{min} can be taken as zero for all dates. These institutions effectively do not face a funding-ratio constraint.

3.5 Equilibrium Conditions

Equilibrium in the economy is defined as the set of optimal fund-withdrawals, $c_{m,t}$, and portfolio policies, $\theta_{m,t}^n$, for all institutions under given constraints, and price processes, $P_{n,t}$, for all financial assets such that the financial markets clear in each state of the world at all points in time, and the aggregate fund-withdrawal equals the aggregate endowment (dividend).[3] Thus the set of unknown variables that characterise the equilibrium consists of $\{c_{m,t}, \theta_{m,t}^n, P_{n,t}\}$, $\forall m, n, t$. The equilibrium conditions can be obtained as the first order Karush-Kuhn-Tucker (KKT) conditions from the Lagrangian of both institutions given in Equations (4) and (5), and are given below in Equations 9 to 14. We write equilibrium conditions for a general case where both minimum-payout and funding-ratio constraints are imposed. The

[3]By Walras's law, clearing of financial markets implies clearing of goods markets, leaving the last condition, which requires the aggregate payout to be equal to the aggregate dividend, redundant.

equilibrium conditions for the case of a single constraint can then be obtained by setting the Lagrange multiplier of the other constraint to be zero.

For the constrained institution, equilibrium conditions are as follows. The first order condition w.r.t. payouts is

$$\lambda_{1,t}^{bc} = u_1'(c_{1,t}) + \lambda_{1,t}^{mw}. \tag{9}$$

The first order condition w.r.t. $\lambda_{1,t}^{bc}$ is

$$c_{1,t} + \theta_{1,t}^0 P_{0,t} + \theta_{1,t}^1 P_{1,t} = \theta_{1,t-1}^0 + \theta_{1,t-1}^1 (d_{1,t-1} + P_{1,t}). \tag{10}$$

The first order condition w.r.t. the riskfree asset allocation is

$$P_{0,t} = \frac{\beta E_t \left[V_{1,t+1}' \right]}{\lambda_{1,t}^{bc} - \lambda_{1,t}^{fr}}. \tag{11}$$

The first order condition w.r.t. the risky asset allocation is

$$P_{1,t} = \frac{\beta E_t \left[V_{1,t+1}' (P_{1,t+1} + d_{1,t+1}) \right]}{\lambda_{1,t}^{bc} - \lambda_{1,t}^{fr}}. \tag{12}$$

The complementary slackness conditions for minimum-withdrawal, and funding-ratio constraints are

$$\lambda_{1,t}^{mw} (c_{1,t} - c_{1,t}^{min}) = 0 \tag{13}$$

$$\lambda_{1,t}^{fr} \left[\theta_{1,t}^0 P_{0,1} + \theta_{1,t}^1 P_{1,t} - \phi^{min} F_{1,t}^{min} \right] = 0. \tag{14}$$

The envelope theorem yields

$$\frac{dV_{1,t}}{dW_{1,t}} = \frac{\partial L_{1,t}}{\partial W_{1,t}} = \lambda_{1,t}^{bc}. \tag{15}$$

The unconstrained institution's equilibrium conditions can be obtained by setting λ^{mw} and λ^{fr} to zero in the constrained institution's set of equilibrium conditions (Equations (9) to (14)).

The market clearing conditions for the two financial assets lead to

$$\sum_{m=0}^{1} \theta^0_{m,t} = 0 \tag{16}$$

$$\sum_{m=0}^{1} \theta^1_{m,t} = 1. \tag{17}$$

It can be seen from the equilibrium conditions that the Lagrange multiplier for the funding-ratio constraint does not affect the first-order condition for payouts (Equation (9)), which determines payouts across states, but only affects the first-order conditions for portfolio allocations, which determine equilibrium prices. Thus, the funding-ratio constraint affects payouts only indirectly through its effect on portfolio allocations (Equations (11) and (12)), but does not affect the choice of payouts given portfolio allocations. In contrast, for minimum-payout constraint and surplus-payout utility models, both the first order conditions for payouts and portfolio allocations are affected. In the case of minimum-payout constraint model, the first-order condition for payouts is affected due to the Lagrange multiplier of payout constraint, λ^{mw}, and in the surplus-payout utility model, the first-order condition is affected due to a change in utility function. Thus, in these models, payouts are affected even after conditioning on portfolio allocations. Intuitively, the difference arises because of the fact the funding-ratio constraint only constrains the amount of wealth invested, but not the payout. Thus, the institution is free to choose payouts once it holds sufficient wealth to satisfy the constraint, and it makes payouts according to its true preferences. The other two modelling choices directly affect the payout choice, and the institution's payout decisions are affected even after it holds sufficient wealth to satisfy the funding-ratio requirement.

4 Results

In this section, we report our results for the effects of the funding-ratio constraint in a pure-exchange economy with two institutions (one unconstrained and one constrained).

For our numerical examples, we set $\gamma = 2$, and $T = 6$ (7 dates from $t = 0$ to $t = T = 6$). A choice of $T = 6$ provides us a simple few-period setting with sufficient dates to show the effect of transitions from the constrained to unconstrained region, upon sufficiently many positive innovations in the aggregate dividend. The details of the solution are presented in Appendix C. The initial dividend is set to $d_{1,0} = 100$, and the values for μ and σ parameters are set to match the growth rate and volatility of the US consumption data, as estimated by Mehra and Prescott (1985)

$$\begin{aligned} \mu_1 &= 1.83\%; \\ \sigma_1 &= 3.57\%. \end{aligned} \tag{18}$$

For minimum-payouts, we consider following scenarios

- A single promised payout at the terminal date T, which allows us to understand the effect of funding-ratio constraint in the simplest setting;
- A deterministic promised payout at every date, which allows us to understand the effects of time-variation in minimum withdrawals.

For the case of a single minimum withdrawal at date T, we set the amount of minimum withdrawal to a fraction, ω_T^{min}, of the lowest aggregate dividend at time T. That is

$$c_{1,T}^{min} = \omega_T^{min} d_{1,T,T+1}, \tag{19}$$

where $d_{1,T,T+1}$ is the aggregate dividend at time T at node $T + 1$ (the node with the lowest aggregate dividend realisation). For the case of non-zero minimum withdrawals at every date, we assume that minimum-withdrawals start from an

initial value of $c_{1,0}^{min} = \omega_0^{min} \times d_{1,0}$ and grow at a constant rate. That is

$$c_{1,t+1}^{min} = c_{1,t}^{min} e^{\mu^{min}}. \tag{20}$$

A single minimum-payout provides us the simplest setting to explore the effects of a funding-ratio constraint, and a deterministic minimum-payout allows us explore the effects of dynamics of minimum-payouts on asset prices. While, in practice, promised institutional payouts may be contingent on stochastic factors, such as inflation, and may not be fully deterministic, we limit ourselves to deterministic payout profiles in this paper for simplicity. Moreover, the values of minimum-payouts are selected such that the minimum-payout never exceed the minimum dividend realisation at any date.

In the case of single promised payout at the terminal date T, we set the minimum-payout to one-fifth of the lowest realisation of the aggregate dividend at time T, i.e.

$$\omega_T^{min} = \frac{1}{5}. \tag{21}$$

Other choices of ω_T^{min} yield qualitatively similar results. And in the case of time-varying minimum-payouts, we set

$$c_{1,0}^{min} = \frac{1}{5} d_{1,0}, \tag{22}$$

and use different values for μ^{min}.

Our main focus is on the funding-ratio constraint (FR), which is compared with three benchmark models: 1) an unconstrained model (UC) or CCAPM, 2) a model where the utility function depends on payouts in excess of a minimum level (MU), which we refer to as a voluntary portfolio insurance model, described in Section A, and 3) a model with a payout (withdrawal) constraint (MW), described in Section A. The comparison with the unconstrained model allows us to highlight the ability of the funding-ratio constraint to improve predictions of CCAPM. The comparison with the voluntary portfolio insurance model allows us to highlight

the effects of tension between the funding-ratio constraint and the institution's true preferences. And the comparison with the payout constraint model allows us to highlight the ability of a funding-ratio constraint to guarantee minimum institutional payouts.

4.1 Allocations

In this section, we explore the two institutions' portfolio-allocation decisions. At the terminal date T, there are no decisions to made, so we start by considering allocations at time $T-1$. In the unconstrained case, both institutions hold a share of equity that is equal to their consumption share

$$\mathring{\theta}^1_{m,T-1} = \omega_{m,T-1}, \tag{23}$$

and invest nothing in the riskfree asset, as both institutions have identical (unconstrained) preferences

$$\mathring{\theta}^0_{m,T-1} = 0. \tag{24}$$

In the case of a funding-ratio constraint, the wealth constraint at time $T-1$ requires the constrained institution to hold a minimum amount of wealth, $F^{min}_{1,T-1}$, and can be satisfied by holding the same fraction of the aggregate wealth, $\mathring{\theta}^1_{m,T-1}$, at which the constrained institution's unconstrained invested wealth, $F_{1,T-1}$, becomes equal to the required wealth, $F^{min}_{1,T-1}$. Therefore, at time $T-1$, both insitutions optimal allocations to the riskfree asset remain zero

$$\widehat{\theta}^0_{m,T-1} = 0. \tag{25}$$

But the allocations to the risky asset deviate from their unconstrained level when the constraint becomes binding, i.e. when $\mathring{F}_{1,T-1} < \mathring{F}^{min}_{1,T-1}$. The point at which

the funding-ratio constraint becomes binding is given by

$$\omega_{1,T-1}^{min} = \frac{E_{T-1}\left[c_T^{min} d_{1,T}^{-\gamma}\right]}{E_{T-1}\left[d_{1,T}^{1-\gamma}\right]}, \tag{26}$$

which simplifies to

$$\begin{aligned}\omega_{1,T-1}^{min} &= \frac{E_{T-1}\left[c_T^{min}\right]}{E_{T-1}\left[d_{1,T}\right]} \\ &= E_{T-1}\left[\omega_T^{min}\right], \text{ where } \omega_T^{min} = \frac{c_T^{min}}{d_{1,T}},\end{aligned} \tag{27}$$

if c_T^{min} is independent of $d_{1,T}$.

Thus, the constrained institution's optimal allocations to the risky asset can be written as

$$\widehat{\theta}_{1,T-1}^{1} = \begin{cases} \omega_{1,T-1}, & \text{if } \omega_{1,T-1} > E_{T-1}\left[\omega_T^{min}\right] \\ E_{T-1}\left[\omega_T^{min}\right], & \text{if } \omega_{1,T-1} \leq E_{T-1}\left[\omega_T^{min}\right], \end{cases} \tag{28}$$

And the unconstrained institution's optimal allocations to the risky asset are then given by

$$\widehat{\theta}_{0,T-1}^{1} = \begin{cases} \omega_{0,T-1}, & \text{if } \omega_{1,T-1} > E_{T-1}\left[\omega_T^{min}\right] \\ \omega_{0,T-1}\frac{1-E_{T-1}\left[\omega_T^{min}\right]}{1-\widehat{\omega}_{1,T-1}}, & \text{if } \omega_{1,T-1} \leq E_{T-1}\left[\omega_T^{min}\right]. \end{cases} \tag{29}$$

The optimal allocations at time $T-1$ for other modelling choices are given in Appendix A.

Figure 1a shows allocations to the riskfree and risky assets for both institutions. In the case of externally imposed constraints, where the preferences are left unchanged, the constrained institution's allocations in the unconstrained region are not affected, and change only in the constrained region. When the amount of payout is directly constrained (denoted by MW in the figure), the institution is forced to invest in the riskfree asset to meet the minimum-payout at time T, which is constant across states of the world. Thus, it increases its allocation to the riskfree

asset while decreasing its allocation to the risky asset in the constrained region. This higher allocation to the riskfree asset for the minimum-payout constraint model is consistent with Basak (1995) Grossman and Zhou (1996).

But when the funding-ratio at time $T-1$ is constrained but not the amount of payout at time T, the constrained institution only needs to maintain the present value of minimum-payouts at time T, but not the actual payout. Therefore, the institution prefers to hold the required wealth in the risky asset, as the risky asset delivers higher expected dividend for a given level of invested wealth, compared to the riskfree asset, thus providing cheaper insurance. As a result, the constrained institution acts as less risk-averse, in the sense that it invests more in the risky asset, in contrast to the minimum-payout constraint model.

However, when the actual payout at time T becomes risky, it can fall short of the minimum-payout in states with lower aggregate dividend realisation. Thus the constrained institution can fail to meet its minimum-payout obligations in some states of the world despite holding sufficient wealth at time $T-1$. However, this shortcoming can be easily overcome by increasing the funding-ratio level above 1.0. This also provides a rational justification for why 'underfunded' institutions may invest more in the risky asset, without invoking limited liability of fund managers that can incentivise excessive risk-taking. If financial difficulties are short term, i.e. the institution only has difficulty meeting near term payments under its unconstrained optimal policies, then a higher allocation to the equity can meet minimum-payout requirements without sacrificing gains in good states of the world. This effect can be stronger for institutions with higher elasticity of intertemporal substitution (EIS) and lower relative risk aversion (RRA), as they may prefer investing more in the equity by foregoing current payouts, and can be better explored with more general preferences that allow different parameters to control for EIS and RRA. Rime (2001) provides some empirical evidence that banks approaching the minimum regulatory capital level increased their capital but did not affect the level of risk.

At time $t < T - 1$, the optimal allocations to the riskfree and risky assets can be written as

$$\widehat{\theta}^0_{m,t} = \frac{\widehat{W}_{m,t+1,u}\Delta_{u,d}\widehat{W}_{tot,t+1} - \widehat{W}_{tot,t+1,u}\Delta_{u,d}\widehat{W}_{m,t+1}}{\Delta_{u,d}\widehat{W}_{tot,t+1}} \tag{30}$$

$$\widehat{\theta}^1_{m,t} = \frac{\Delta_{u,d}\widehat{W}_{m,t+1}}{\Delta_{u,d}\widehat{W}_{tot,t+1}}, \tag{31}$$

where $\Delta_{u,d}X_t \equiv X_{t+1,u} - X_{t+1,d}$, and the allocations to the two assets are determined by sensitivities of the investor's total wealth, W_m, and the aggregate wealth, W_{tot}, to the innovations in the aggregate dividend.

Figure 2a shows portfolio-allocations at time 0. At time 0, the constrained institution always invests more in the riskfree asset, irrespective of the modelling choice. This is because, at time 0, the funding-ratio is more sensitive to innovations in the aggregate dividend, due to longer duration of minimum-payouts. Upon a negative innovation to the aggregate dividend, the stream of minimum-payouts does not change but a given stream of payouts becomes more valuable, as the discount rate goes down, increasing the required level of funding, while the equity value goes down, due to a decrease in expected future dividends. Thus, the ratio of the required funding to equity value goes down. The higher the duration, the more pronounced this effect. As a result, the equity provides a poor hedge against a funding-shortfall risk at time 0.

This can also be seen in Figure 5a, which shows the time evolution of constrained institution's allocation to bond and equity along the worst and best possible paths of the aggregate dividend. The worst (best) possible path is the one where the innovation to aggregate dividend is always negative (positive). We see that the constrained institution gradually decreases its allocation to the risky asset as it moves out of the constrained region along the best possible path, and increases its allocation to the riskfree asset as it moves further into the constrained region along the worst possible path. This is because as the institution gets nearer to transitioning out of the constrained region, it chooses to meet its funding-ratio

requirement by allocating more to the risky asset, as it promises higher future payoff for a given amount of invested wealth.

4.2 Withdrawals

Given that the optimal allocations to the riskfree asset at time $T-1$ are zero, the optimal payouts at time T are given by

$$\widehat{c}_{m,T} = \widehat{\theta}^{1}_{m,T-1} d_{1,T}. \tag{32}$$

Figure 1b shows the actual payouts at time T. As discussed before, in the case of externally imposed constraint (FR and MW in the figure), the constrained institution's optimal decisions are modified only in the constrained region, and remain unaffected in the unconstrained region. As a result, payouts change non-smoothly, i.e. in a non-differentiable way, both as a function of the state of the world (realisation of the aggregate dividend) at time T, and as a function of the unconstrained institution's share of aggregate endowment at time $T-1$.

In the case of only one non-zero minimum-payout at the terminal date, the two institutions' payouts for $t < T$ is given by

$$\widehat{c}_{m,t} = \omega_{m,t-1} d_{1,t}, \tag{33}$$

Thus, the two institutions maintain the same consumption-share across different realisations of the aggregate dividend for all dates at which $c_t^{min} = 0$, as they would in the absence of a funding-ratio constraint, and the effect of the constraint only affects the level of endowment-share, $\widehat{\omega}_{m,t-1}$. Figure 2b shows the constrained institution's wealth and payout/wealth ratio at time 0. As expected, the amount of wealth held by the constrained institution is higher than the corresponding level in the unconstrained case in the constrained region, and the payout/wealth ratio is lower, as the constrained institution is willing to postpone its payouts until later.

4.3 Analysis of the Stochastic Discount Factor

In this section we analyse the stochastic discount factor (SDF) in our model. For the case of a single non-zero minimum payout at date T, the SDF between time 0 and T can be written as

$$\widehat{SDF}_{0,T} = \beta^T \frac{u'(c_{0,T})}{u'(c_{0,0})} = \beta^T \left(\frac{1-\widehat{\omega}_{1,T}}{1-\omega_{1,0}}\right)^{-\gamma} \left(\frac{d_{1,T}}{d_{1,0}}\right)^{-\gamma}, \tag{34}$$

where

$$\widehat{\omega}_{1,T} = \max\left(E_{T-1}\left[\omega_T^{min}\right], \widehat{\omega}_{1,0}\right). \tag{35}$$

The left panel of Figure 3 shows that the SDF exhibits strong dependence on the state variable, the unconstrained institution's share of aggregate endowment (ω_0), in the constrained region, and the SDF increases as the endowment share of the constrained institution, $1-\omega_0$, goes to zero. This is because the constrained institution's share of aggregate endowment is smaller when it starts with smaller initial asset endowments. All else being equal, a smaller initial endowment promises smaller payoffs in the future, and a higher likelihood that the constrained institution's share of aggregate endowment may fall below its minimum-payout at date T. Thus, in order to maintain its minimum level of wealth, the constrained institution is more willing to forego its current payouts for future payouts, thus increasing the state-prices for payoffs at time T.

As a result, the increase in demand for future payoffs is independent of the size (wealth) of the constrained institution, and is determined by the shadow price of the constraint. This highlights that even a small institution with relatively fewer assets under management can have a significant impact on asset prices if it is sufficiently constrained. Kogan, Ross, Wang, and Westerfield (2006) obtain a similar result for irrational traders, where irrational traders can have a significant impact on asset prices despite their low wealth, and low survival probability. The main difference of their result from ours is that in their model the price impact of irrational traders stems from their inaccurate beliefs about extreme states, such as highly improbable states with very low aggregate endowment, while in our model

the price impact stems from the severity of constraint, and the constrained institution's subjective state prices can dominate in any state as long as the constraint can be sufficiently binding in that state.

The log of SDF can be written as

$$m_{0,T} \equiv \log SDF_{0,T} = T \log \beta - \gamma \log \min \left(1, \frac{1 - E_{T-1}\left[\omega_T^{min}\right]}{1 - \omega_{1,0}}\right) - \gamma \log \frac{d_{1,T}}{d_{1,0}} \tag{36}$$

$$\Rightarrow m_{0,T} = \underbrace{T \log \beta - \gamma \log \frac{d_{1,T}}{d_{1,0}}}_{\text{CCAPM}} - \underbrace{\gamma \log f_{\omega_{T-1}}}_{\text{funding-ratio factor}}, \because \widehat{\omega}_{1,T-1} = \widehat{\omega}_{1,0}. \tag{37}$$

Given that $f_{\omega_{T-1}} = \min\left(1, \frac{1-E_{T-1}\left[\omega_T^{min}\right]}{1-\omega_{1,0}}\right)$ is a non-linear function of $d_{1,T-1}$, the log of SDF depends non-linearly on the dividend growth rate. The right panel of Figure 3 shows this non-linear relation between the log of SDF and the growth rate. When dividend growth is sufficiently high, the constrained institution moves out of the constrained region, and the standard CCAPM holds, as the effect of $f_{\omega_{T-1}}$ on the SDF vanishes. But when the endowment growth is sufficiently low, the constrained institution remains in the constrained region, and its demand for future payoffs increases, thus increasing the state-prices for future payoffs above their unconstrained level. Thus, the relation between the log of SDF and dividend growth becomes non-linear.

To see the implications of this non-linearity on asset premia, note that the expected excess return on any asset can be written as

$$\begin{aligned}\frac{E_t(r_{i,T}-r_{f,T})}{(1+r_{0,T})} &= -Cov_t(m_T, r_{i,T}-r_{f,T}) \\ &= -Cov_t(-\gamma \log f_{\omega_{T-1}} - \gamma \log \frac{d_{1,T}}{d_{1,0}}, r_{i,T}-r_{f,T}) \\ &= \gamma \underbrace{Cov_t(\log f_{\omega_{T-1}}, r_{i,T}-r_{f,T})}_{\text{shortfall premium}} + \underbrace{\gamma Cov_t(\log \frac{d_{1,T}}{d_{1,0}}, r_{i,T}-r_{f,T})}_{\text{CCAPM premium}}.\end{aligned} \tag{38}$$

Thus, the funding-ratio constraint gives rise to a two-factor asset-pricing model, where the second factor, $\log f_{\omega_{T-1}}$, is related to evolution of funding-ratio over time. This will increase the risk premia for assets that covary with this factor. Moreover, given that the funding-ratio factor only varies in the bad states of the world (low aggregate dividend realisations) and remains constant during good states of the world, it will affect the premia of assets with different conditional variances differently. That is, assets that have a higher variance in bad states of the world will earn higher premia compared to assets that have lower variance conditional to bad states of the world, even if they have the same unconditional variance.

The contribution of this second term in Equation (38) vanishes both when the constrained institution has no probability of being underfunded, i.e. the institution is in the unconstrained region, and when the institution has no probability of getting out of the constrained region (in both these cases $\widehat{\omega}_{1,T} = \widehat{\omega}_{1,0}$). In the former case, the log of SDF never deviates from its unconstrained level (dotted black line in the figure), which is a straight line. While in the case of latter, the log of SDF almost converges to the surplus-payout utility model (green line in the figure), which is also a straight line, implying that the economy behaves as if it was populated with institutions with power preferences, albeit with a higher risk aversion. Thus, the unconstrained region and the surplus-payout utility model can be understood as the two limits of an economy with a funding-ratio constraint in which the constraint is never and always binding, respectively.

We highlight the following implications of this nonlinearity

- Even assets with a low covariance with the aggregate dividend (consumption-beta) may require higher risk premia, without higher levels of risk aversion;
- The return on aggregate dividend and the return on equity will not be jointly normally distributed, even if the return on aggregate dividend is normally distributed.

In a linear SDF model, a higher risk aversion increases the slope of the log of SDF as a function of the aggregate endowment growth, and hence the required asset risk premia, for a given covariance between asset returns and the growth of aggregate dividend. In the case of a non-linear SDF, however, the effective slope between the log of SDF and the aggregate endowment growth can increase as the log of SDF deviates from a straight line, without any increase in risk aversion. Thus, assets can require higher premia compared to a linear SDF model for a given level of risk aversion.

4.4 Asset Prices

In this section, we report our results for asset prices, returns, and risk-return ratios. Starting from the terminal date T, where prices for all assets are zero, prices of debt and equity claims at time t are computed in a recursive manner

$$P_{n,t} = \beta E_t \left[\frac{u'(c_{0,t+1})(P_{1,t+1} + d_{n,t+1})}{u'(c_{0,t})} \right],$$

which leads to

$$\widehat{P}_{0,t} = \beta \frac{E_t[d_{1,t+1}^{-\gamma}]}{d_{1,t}^{-\gamma}} \tag{39}$$

$$\begin{aligned}
\widehat{P}_{1,t} &= \frac{1}{d_{1,t}^{-\gamma}} E_t \left[\sum_{i=t+1}^{T-1} \beta^{i-t} d_{1,i}^{1-\gamma} + \beta^{T-t} d_{1,T}^{1-\gamma} f_{\omega_{T-1}}^{-\gamma} \right] \\
&= \underbrace{\frac{1}{d_{1,t}^{-\gamma}} E_t \left[\sum_{i=t+1}^{T} \beta^{i-t} d_{1,i}^{1-\gamma} \right]}_{\text{CCAPM Price}} + \underbrace{\frac{1}{d_{1,t}^{-\gamma}} E_t \left[\beta^{T-t} d_{1,T}^{1-\gamma} \left(f_{\omega_{T-1}}^{-\gamma} - 1 \right) \right]}_{\text{Value of insurance against shortfall risk}} \\
&= \mathring{P}_{1,t} + \Delta \widehat{P}_{1,t},
\end{aligned} \tag{40}$$

where the second term, $\Delta \widehat{P}_{1,t} \geq 0$, can be seen as the present value of the demand for insurance arising due to funding-shortfall risk, which inflates the price of the risky asset.

Once the prices are determined, expected bond and equity returns at time t are computed as

$$R_{0,t} = \log \frac{1}{P_{0,t}} \tag{41}$$

$$E_t[R_{1,t}] = E_t \left[\log \frac{(P_{1,t+1} + d_{1,t+1})}{P_{0,t}} \right], \tag{42}$$

and the conditional volatility and Sharpe ratio of equity return are computed as

$$\sigma_{1,t} = \sqrt{E_t \left[(R_{1,t} - E_t[R_{1,t}])^2 \right]} = \sqrt{E_t \left[R_{1,t}^2 \right] - \left(E_t \left[R_{1,t} \right] \right)^2} \tag{43}$$

$$SR_{1,t} = \frac{E_t[R_{1,t}] - R_{0,t}}{\sigma_{1,t}}. \tag{44}$$

Using Equation 40, the return on equity can be written as

$$\widehat{R}_{1,t} = \log \left(\frac{d_{1,t+1} + \mathring{P}_{1,t+1}}{\mathring{P}_{1,t}} \right) + \log \left(1 + \frac{\Delta \widehat{P}_{1,t+1}}{d_{1,t+1} + \mathring{P}_{1,t+1}} \right) - \log \left(1 + \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}} \right). \tag{45}$$

Using $E_t\left[\log(1+X)\right] \approx \log(1+E_t\left[X\right]) - Var_t\left(X\right)$, the expected equity return can be approximated as

$$E_t\left[\widehat{R}_{1,t}\right] \approx E_t\left[\mathring{R}_{1,t}\right] + \log\left(1 + \frac{E_t\left[\beta\left(\frac{d_{1,t+1}}{d_{1,t}}\right)^{-\gamma}\frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t}}\frac{\mathring{P}_{1,t}-\mathring{P}_{1,t}(\mathcal{F}_{t+1})}{\mathring{P}_{1,t}(\mathcal{F}_{t+1})}\right]}{1+\frac{\Delta\widehat{P}_{1,t}}{\mathring{P}_{1,t}}}\right)$$
$$- Var_t\left(\frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t+1}+d_{1,t+1}}\right), \tag{46}$$

where $\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)$ denotes the time-t value of the dividend stream, $d_{1,t+1:T}$, conditional on time-$t+1$ information, $\mathcal{F}_{t+1}$. Thus, expected equity return inherits two additional terms, which vanish when $\frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t}}$ goes to zero. $\frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t}}$ can be interpreted as the value of insurance against shortfall risk as fraction of the aggregate wealth. The demand for insurance goes up (down) in bad (good) states of the world, as the constrained institution becomes more (less) constrained. Therefore, $\frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t}}$ increases (decreases) when $\mathring{P}_{1,t+1} - E_t\left[\mathring{P}_{1,t+1}\right]$ is negative (positive). Similarly, the stochastic discount factor, $\left(\frac{d_{1,t+1}}{d_{1,t}}\right)^{-\gamma}$, increases (decreases) upon a negative (positive) shock to the aggregate dividend. Hence, the expectation term can be interpreted as the covariance of $\left(\frac{d_{1,t+1}}{d_{1,t}}\right)^{-\gamma}\mathring{P}_{1,t+1} - E_t\left[\mathring{P}_{1,t+1}\right]$ with loss in unconstrained equity value, $\mathring{P}_{1,t} - \mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)$, upon positive and negative shocks to aggregate dividend at time $t+1$. Because the SDF and the demand for insurance are higher upon a negative shock, when the loss in equity value is higher, the expectation term in Equation (46) is positive, and increases the expected equity return. In other words, the loss in equity value is penalised because it covaries positively with the demand for insurance. The last term, $Var_t\left(\frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t+1}+d_{1,t+1}}\right)$, arises due to the concavity of log returns, which serves to decrease the expected log return, $E_t\left[\widehat{R}_{1,t}\right]$, as the variance of insurance value increases, for a given expected insurance value.

The volatility of equity return can then be approximately written as

$$\widehat{\sigma}_{1,t}^2 \approx \mathring{\sigma}_{1,t}^2 + 2Cov_t\left(\mathring{R}_{1,t}, \frac{\Delta\widehat{P}_{1,t+1}}{d_{1,t+1}+\mathring{P}_{1,t+1}} - \frac{\Delta\widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right) + Var_t\left(\frac{\Delta\widehat{P}_{1,t+1}}{d_{1,t+1}+\mathring{P}_{1,t+1}} - \frac{\Delta\widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right), \tag{47}$$

where the covariance in the above expression is negative, as $\frac{\Delta\widehat{P}_{1,t+1}}{d_{1,t+1}+\mathring{P}_{1,t+1}}$, which is the ratio of the insurance value to the aggregate wealth at time $t+1$, increases (decreases) upon a negative (positive) innovation to the aggregate dividend, while $\mathring{R}_{1,t}$ decreases (increases) upon a negative (positive) innovation to the aggregate dividend, bringing volatility down from its unconstrained level. The variance term is positive, and overall change in volatility depends on the relative contributions of the two terms.

Figure 4a shows expected bond return (left panel) and equity premium (right panel) at time $T-1$. The constraint effectively introduces a demand for insurance, as the institution seeks to insure its payout at time T. This demand for insurance becomes stronger as the constrained institution's share of aggregate endowment at time $T-1$ goes down. As a result, asset returns between periods $T-1$ to T inherit an insurance premium, making the asset returns fall below their benchmark levels. When the endowment-share of the constrained institution becomes sufficiently low, the demand for insurance is strong enough to make the riskfree rate negative, as the institution is willing to forego more than one unit of current period's payout in exchange for one unit of payout in the next period (time T).

The lower expected equity return at time $T-1$ in turn inflates the equity price in all periods prior to time T. This creates a bubble-like behaviour in equity prices, as they go up when the constraint is binding, and revert to their unconstrained level as the constrained institution moves out of the constrained region. This can be seen in Figure 5b, which shows the evolution of the ratio of equity prices in constrained and unconstrained economies. In this figure, the price-ratio constantly decreases and reaches 1 at date 4, along the best possible path.

Figure 4b shows the bond return and equity premium at time 0. Bond return at time 0 is unchanged compared to its unconstrained level. However, the equity premium increases for both minimum-payout and funding-ratio constraints. Figure 6 shows the evolution of bond return, equity premium, equity return volatility and Sharpe ratio. We see that both equity premium and Sharpe ratio increase as the constrained institution gets deeper into the constrained region (green curve in the figure), and converge to their unconstrained level as the constrained institution moves out of the constrained region, either due to positive innovations in the aggregate dividend or due to the expiration of constraint. In contrast, the volatility of equity return always remains below its unconstrained level and goes up (down) as the constrained institution moves out (in) of the constrained region.

Thus, the equity premium and Sharpe ratio change counter-cyclically, while the conditional volatility changes in a cyclical manner. This counter-cyclical change is similar to the one observed in the case of heterogeneous agent models, such as Chan and Kogan (2001). In Chan and Kogan (2001), time variation stems from fluctuations in wealth distribution between heterogeneous agents. In our model, it stems both from fluctuations in wealth distribution and state dependence in the constrained institution's subjective state prices. As the constrained institution gets deeper into or out of the constrained region, its demand for the riskfree and risky asset changes, changing asset prices even in the absence of any redistribution of wealth between the two institutions.

From this analysis of asset prices we can see that the effect of institutions that are in the constrained region is not necessarily cancelled out by institutions in the unconstrained region, as the behaviour of institutions is highly asymmetrical across the two regions. Thus the presence of a small number of institutions at a high risk of funding-ratio shortfall can affect the overall behaviour of prices, even if most of the institutions are well-funded.

4.5 Put Option Prices

The demand for portfolio insurance is often used to provide a rationale for the empirically observed volatility smile, because out-of-the-money (OTM) put options can be used to insure portfolio against large losses, making these OTM options relative more attractive (Grossman and Zhou (1996), Bates (2000)). While portfolio insurance provides an intuitive explanation for volatility smile, the implied volatility exhibits several other stylised regularities. Here we explore the ability of the funding-ratio constraint to explain any stylised empirical regularities regarding the prices of put options. We use the risk-neutral probability of a down move (negative innovation) in the aggregate dividend as a proxy for Black Scholes implied volatility, as our model is set in discrete time, while the Black-Scholes option pricing formula is obtained in continuous time. The higher the risk-neutral probability of a downward move, the more expensive the put option is relative to the unconstrained level. The risk-neutral probability is computed as the implied probability of a down move that makes the expected present value of the put option's payoff equal to its price, i.e.

$$\begin{aligned}
&\sum_{j=0}^{T+1} \binom{T}{j} (1-p)^j (p)^{n-j} \beta^{T-t} \frac{u'(c_{0,T,j})}{u'(c_{0,t})} max(K - d_j, 0) \\
=&\sum_{j=0}^{T+1} \binom{T}{j} (1-p^*)^j (p^*)^{n-j} e^{-r^f_{t,T}(T-t)} max(K - d_j, 0) \qquad (48)
\end{aligned}$$

where $\beta^{T-t}\frac{u'(c_{0,T,j})}{u'(c_{0,t})}$ is the equilibrium SDF, $p = 0.5$ is the physical probability of a down move, p^* is the risk-neutral probability of a down move, T is the maturity of the put option, $r^f_{t,T}$ is the riskfree rate for maturity T at time t, and K is the strike price of the put option.

To study both the behaviour of implied volatility across strikes, and across maturities, we consider nine put options for maturities between $T = 1$ to $T = 6$.

$$K_{T,n} = d_{T,T+1} + n\frac{d_{T,1} - d_{T,T+1}}{9},\ n \in \{1, 2, \cdots, 9\}, \qquad (49)$$

where $d_{T,1}$ denotes the highest dividend (dividend at node 1) at time T, and $d_{T,T+1}$ denotes the lowest dividend (dividend at node $T+1$) at time T.

Figure 8a plots the implied risk-neutral probability of a down move for put options of various strikes and maturity T. The figure shows that the implied probability of a negative innovation in the aggregate dividend obtained using time-0 prices of put options for various strike prices, all maturing at time T. The blue curve in the left panel corresponds to the funding-ratio constraint. The implied probability of a downward move is higher for put options with lower strike prices, reflecting a demand for insurance.

The shape and slope of the implied probability curve varies with the constrained institution's share of aggregate endowment. As the institution gets poorer (richer), its decisions are more (less) influenced by the constraint compared to its unconstrained preferences, and its demand for a protection against shortfall goes up (down). This can be seen by comparing the left and right panels of Figure 8a that plot implied probabilities for two different initial endowment shares of the constrained institution. As a consequence, the funding-ratio constraint can not only generate a volatility-smile, but it can also generate time-variation in the slope of the volatility-smile curve due to variations in endowment-shares of the two institutions. Explaining such time variation can be challenging under jump-risk based explanations of volatility smile, unless the jump-risk or jump-risk premium is assumed to vary significantly over time, as in Pan (2002). In contrast, the funding-ratio constraint may provide a more plausible explanation of time-variation in the volatility-smile curve, by linking it to the variation in the constrained institution's demand for insurance as it moves further in or out of the constrained region.

Moreover, the slope of implied risk-neutral probability can also exhibit an empirically documented negative correlation with the equity premium (Yan, 2011). This is because the slope is the largest when the constrained institution is at the threshold of the constrained region. In this case, only deep OTM put options that pay in the worst states of the world are more valuable, but not in-the-money (ITM) put options, as the constrained institution will only require protection in extreme bad states of the world. But as the constrained institution's share of aggregate

endowment decreases, and it gets deeper into the constrained region, its probability of staying in the constrained region increases. As a result, all put options become more valuable, and the implied probability curve moves upward, but the slope of the curve decreases. However, the equity premium continues to increase as the constrained institution gets deeper into the constrained region, yielding a negative relation between the equity premium and the slope of implied volatility curve. This can be seen in the left panel of Figure 8b. The slope of implied probability is computed as the difference in the implied probabilities of put options with strike prices of 95 and 135. The figure is obtained by plotting the equity premium and the slope of implied probability in the region $0.80 < \omega_{0,0} < 1$. Initially, as the constrained institution enters the constrained region, both the slope of implied probability and the equity premium increase simultaneously, exhibiting a positive correlation in $0.80 < \omega_{0,0} < 0.85$ region. But as the constrained institution gets deeper into the constrained region, the relationship is reversed in $0.85 < \omega_{0,0} < 1$ region.

4.6 Time-Varying Minimum-Payouts

The single-payment model discussed above does not shed any light on how the dynamic profile of minimum institutional payouts may affect financial markets. To address this question, in this section we consider a more general setting by allowing a non-zero minimum-payout at every date, as described in Equation (20), and explore the effects of time-varying minimum-payouts on asset returns, volatilities, Sharpe ratios, as well as the term structure of risk free rates and implied volatility.

With time-varying minimum-payouts, the equilibrium price of a τ-maturity bond can be written as

$$\widehat{P}^{\tau}_{0,t} = \underbrace{\frac{1}{d_{1,t}^{-\gamma}} E_t \left[\sum_{i=t+1}^{\tau} \beta^{i-t} d_{1,i}^{-\gamma} \right]}_{\text{CCAPM Price: } \mathring{P}^{\tau}_{0,t}} + \underbrace{\frac{1}{d_{1,t}^{-\gamma}} E_t \left[\sum_{i=t+1}^{\tau} \beta^{i-t} d_{1,i}^{-\gamma} \left(\prod_{j=t}^{i-1} f_{\omega_j}^{-\gamma} - 1 \right) \right]}_{\text{Price deviation due to the constraint: } \Delta \widehat{P}^{\tau}_{0,t}}, \tag{50}$$

and the price of the risky asset can be written as

$$\widehat{P}_{1,t} = \underbrace{\frac{1}{d_{1,t}^{-\gamma}} E_t \left[\sum_{i=t+1}^{T} \beta^{i-t} d_{1,t+i}^{1-\gamma} \right]}_{\text{CCAPM price: } \mathring{P}_{1,t}} + \underbrace{\frac{1}{d_{1,t}^{-\gamma}} E_t \left[\sum_{i=t+1}^{T} \beta^{i-t} d_{1,t+i}^{1-\gamma} \left(\prod_{j=t}^{i-1} f_{\omega_j}^{-\gamma} - 1 \right) \right]}_{\text{Price deviation due to the constraint: } \Delta\widehat{P}_{1,t}}. \quad (51)$$

Hence, the return on one-period riskfree asset can be written as

$$\widehat{R}_{0,t} = \mathring{R}_{0,t} - \log \left(1 + \frac{\Delta \widehat{P}_{0,t}}{\mathring{P}_{0,t}} \right), \quad (52)$$

which falls below its unconstrained level, $\mathring{R}_{0,t}$, as $\frac{\Delta \widehat{P}_{0,t}}{\mathring{P}_{0,t}} \geq 0$. Similarly, the expected return on risky asset can be written as

$$E_t \left[\widehat{R}_{1,t} \right] = E_t \left[\mathring{R}_{1,t} \right] + \log \left(1 + \frac{E_t \left[\widehat{SDF} \frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}} \frac{\mathring{P}_{1,t} - \mathring{P}_{1,t}(\mathcal{F}_{t+1})}{\mathring{P}_{1,t}(\mathcal{F}_{t+1})} \right] - E_t \left[\mathring{SDF} \left(f_{\omega_t}^{-\gamma} - 1 \right) \frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}(\mathcal{F}_{t+1})} \right]}{1 + \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}}} \right)$$
$$- Var_t \left(\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t+1} + d_{1,t+1}} \right), \quad (53)$$

and has an additional term compared to Equation (46) that depends on $\left(f_{\omega_t}^{-\gamma} - 1 \right)$, which measures the value of insurance against the immediate minimum-payout, $c_{1,t+1}^{min}$, and vanishes when the expected next period's payout already exceeds the minimum-payout, i.e. $E_t \left[\widehat{c}_{1,t+1} \right] \geq c_{1,t+1}^{min}$. Because $\left(f_{\omega_t}^{-\gamma} - 1 \right) \geq 0$, this additional expectation is positive, and the demand for insurance against the immediate minimum-payout decreases the expected equity return. That is, while the positive covariance of present value of insurance with equity losses increases the expected equity return, the demand for insurance against the immediate minimum-payout tends to decrease expected equity return.

Thus, when minimum-payouts are allowed to be possibly non-zero at every date, the dynamics of asset-return moments are affected by the dynamics of minimum-payouts. And prices, returns, and Sharpe ratios become a function of the dynamics of minimum-payouts. As a result, these variables will move even in the absence of any news about assets' dividends, as the present value of the required level of fund-

ing changes due to the passage of time. While the changes in minimum-payouts are assumed to be deterministic, if the true underlying dynamic process is not public information, as is likely to be the case, these changes in asset return moments will appear to be stochastic. Therefore, this time-variation can introduce an additional source of uncertainty in the market, which is unrelated to shocks to dividends. Moreover, the higher the time-variation, e.g. the growth rate of minimum-payouts, the stronger is their effect on the time-variation of asset-return moments. Thus a higher growth rate of minimum-payouts can make markets look more noisy, despite the fact that the conditional volatility of equity return actually goes down more as the growth rate of minimum-payouts increases. Figures 7a and 7b compares the price of equity, equity premium, equity premium volatility, and Sharpe ratio for different growth rates of minimum-payouts. As the growth rate of minimum-payouts increases, the equity premium and Sharpe ratio tends to go up, while the conditional volatility goes down, because starting from a fixed minimum-payout, a higher growth rate of minimum-payouts increases the average level of minimum-payouts, making the constraint more tightly binding. The trend in minimum-payouts also introduces a predictable trends in asset return moments, leading to predictability in asset returns, return volatilities, and prices of risk.

4.6.1 Term Structure of Riskfree Interest Rates

The most noticeable effect of dynamic minimum-payouts is on the term structure of riskfree rates. When we only have a single non-zero minimum-payout at the horizon, the term structure of the riskfree rates is downward sloping, as only the demand for the longest maturity bond increases. However, when we allow for a minimum-payout at every date, interest rates at all maturities are affected. The riskfree interest rate for any maturity, τ, at time t can be computed as

$$1 + R^{\tau}_{0,t} = \frac{1}{P^{\tau}_{0,t}}, \tag{54}$$

where $P^{\tau}_{0,t}$ is given in Equation (50). Figure 7c compares the riskfree term structures for different growth rates of minimum-payouts relative to the expected growth rate of aggregate dividend. Different growth rates for minimum-payouts allow us to explore the effects of changes in the expected growth rate of aggregate dividend on the term structure of interest rates. For instance, for a given growth rate of minimum-payouts, an increase in the expected growth rate of aggregate dividend will decrease the relative growth rate of minimum-payouts. The figure shows that the term structure is predominantly upward sloping, but can be downward sloping as well, depending on the relative expected growth rates of aggregate dividend and minimum-payouts, and the endowment shares of the two institutions. The term structure is more likely to be upward sloping when the growth rate of minimum-payouts is less than the expected growth rate of aggregate endowment. In this case, the constrained institution tends to move out of the constrained region due to the relatively higher expected growth rate of aggregate dividend. As a result, the demand for shorter maturity bonds is higher compared to the demand for longer maturity bonds, for all values of the endowment shares of the two institutions, leading to an upward sloping term structure. But when the growth rate of minimum-payouts exceeds the expected growth rate of aggregate dividend, the term structure can be both upward and downward sloping, depending on the constrained institution's share of aggregate endowment. When the constrained institution's share of aggregate endowment is large enough, the institution is only likely to need protection later in the future, when minimum-payouts are a bigger fraction of the aggregate dividend due to higher relative growth rate of minimum-payouts. As a result, the demand for shorter term bonds is much smaller compared to the longer term bonds, leading to a downward sloping term structure. However, as the constrained institution's share of aggregate endowment decreases, i.e. as it moves deeper into the constrained region, the demand for shorter maturity bonds increases, pulling the lower end of the term structure downward. A higher allocation to the riskfree asset in the first few periods makes the institution's wealth less sensitive to innovations in the aggregate dividend, and as a result the demand for the riskfree asset at future dates goes down. Thus, the term structure again becomes upward sloping.

This may provide a possible rationale for downward sloping term structure at the onset of booms, and upward sloping term structure at the onset of recessions, documented in the empirical literature (Harvey, 1988). A negative (positive) news regarding the expected growth rate of aggregate dividend will increase (decrease) the growth rate of minimum-payouts relative to the expected growth rate of aggregate dividends, making a downard (upward) sloping term structure more likely than it was prior to the news. While the term structure may exhibit the empirically observed shape, the term spread, however, does not fit the empirically observed pattern. As the term spread increases when the constrained institution gets deeper into the constrained region, the term spread would be lower during good times, and higher during bad times, which is contrary to what is observed in practice. Nevertheless, the model provides a new channel through which the business cycle (shocks to the growth rate of aggregate dividend) can affect the level and the slope of the term structure.

4.6.2 Term Structure of Implied Volatility

Next, we look at the term structure of implied volatility, with time-varying minimum-payouts. The right panel of Figure 8b shows the term structure of implied probabilities for put options with a given level of 'moneyness'. The strike price for options of all maturities is set equal to the maximum dividend realisation at the corresponding maturity, i.e. $K_T = d_{T,1}$, where $d_{T,1}$ is the highest dividend (dividend at node 1) at time T. This figure shows that while a minimum funding-ratio requirement can explain why OTM put options are expensive, as discussed in Section 4.5, it does not necessarily explain the term structure of implied volatility, which is empirically observed to be downward sloping. Only when constrained institutions are assumed to have short-duration liabilities (minimum-payouts), it can generate a downward sloping term structure of implied volatility. Such a situation, however, is not entirely improbable. If some institutions are facing severe but temporary financial difficulties, then their excessive demand for short term protection from a funding-shortfall can cause the term structure of implied volatility to be downward sloping. This becomes more plausible if restructuring of minimum-payouts

is allowed. As institutions facing financial difficulties restructure their minimum-payouts, they may only stay in severe financial distress for short periods of time. Moreover, as we have argued earlier, the size of these institutions does not need to be significant in order to have a non-negligible price-impact.

4.7 The Model with Two Risky Assets

One noticeable difference of our model from heterogeneous risk aversion models can be seen in the underlying reason for the increase in the Sharpe ratio, discussed in Section 4.4. In heterogeneous risk aversion models, changes in the Sharpe ratio arise due to changes in the representative investor's risk aversion, and affect all risky assets in an identical manner, i.e. the Sharpe ratio for all risky assets changes by the same amount. In the funding-ratio constraint model, however, the change in Sharpe ratio is driven by the additional priced factor, discussed in Section 4.3. Since this factor may have different covariances with different risky assets, Sharpe ratios for different risky assets may change differently, creating a cross-sectional effect in risky assets' prices of risk. We explore this possibility in this section with two risky assets that are perfectly correlated. Both assets' dividends are assumed to follow binomial process, with identical growth rates, starting from an initial value of 50. The second risky asset is assumed to have a dividend that has a higher variance, which is 1.2 times the variance of the first asset risky asset. That is, the two risky assets' dividend processes can be written as

$$d_{n,t} = d_{n,t-1} e^{\mu_n - \frac{1}{2}\sigma_n^2 + \sigma_n z}, n \in [1, 2], \tag{55}$$

with the following parameters

$$\begin{aligned} d_{1,0} &= d_{2,0} = 50; \\ \mu_1 &= \mu_2 = 1.83\%; \\ \sigma_1 &= 3.57\%; \\ \sigma_2 &= 1.2 \times \sigma_1. \end{aligned} \tag{56}$$

Figure 9b shows the constrained institution's allocations to the two risky assets at time $T-1$ under the funding-ratio constraint. In the case of funding-ratio constraint, the constrained institution holds more of the first risky asset compared to the unconstrained institution. This is because the constrained institution is more concerned about the higher variance, as it increases the probability of its funding-ratio falling below the required level, and as a result holds more of the less-risky asset. In sharp contrast to this, at time 0, the constrained institution holds less of the more risky asset, while it increases its allocation to the riskfree asset. Figures 12a and 12b show the time evolution of the constrained institution's portfolio allocations. We see that the constrained institution gradually decreases its allocation to the less-risky asset as it increases its allocation to the riskfree asset, while it keeps its allocation to the more risky-asset unchanged from its unconstrained level until time $T-1$. Thus, in comparison to the unconstrained case, we see that the constrained institution increases its allocation to the riskfree asset, decreases its allocation to the safer of the two risky assets, and keeps its allocation to the riskier of the two assets unchanged. Hence, the constrained institution's allocation to the high-variance asset increases relative to its allocation to the low-variance asset, indicating a simultaneous increase in the demand for the riskfree asset and the more-risky asset relative to the less-risky asset.

In the case of funding-ratio constraint, the riskfree and less-risky assets become substitutes. When the constrained institution is more heavily invested in the riskfree asset, its demand for the less-risky asset is lower, but increases at time $T-1$ when the constrained institution is about to move out of the constrained region and no longer needs the riskfree asset to protect itself against a funding-shortfall. This can be understood by the fact that the constrained institution moves out of the constrained region upon positive innovations in the aggregate endowment, it acts as less risk averse in good states, and more risk averse in bad states of the world. The riskfree asset provides a better protection against a funding-shortfall in bad states of the world, and the more-risky asset provides higher payoff in good states of the world.[4] The less risky of the two risky assets, thus, becomes

[4]Since both assets have identical expected dividend growth rates, i.e. $\mu_1 = mu_2$, starting from the same initial value, the more-risky asset provides a higher (lower) dividend in the up (down) state, compared to the less-risky asset, due to its higher variance.

less attractive, as it neither provides a good hedge against bad states, nor higher payoff in good states. In earlier periods, when the funding-ratio is more sensitive to innovations in the aggregate dividend due to higher duration minimum-payouts, as discussed in Section 4.1, the riskfree asset is more desirable, and the demand for the less-risky asset is lower. Moreover, as the constrained institution's share of aggregate endowment decreases in earlier periods, its demand for riskfree asset increases, and consequently the demand for the less-risky asset decreases, and it can even become optimal to short the less-risky asset, as investment in the more risky asset can provide sufficient endowment in good states of the world to cover any losses on the short position. This substitutability effect between the riskfree and less risky assets decreases the demand for the less risky asset relative to the more risky asset.

This relative increase in demand for the more risky asset decreases the relative premium between the more- and less-risky assets. Thus, the price of the more-risky asset gets more inflated compared to its unconstrained level, exhibiting a bubble-like behaviour.[5] This is shown in the left panel of Figure 12c, which shows that the ratio of expected excess returns for the two assets can change significantly over time, in sharp contrast to the unconstrained case, where the ratio stays constant throughout at a level determined by the ratio of the variances of the two assets' dividends. That is, in the case of funding-ratio constraint, the two assets' expected returns vary over time, even when their cashflow covariances with the aggregate dividend remains unchanged.

The right panel of Figure 12c shows the ratio of the Sharpe ratios of the two assets. In the unconstrained case, this ratio is always one, as both asset are perfectly correlated with each other and with the aggregate dividend, which is merely the sum of both assets' dividends. In the constrained case, however, the Sharpe ratios of the two assets can diverge, indicating that the two assets are required to pay different prices of risk in equilibrium. Moreover, the ratio of Sharpe ratios mostly lie below 1.0, indicating that the second asset, which has a higher dividend volatility, requires a lower price of risk due to an increase in the relative demand

[5] We use the term "bubble" somewhat loosely to refer to a temporary increase in the price of an asset above its unconstrained level.

of the more risky asset. This creates a bubble-like behaviour in the price of high-variance (or equivalently high consumption-beta) assets. As higher-beta assets are charged lower prices of risk, their price levels can get more inflated compared to the low-beta assets, leading to a bubble in these high-beta assets that persists until the constrained institution remains in the constrained region. This can be seen in Figure 11, which plots the difference in price/dividend ratio of the two risky assets, and shows that the difference increases in the constrained region. That is, the more-risky asset becomes more expensive as the constrained institutions gets deeper into the constrained region (i.e. its endowment-share decreases).

This relative increase in the attractiveness of the more-risky asset can be understood from our analysis of the SDF, discussed in Section 4.3, as well. If the more-risky asset covaries less with the funding-ratio factor compared to the less-risky asset, then it's premium relative to the less-risky asset can be lower compared to the unconstrained case. For instance, in an extreme case, where the more-risky asset only pays in those (good) states of the world where the constrained institution moves out of the constrained region, then its covariance with the funding-ratio factor would be zero, as the funding-ratio factor remains zero in these states and is only non-zero in the constrained region. As a result, its premium relative to the less-risky may decrease.

Thus, the funding-ratio constraint may provide a rationale for under-diversification, as well as a simultaneous demand for riskfree and high-risk assets, both of which have been documented in the empirical literature. Notice, however, that in our model, both assets are perfectly correlated and hence there is no benefit from diversification. Nevertheless, the constraint introduces a demand for an under-diversified portfolio that would prevail even when the two assets are not perfectly correlated. And the level of under-diversification will depend on the trade-off between relative benefits of more and less diversified portfolios.[6] Mitton and Vorkink (2007) also obtains under-diversification of risky assets when asset returns are skewed and investors have heterogeneous preferences over skewness. Their model is a partial equilibrium one, and the skewness in asset returns is exogenously given.

[6] We obtain qualitatively similar results, which are not reported in this paper, with imperfectly correlated assets.

In contrast to them, we neither explicitly assume any preferences over skewness or any other higher moments, nor assume any skewness in dividend dynamics, and under-diversification arises as an equilibrium consequence of the funding-ratio constraint.

5 Extension to a Large Number of Investors

Here, we show how the results of the two-institution model can be extended to a large number of investors. Consider a large number of investors, M, some of which face a funding-ratio constraint. For each constrained investor, the optimal allocations to the risky asset are given by

$$\theta^1_{i,T-1} = \begin{cases} \omega_{i,T-1}, & \text{if } \omega_{i,T-1} \geq E_{T-1}\left[\omega^{min}_{i,T}\right] \\ E_{T-1}\left[\omega^{min}_{i,T}\right], & \text{if } \omega_{i,T-1} < E_{T-1}\left[\omega^{min}_{i,T}\right] \end{cases} \tag{57}$$

where $\omega^{min}_{i,T} = \phi^{min}_i \frac{c^{min}_{i,T}}{d_{1,T}}$. If there are $N_{c,T-1}$ investors for which the funding-ratio constraint is binding at time $T-1$, then the aggregate demand for the risky asset of these constrained investors at time $T-1$ is

$$\sum_i^{N_{c,T-1}} \widehat{\theta}^1_{i,T-1} = \sum_i^{N_{c,T-1}} E_{T-1}\left[\omega^{min}_{i,T}\right] = N_{c,T-1}\overline{\omega}^{min}_{T-1}, \tag{58}$$

where

$$\overline{\omega}^{min}_{T-1} = \frac{1}{N_c} \sum_i^{N_{c,T-1}} E_{T-1}\left[\omega^{min}_{i,T}\right],$$

is the average minimum-payout share of the aggregate dividend for a constrained investor. And the allocation to the risky asset of an investor, which is either unconstrained or for which the constraint is not binding at time $T-1$ is

$$\widehat{\theta}^1_{i,T-1} = \widehat{\omega}_{i,T-1}\left(\frac{1 - N_{c,T-1}\overline{\omega}^{min}_{T-1}}{1 - N_c\overline{\omega}_{c,T-1}}\right), \tag{59}$$

where

$$\overline{\omega}_{c,T-1} = \frac{1}{N_{c,T-1}} \sum_{i}^{N_{c,T-1}} \widehat{\omega}_{i,T-1},$$

denotes the average consumption share of a constrained investor. Thus, the solution for a large number of investors can be obtained by writing $f_{\omega_{T-1}}$ as

$$f_{\omega_{T-1}} = \min\left(1, \frac{1 - N_{c,T-1}\overline{\omega}_{T-1}^{min}}{1 - N_{c,T-1}\overline{\omega}_{c,T-1}}\right). \tag{60}$$

For $t < T - 1$, the effect of the funding-ratio can be captured similarly by writing

$$f_{\omega_t} = \min\left(1, \frac{1 - N_{c,t}\overline{\omega}_{t}^{min}}{1 - N_{c,t}\overline{\omega}_{c,t}}\right). \tag{61}$$

All the expressions for equilibrium quantities (allocations, consumptions, and prices) in the funding-ratio model, denoted by $\hat{x}$, written in terms of their unconstrained counterparts, denoted by $\mathring{x}$, and funding-ratio factor, $f_{\omega_{t:T}}$, then remain valid. Thus, the main insights of the two-institution model would carry over to a setting with a large number of investors.

6 Conclusion

In this paper, we have explored the effects of funding-ratio constrained institutions with intertemporal consumption, when there is a tension between the institution's unconstrained preferences and the demands of an externally imposed funding-ratio constraint. We have shown that this setting creates new effects that do not arise in the portfolio insurance literature, where this tension between unconstrained preferences and funding-ratio constraint is absent. One main channel through which these new effects arise is the possibility of transitioning from the constrained region, where the decisions are influenced by the constraint, to the unconstrained region, where the constraint has no impact on the institution's optimal decision,

which leads to a two-factor asset pricing model. This regime-switching effect can be summarised by an additional asset-pricing factor that is related to the funding-ratio, and leads to higher risk premia and Sharpe ratios for assets that covary more with this factor. An additional, and distinct, effect on asset prices arises from the dynamics of contractually promised institutional payouts. The dynamics of promised institutional payouts creates an additional source of time-variation, which may also lead to predictability in asset-return moments. The demands for different maturity bonds, and, hence, the shape of the term structure of riskfree rates, is affected by the relative growth rates of minimum-payouts and aggregate dividend. Thus, shocks to growth rates of aggregate dividend (business cycle) or mimimum-payouts can affect the shape of the term structure.

Appendices

A Alternative Modelling Choices

Given that the goal of a funding-ratio constraint is to insure that the required payouts in future periods can be met, we also consider the case where the constraint is imposed directly on the amount of payout. That is, instead of constraining institutions to hold sufficient wealth, this constraint requires institutions to pay out a minimum amount. Thus, the minimum amount that the constrained institution withdraws at date t always exceeds its minimum-payout

$$c_{1,t} \geq c_{1,t}^{min}, \tag{62}$$

where $c_{1,t}$ denotes the amount withdrawn from the fund, and $c_{1,t}^{min}$ denotes required minimum-payout for time t. We refer to the minimum-withdrawal constraint with MW in our figures.

The constrained institution's Lagrangian in the case of minimum-payout constraint is given by

$$\begin{aligned}\tilde{L}_{1,t} = & u_1(c_{1,t}) + \beta E_t \left[V_{1,t+1}(W_{1,t+1})\right] \\ & + \lambda_{1,t}^{bc} \left(\theta_{1,t-1}^0 + \theta_{1,t-1}^1(d_{1,t-1} + P_{1,t}) - (c_{1,t} + \theta_{1,t}^0 P_{0,t} + \theta_{1,t}^1 P_{1,t})\right) \\ & + \lambda_{1,t}^{mw} \left[c_{1,t} - c_{1,t}^{min}\right],\end{aligned} \tag{63}$$

where $\lambda_{1,t}^{mw}$ is the Lagrange multiplier for the minimum-payout constraint.

Notice that unlike this minimum-payout constraint, the funding-ratio constraint does not constrain the amount of payout at any date. Thus, the funding-ratio constraint is less strong than the minimum-payout constraint, which naturally requires the constrained institution to hold more wealth than the present value of minimum-payouts, but also additionally constrains the amount of payout at each date.

In order to highlight the effects of tension between the constrained institution's true preferences and the funding-ratio constraint, we also consider a model where we impose a minimum-payout constraint by writing the institution's utility function over its surplus-payout (above the level of minimum-payout) such that there is never any mismatch between the institution's preferences and the constraint. This utility function is written as

$$\breve{u}_{1,t} = \frac{(c_{1,t} - c_{1,t}^{min})^{1-\gamma}}{1-\gamma}. \tag{64}$$

We refer to this model as surplus-payout utility model, and for reasons discussed in Section A.0.1, we interpret this model as a model of voluntary portfolio insurance. This model is denoted by MU in figures.

The Lagrangian for $m = 1$ institution in the case of surplus-payout utility model is given by

$$\begin{aligned}\breve{L}_{1,t} = \breve{u}_1(c_{1,t}) + \beta E_t \left[\breve{V}_{1,t+1}(W_{1,t+1})\right] \\ + \lambda_{1,t}^{bc} \left(\theta_{1,t-1}^0 + \theta_{1,t-1}^1(d_{1,t-1} + P_{1,t}) - (c_{1,t} + \theta_{1,t}^0 P_{0,t} + \theta_{1,t}^1 P_{1,t})\right).\end{aligned} \tag{65}$$

A.0.1 Interpretation of Modelling Choices

Below we discuss the differences between these different modelling choices, and the use of the funding-ratio model as our model of interest.

In the case of surplus-payout utility (MU) model, the institution's true preferences are assumed to be defined over its surplus payout, i.e. its payout above the required payout. In this case, the institution, even when unconstrained, always seeks to maintain its payouts above the level of minimum-payouts. As a result, it is always optimal for the institution to maintain a wealth that does not fall below the present value of its minimum-payouts. Such an assumption is at odds with the observation that the UK pension funds' allocations considerably changed after the imposition of a funding-ratio constraint in the pension reform act of 2004 (Gromb and Vayanos, 2010). Under this assumption, any observed funding-shortfall can be attributed

only to unidentified risks, but not to the institution's optimal decisions given the information the institution has about various risks. And since the imposition of a funding-ratio constraint can only modify an institution's optimal decisions given its information about different risks, but cannot improve the institution's information set, this assumption does not provide any justification for why the funding-ratio constraint may be imposed upon institutions through regulation.

One can instead assume that the institution's true preferences are not defined over its surplus payout, but are modified after the imposition of the constraint. In this case, the institution's preferences, and hence its optimal decisions, should only change in those states where the constraint can be violated under its unconstrained preferences, and should remain unchanged in those states of the world where the constraint is respected even under the institution's unconstrained optimal decisions. But this is not the case in the surplus-payout utility model. When the institution's utility function is defined over surplus payout, then the institution's optimal decisions will change in all states of the world, including those states where the constraint was satisfied by the institution's optimal decisions under its true preferences prior to the imposition of the constraint.

Moreover, defining an institution's preferences over surplus payout implicitly assumes that the institution is infinitely averse to falling short of its minimum-payouts, i.e. it's marginal utility goes to infinity as its payouts approache the level of required payouts. This, however, may not be the case due to limited liability, which may limit the institution's dis-utility from a default on its payments. Thus, both these implications of the surplus-payout utility model—changing preferences in every state of the world, and assigning infinite dis-utility to a missed payment—seem overly restrictive, and are at odds with all those reasons discussed above and earlier in Section 1, which may cause an institution to hold lower wealth than the present value its future minimum-payouts. These implications of the surplus-payout utility model are more realistic from a voluntary portfolio insurance perspective, where an agent or institution may voluntarily wish to preserve its wealth above a certain threshold. As a result, the demand for a minimum funding-ratio would be a direct consequence of its true preferences, leaving no tension between its preferences and the funding-ratio requirement. Therefore, we

interpret the surplus-payout utility model as an example of voluntary portfolio insurance.

In the funding-ratio (FR) model, we assume that the institution's true preferences are not always compliant with the demands of the funding-ratio constraint, making it possible for the funding-ratio to fall below its required level, if the institution is allowed to act in an unconstrained way under its true preferences. This can happen for several reasons. The actual promised payments of these institutions typically lie in the distant future, and a temporary decline in funding-ratio does not necessarily make them insolvent. An institution may also be willing to withdraw funds at the expense of a lower funding-ratio, if the institution has a higher demand for current expenditures. For instance, in the case of a corporate pension plan, the plan sponsor may withdraw funds for its corporate investments if it has lucrative investment opportunities at its disposal. Financial institutions' limited liability may also incentivise them to take excessive risk in hopes of benefiting from higher surplus wealth in good states of the world, while leaving the institution underfunded in bad states of the world, where limited liability protects their losses from increasing beyond a threshold. Asset managers who manage institutional assets may have shorter employment tenures than the horizon of institution's promised payouts to end-investors, and may not be adversely affected by the institution's inability to make good on its promised payouts.

This discrepancy between unconstrained preferences and the demands of a funding-ratio constraint then justifies the imposition of the constraint, which then forces the institution to deviate from its optimal unconstrained decisions in those states where the constraint has a probability of being violated, but leaving its optimal decisions unchanged in those states where the constraint is satisfied even under its true preferences. That is, the institution sticks to its optimal decisions under its true preferences whenever possible, but modifies them to meet the demands of the externally imposed funding-ratio constraint, whenever the constraint has a probability of being violated. Thus, the constraint and the institution's true preferences can have opposing effects on the constrained institution's optimal decisions, allowing us to explore the novel effects of this tension on asset prices.

In the minimum-withdrawal (MW) model, we retain this tension between the institution's true preferences and the constraint, but instead of constraining the funding-ratio, we constrain the amount of actual payout (fund-withdrawal). This constraint is more restrictive than the funding-ratio constraint, as it constrains both the payouts and the wealth of the constrained institution, while the funding-ratio constraint only constrains the institution's wealth, but not its payouts. This constraint is however less useful as it relates to future payouts, which are only observable ex-post, and the constraint cannot be satisfied ex-ante.

This tension between an institution's unconstrained preferences and the demands of the funding-ratio constraint, which may stem from any number of reason discussed above and in Section 1, is incorporated through through standard power preferences defined over aggregate payout of the institution, so that the model does not deviate too much from the canonical consumption capital asset pricing model (CCAPM) of Breeden (1979). This choice of preferences can be rationalised by viewing the institution as serving the interests of various stakeholders, which collectively benefit from the institution's payout in every period, and may include end-investors, fund-managers, corporate sponsors etc. Thus, the institution's utility function is defined over its aggregate payout to all these stakeholders, irrespective of how the payout may be distributed between individual stakeholders, creating the potential for an agency conflict between these different stakeholders.

Thus, different stakeholders may wish to impose constraints on institution's decisions to safeguard their interests. And, hence, the funding-ratio constraint can be seen as an instrument to safeguard end-investors' interests, as it ensures that other stakeholders, such as fund-managers and plan sponsors, cannot deviate funds for their use at the expense of future minimum-payouts to end-investors.

A similar rationale for such a constraint can be found in Panageas (2011), which argues that a government may impose a minimum consumption constraint, if the government believes that a consumption drop below a certain level will have an adverse impact on society. In this case the government could impose the funding-ratio constraint to ensure that institutions can maintain a minimum level of payouts to end-investors, even if end-investors themselves would be willing to accept a lower

level of consumption in the future. In this case, the funding-ratio constraint can be seen to arise from a conflict of interest between end-investors, who act in their individual interests, and the government that is acting on the behalf of the whole society.

A.1 Comparison of Different Modelling Choices

In this section, we highlight the main differences between different modelling choices. The main difference between the funding-ratio constraint and payout constraint models is in terms of asset allocations near the horizon. In the case of funding-ratio constraint, the constrained institution prefers to satisfy the constraint by holding more equity, while in the case of minimum-payout constraint, it holds more riskfree asset. Besides this, the behaviour under both these models is largely identical, but is different compared to the surplus-payout utility model.

Comparing the funding-ratio constraint model with the portfolio insurance benchmark (denoted by MU in the figures), we notice that in the funding-ratio constraint model:

- The economy exhibits two distinct regions, one which coincides with the unconstrained model (CCAPM);
- Asset prices exhibit stronger dependence to the innovations in aggregate dividend;
- Equity premium and Sharpe ratios are higher;
- The optimality of the market portfolio no longer holds.

All of the above conclusions are valid for both funding-ratio and minimum-withdrawal constraints, and emerge due to the tension between the constrained institution's true preferences and the externally imposed constraint.

For the portfolio insurance model, the portfolio insurer effectively acts as more risk averse, and always invests more in the riskfree asset, compared to the un-

constrained model (denoted by UC in the figures).[7] This higher allocation to the riskfree asset can be seen in Figures 1a and 2a, and is consistent with Basak (1995). Thus, the portfolio insurer's allocations change for all values of the state variable, ω_0, causing asset prices to change, and we no longer have a constrained and an unconstrained region.

The stronger dependence to the innovations in aggregate dividend can be seen in increased divergence between blue and green curves in Figure 6 for the case of funding-ratio constraint. While solid blue and green curves, which correspond to the funding-ratio model, diverge considerably, dotted blue and green curves, which correspond to the portfolio insurance model, do not. That is, while equity premium, volatility, and Sharpe ratio are considerably different along the best (blue) and worst (green) paths of the aggregate dividend for the funding-ratio model, they are not so different for the portfolio insurance model. This happens because the tension that the constrained institution faces between unconstrained preferences and the funding-ratio constraint becomes more or less strong as the institution moves in or out of the constrained region. For the portfolio insurance model, there is no such tension, and the institution always behaves more or less similarly irrespective of the innovations in the aggregate dividend.

The reason for the higher equity premium and Sharpe ratio in the funding-ratio model compared to the portfolio insurance model is the non-linear asset-pricing model, discussed in Section 4.3. In the case of portfolio insurance model, the asset pricing relation remains linear, and the economy can still be described by power preferences, despite a higher risk aversion. Under power preferences, the market price of risk (Sharpe ratio) is given by the relative risk aversion times the volatility of equity premium. Thus an increase in risk aversion does not need to increase the Sharpe ratio or equity premium as long as the volatility decreases sufficiently. But for externally imposed constraints, the tension between unconstrained pref-

[7] In the case of surplus-payout utility model, the relative risk aversion of the constrained institution is given by

$$RRA^{\mathrm{MU}} = -c\frac{u''(c)}{u'(c)} = \gamma\frac{c}{c - c^{min}} = \gamma(1 + \frac{c^{min}}{c} + \cdots) \geq \gamma.$$

erences and the constraint creates an additional priced risk factor, which is only imperfectly correlated with the aggregate dividend risk, and causes both the equity premium and Sharpe ratio to go up. It can also be seen from the figure that when the endowment share of the constrained institution approaches zero, the economy approximately converges to the portfolio insurance case. That is, for very small values of constrained institution's endowment share, the constraint governs the institution's behaviour almost fully, and any tension between the constraint and unconstrained preferences disappears, and the funding-shortfall risk becomes perfectly correlated with the aggregate dividend risk, causing the equity premium and Sharpe ratio to fall back to their portfolio insurance level.

The additional risk factor, namely the funding-shortfall risk, makes two institutions care about different risks, making them hold heterogenous portfolios of risky assets, described in Section 4.7. This heterogeneity of risky portfolios that is observed for the funding-ratio constraint is in sharp contrast to standard heterogeneous agent models, where changes in risk aversion and time discount rates only make investors change their allocations between risky and riskfree assets, but not within risky assets. This can be seen in the portfolio insurance model, where a change in the portfolio insurer's preferences only makes it reallocate its wealth between risky and riskfree assets, but not within risky assets, as can be seen in Figure 10b.

B Capital Preservation Rule

Here we show that the capital preservation rule can be written as a special case of the funding-ratio constraint. The minimum level of wealth that the funding-ratio constrained investor is required to hold can be written as

$$\begin{aligned} F_{1,t}^{min} &= \phi^{min} E_t \left[\sum_{i=t+1}^{T} \frac{\beta u'(c_{1,i})}{u'(c_{1,t})} c_{1,i}^{min} \right] \\ &= \phi^{min} \sum_{i=t+1}^{T} \frac{c_{1,i}^{min}}{(1+IRR)^{i-t}}, \end{aligned}$$

where IRR is a constant discount rate that makes the discounted present value of $\phi^{min} c_{1,i}^{min}$ equal to $F_{1,t}^{min}$. If required payouts are constant in time, i.e. $c_{1,i}^{min} \equiv c^{min}$, and the investment horizon is infinitely long, i.e. T approaches infinity, then assuming $x = \frac{1}{1+IRR}$, and $n = i - t$, we get

$$\begin{aligned}
\sum_{i=t+1}^{T} \left(\frac{1}{(1+IRR)^{i-t}} \right) &= \sum_{n=1}^{\infty} x^n \\
&= \frac{1}{1-x} - 1 \\
&= \frac{1}{1 - \frac{1}{1+IRR}} - 1 \\
&= \frac{1+IRR}{IRR} - 1 \\
&= \frac{1}{IRR}.
\end{aligned}$$

Thus, the required level of wealth can be written as

$$F_{1,t}^{min} = \phi^{min} \frac{c^{min}}{IRR},$$

and the funding-ratio constraint implies

$$\begin{aligned}
F_{1,t} &\geq \phi^{min} \frac{c^{min}}{IRR} \\
\Rightarrow \frac{c^{min}}{F_{1,t}} &\leq \frac{IRR}{\phi^{min}},
\end{aligned}$$

which states that, for $\phi^{min} = 1$, the spending-ratio should be no greater than a of the constant discount rate used by the fund to discount its future payouts.

C Analytical Solution

In this section, we provide more details of the analytical results used in the text, for all three modelling choices.

C.1 Funding-Ratio Constraint Model

C.1.1 Optimal Allocations at time $T-1$ and Optimal Consumptions at time T

When the constraint is not binding, the optimal allocations are given by

$$\begin{aligned} \widehat{\theta}^0_{m,T-1} &= 0 \\ \widehat{\theta}^1_{m,T-1} &= \omega_{m,T-1}. \end{aligned} \tag{66}$$

This can be easily verified that this solution satisfies all the kernel and market-clearing conditions.

To obtain optimal allocations when the constraint is binding, first, we guess that the allocations to the riskfree asset at time $T-1$ are zero for all investors

$$\widehat{\theta}^0_{m,T-1} = 0, \tag{67}$$

and, hence, the market clearing conditions for the riskfree asset are satisfied. We conjecture that the constrained investor can maintain its funding-ratio requirement by holding a fraction, $\omega^{min}_{T-1} = E_{T-1}\left[\omega^{min}_T\right]$, of the aggregate wealth, which can be achieved by holding ω^{min}_{T-1} shares of the equity. Hence,

$$\widehat{\theta}^1_{1,T-1} = \omega^{min}_{T-1}. \tag{68}$$

By market clearing, the unconstrained institution's allocation to the risky asset is then given by

$$\widehat{\theta}^1_{0,T-1} = 1 - \omega^{min}_{T-1}. \tag{69}$$

The remaining two equations are the kernel conditions for riskfree and risky assets for $m = 0$ and $m = 1$ investors. For the riskfree asset, it is given by

$$E_{T-1}\left[\left(\frac{1-\omega_{T-1}^{min}}{1-\omega_{1,T-1}}\frac{d_{1,T}}{d_{1,T-1}}\right)^{-\gamma}\right] = E_{T-1}\left[\frac{\left(\widehat{\theta}_{1,T-1}^{1}d_{1,T-1}\right)^{-\gamma}}{c_{1,T-1}^{-\gamma}-\lambda_{T-1}^{fr}}\right].$$

For the risky asset, it is given by

$$E_{T-1}\left[\left(\frac{1-\omega_{T-1}^{min}}{1-\omega_{1,T-1}}\frac{d_{1,T}}{d_{1,T-1}}\right)^{-\gamma}d_{1,T}\right] = E_{T-1}\left[\frac{\left(\widehat{\theta}_{1,T-1}^{1}d_{1,T-1}\right)^{-\gamma}}{c_{1,T-1}^{-\gamma}-\lambda_{T-1}^{fr}}d_{1,T}\right].$$

Both of these equations can be satisfied by setting

$$\lambda_{T-1}^{fr} = \left(\omega_{1,T-1}d_{1,T-1}\right)^{-\gamma} - \left(\omega_{T-1}^{min}d_{1,T-1}\right)^{-\gamma}\left(\frac{1-\omega_{T-1}^{min}}{1-\omega_{1,T-1}}\right)^{-\gamma}.$$

Thus the assumed solution satisfies all equilibrium conditions.

C.1.2 Consumptions at $t < T$

In the case of only a single non-zero minimum-payout at the horizon, the constraint at time $T-1$ is satisfied by holding sufficient wealth at time $T-1$, and, thus, the payouts across states (dividend realisations) at time $T-1$ remains unaffected by the constraint

$$\widehat{c}_{m,t} = \widehat{\omega}_{m,t}d_{1,t}, \tag{70}$$

and the effect of the constraint only affects the level of endowment-share $\widehat{\omega}_{m,T-1}$. Thus, the constraint would only affect the choice of initial endowment-share, $\omega_{m,0}$, and the two institutions would maintain the same endowment-share, $\omega_{m,0}$, for all $t < T$. At $t = T$, however, the endowment-shares of the institutions will vary across different realisations of the aggregate dividend, depending on whether or not the constraint was binding at time $T-1$.

The initial payout at date 0 is determined by the budget constraint and exogenous endowments in the risky asset (denoted by $\overline{\theta}_m^1$)

$$c_{m,0} + F_{m,0} = \overline{\theta}_m^1 \left(d_{1,0} + P_{1,0}\right) \tag{71}$$

which can be solved for $c_{m,0}$, because $F_{m,0}$ can be written as a function of $c_{m,0}$ by writing the optimal allocations and prices as a function of $c_{m,0}$ as shown in Sections C.1.4 and C.1.3.

C.1.3 Prices at $t < T$

Equilibrium prices at time $T-1$ can be written as

$$\widehat{P}_{0,T-1} = \begin{cases} \beta \frac{E_{T-1}[d_{1,T}^{-\gamma}]}{d_{1,T-1}^{-\gamma}}, & \text{if } \omega_{1,T-1} > E_{T-1}[\omega_T^{min}] \\ \beta \left(\frac{1-E_{T-1}[\omega_T^{min}]}{1-\widehat{\omega}_{1,T-1}}\right)^{-\gamma} \frac{E_{T-1}[d_{1,T}^{-\gamma}]}{d_{1,T-1}^{-\gamma}}, & \text{if } \omega_{1,T-1} \leq E_{T-1}[\omega_T^{min}] \end{cases} \tag{72}$$

$$\widehat{P}_{1,T-1} = \begin{cases} \beta \frac{E_{T-1}[d_{1,T}^{1-\gamma}]}{d_{1,T-1}^{-\gamma}}, & \text{if } \omega_{1,T-1} > E_{T-1}[\omega_T^{min}] \\ \beta \left(\frac{1-E_{T-1}[\omega_T^{min}]}{1-\widehat{\omega}_{1,T-1}}\right)^{-\gamma} \frac{E_{T-1}[d_{1,T}^{1-\gamma}]}{d_{1,T-1}^{-\gamma}}, & \text{if } \omega_{1,T-1} \leq E_{T-1}[\omega_T^{min}] \end{cases}. \tag{73}$$

For $t < T-1$, the equilibrium prices can be computed recursively as

$$P_{n,t} = \beta E_t \left[\frac{u'(c_{0,t+1})}{u'(c_{0,t})} \left(d_{n,t+1} + P_{n,t+1}\right)\right],$$

which leads to

$$\widehat{P}_{0,t} = \beta \frac{E_t[d_{1,t+1}^{-\gamma}]}{d_{1,t}^{-\gamma}} \tag{74}$$

$$\widehat{P}_{1,t} = \frac{1}{d_{1,t}^{-\gamma}} E_t \left[\sum_{i=t+1}^{T-1} \beta^{i-t} d_{1,i}^{1-\gamma} + \beta^{T-t} d_{1,T}^{1-\gamma} f_\omega^{-\gamma}\right], \tag{75}$$

where

$$\Delta_{u,d} X = X_u - X_d. \tag{76}$$

C.1.4 Optimal Allocations for $t < T - 1$

Using the budget constraint

$$c_{m,t+1,\xi} + F_{m,t+1,\xi} = \theta^0_{m,t} + \theta^1_{m,t} \left(d_{1,t+1} + P_{1,t+1} \right), \tag{77}$$

for $\xi \in \{u, d\}$. This yields two equations, which can be solved to express asset allocations at time t as

$$\begin{aligned} \widehat{\theta}^0_{m,t} &= \frac{\widehat{W}_{m,t+1,u} \Delta_{u,d} \widehat{W}_{tot,t+1} - \widehat{W}_{tot,t+1,u} \Delta_{u,d} \widehat{W}_{m,t+1}}{\Delta_{u,d} \widehat{W}_{tot,t+1}} \\ \widehat{\theta}^1_{m,t} &= \frac{\Delta_{u,d} \widehat{W}_{m,t+1}}{\Delta_{u,d} \widehat{W}_{tot,t+1}}, \end{aligned} \tag{78}$$

where $\widehat{W}_{tot,t+1} = d_{1,t+1} + P_{1,t+1}$ and $\widehat{W}_{m,t+1} = \widehat{c}_{m,t+1} + F_{m,t+1}$ denotes the aggregate wealth, and the total wealth of the mth institution at time $t+1$, respectively. Thus, the optimal allocations at any date t can be computed recursively.

C.1.5 Solution with Time-Varying Minimum-Payouts

The effect of wealth distribution can be summarised in one function

$$\widehat{f}_{\omega_{T-1}} = \min \left(1, \frac{1 - E_{T-1}\left[\omega_T^{min}\right]}{1 - \omega_{1,T-1}} \right) = \begin{cases} 1, & \text{if } \omega_{1,T-1} > E_{T-1}\left[\omega_T^{min}\right] \\ \frac{1 - E_{T-1}\left[\omega_T^{min}\right]}{1 - \omega_{1,T-1}} & \text{if } \omega_{1,T-1} < E_{T-1}\left[\omega_T^{min}\right]. \end{cases} \tag{79}$$

The equilibrium quantities can then be written more succinctly in terms of their unconstrained counterparts, which we denote by $\mathring{}$, as

$$\begin{aligned}
\widehat{\theta}^0_{m,T-1} &= \mathring{\theta}^0_{m,T-1} = 0 \\
\widehat{\theta}^1_{0,T-1} &= \mathring{\theta}^1_{0,T-1} f_{\omega_{T-1}} \\
\widehat{\theta}^1_{1,T-1} &= \mathring{\theta}^1_{1,T-1} f_{\omega_{T-1}} + 1 - f_{\omega_{T-1}} \\
\widehat{c}_{m,T} &= \mathring{c}_{m,T-1} f_{\omega_{T-1}} \\
\widehat{P}_{n,T-1} &= \mathring{P}_{n,T-1} f^{-\gamma}_{\omega_{T-1}}.
\end{aligned} \tag{80}$$

With time-varying minimum-payouts, the above structure can be generalised by writing

$$\widehat{f}_{\omega_t} = \min\left(1, \frac{1 - E_t\left[\omega^{min}_{t+1}\right]}{1 - \omega_{1,t}}\right). \tag{81}$$

Then, the optimal payouts of constrained institution at every date t can be written as

$$\widehat{c}_{1,t} = \max\left(E_t\left[\omega^{min}_{t+1}\right], \widehat{\omega}_{1,t-1}\right) d_{1,t}, \tag{82}$$

and equilibrium prices can be written as

$$\widehat{P}^{\tau}_{0,t} = \frac{1}{d^{-\gamma}_{1,t}} E_t\left[\sum_{i=t+1}^{\tau} \beta^{i-t} d^{-\gamma}_{1,i} \prod_{j=t}^{i-1} \widehat{f}^{-\gamma}_{\omega_j}\right] \tag{83}$$

$$\widehat{P}_{1,t} = \frac{1}{d^{-\gamma}_{1,t}} E_t\left[\sum_{i=t+1}^{T} \beta^{i-t} d^{1-\gamma}_{1,t+i} \prod_{j=t}^{i-1} \widehat{f}^{-\gamma}_{\omega_j}\right], \tag{84}$$

where $\widehat{P}^{\tau}_{0,t}$ is a τ maturity riskfree bond.

The log of SDF between time t to T can then be written as

$$\begin{aligned}\hat{m}_{t,T} &= \log\left(\beta^{T-t}\frac{u'(\hat{c}_{0,T})}{\hat{c}_{0,t}}\right)\\ &= \log\left(\beta^{T-t}\left(\frac{d_{1,T}-\hat{c}_{1,T}}{d_{1,t}-\hat{c}_{1,t}}\right)^{-\gamma}\right)\\ &= \log\left(\beta^{T-t}\left(\frac{1-\hat{\omega}_{1,T}}{1-\hat{\omega}_{1,t}}\right)^{-\gamma}\left(\frac{d_{1,T}}{d_{1,t}}\right)^{-\gamma}\right).\end{aligned}$$

Given that

$$\begin{aligned}\hat{\omega}_{1,T} &= \max\left(E_{T-1}\left[\omega_T^{min}\right],\hat{\omega}_{1,T-1}\right)\\ \hat{\omega}_{1,T-1} &= \max\left(E_{T-2}\left[\omega_{T-1}^{min}\right],\hat{\omega}_{1,T-2}\right)\\ &\vdots\\ \hat{\omega}_{1,t} &= \max\left(E_{t-1}\left[\omega_t^{min}\right],\omega_{1,t-1}\right) = \max\left(\omega_t^{min},\mathring{\omega}_{1,t}\right)\end{aligned}$$

Thus

$$\hat{\omega}_{1,T} = \max\left(E_{T-1}\left[\omega_T^{min}\right],E_{T-2}\left[\omega_{T-1}^{min}\right],\cdots,\omega_t^{min},\mathring{\omega}_{1,t}\right), \tag{85}$$

and

$$\hat{m}_{t,T} = (T-t)\log\beta - \gamma\log\frac{d_{1,T}}{d_{1,t}} - \gamma\log\frac{1-\hat{\omega}_{1,T}}{1-\hat{\omega}_{1,t}}. \tag{86}$$

C.1.6 Expected Equity Return and Volatility

Here we provide the details of the approximte expression used for expected return in Equations (46) and (53). Starting from

$$\widehat{R}_{1,t} = \mathring{R}_{1,t} + \log\left(1+\frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t+1}+d_{1,t+1}}\right) - \log\left(1+\frac{\Delta\widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right),$$

and using $E_t\left[\log\left(1+X\right)\right] \approx \log\left(1+E_t\left[X\right]\right) - Var_t\left(X\right)$, expected equity return can be approximated as

$$E_t\left[\widehat{R}_{1,t}\right] \approx E_t\left[\mathring{R}_{1,t}\right] + \log\left(1+E_t\left[\frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t+1}+d_{1,t+1}}\right]\right) - Var_t\left(\frac{\Delta\widehat{P}_{1,t}}{\mathring{P}_{1,t}+d_{1,t+1}}\right) - \log\left(1+\frac{\Delta\widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right).$$

$\frac{\Delta\widehat{P}_{1,t}}{\mathring{P}_{1,t}+d_{1,t+1}}$ can be written as

$$\frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t}+d_{1,t+1}} = \frac{\Delta\widehat{P}_{1,t+1}}{\mathring{P}_{1,t}}\frac{\mathring{P}_{1,t}}{\mathring{P}_{1,t}+d_{1,t+1}},$$

where $\frac{\mathring{P}_{1,t}}{\mathring{P}_{1,t}+d_{1,t+1}}$ is the inverse of gross return on equity in the unconstrained case, and is given by

$$\begin{aligned}\frac{\mathring{P}_{1,t}}{\mathring{P}_{1,t}+d_{1,t+1}} &= \left(\frac{1}{\beta}\left(\frac{d_{1,t}}{d_{1,t+1}}\right)^{-\gamma}\frac{E_{t+1}\left[\sum_{i=t+1}^{T}\beta^{i-t}d_{1,i}^{1-\gamma}\right]}{E_t\left[\sum_{i=t+1}^{T}\beta^{i-t}d_{1,i}^{1-\gamma}\right]}\right)^{-1} \\ &= S\mathring{D}F_{t,t+1}\left[1+\frac{\mathring{P}_{1,t}-\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)}{\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)}\right],\end{aligned}$$

where $\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)$ is the value at time t of all dividends from time $t+1$ to time T given time-$t+1$ information

$$\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right) = \frac{1}{d_{1,t}^{-\gamma}}E_{t+1}\left[\sum_{i=t+1}^{T}\beta^{i-t}d_{1,i}^{1-\gamma}\right],$$

and $\mathring{P}_{1,t}$ is time-t price of all dividends from time $t+1$ to time T given time-t information

$$\mathring{P}_{1,t} = \frac{1}{d_{1,t}^{-\gamma}}E_t\left[\sum_{i=t+1}^{T}\beta^{i-t}d_{1,i}^{1-\gamma}\right].$$

Hence, $\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}+d_{1,t+1}}$ can be written as

$$\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}+d_{1,t+1}} = S\mathring{D}F_{t,t+1}\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}}\left[1+\frac{\mathring{P}_{1,t}-\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)}{\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)}\right] \tag{87}$$

In the case of a single minimum-payout at the terminal date T, the SDF between time t and $t+1$ is unaffected by the presence of the constraint, i.e. $\widehat{SDF}_{t,t+1} = S\mathring{D}F_{t,t+1}$, and we get

$$E_t\left[\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}+d_{1,t+1}}\right] = \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}} + E_t\left[\widehat{SDF}_{t,t+1}\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}}\frac{\mathring{P}_{1,t}-\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)}{\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)}\right], \tag{88}$$

because

$$E_t\left[S\mathring{D}F_{t,t+1}\Delta \widehat{P}_{1,t+1}\right] = \Delta \widehat{P}_{1,t}.$$

Thus, using Equation (88) the expected equity return can be written as in Equation 46.

In the case of non-zero minimum-payouts at every date, the SDF between time t and $t+1$ deviates from its unconstrained case, and is given by

$$\begin{aligned}\widehat{SDF}_{t,t+1} &= \beta\left(\frac{d_{1,t+1}}{d_{1,t}}\right)^{-\gamma} f_{\omega_t}^{-\gamma} \\ &= S\mathring{D}F_{t,t+1} + \beta\left(\frac{d_{1,t+1}}{d_{1,t}}\right)^{-\gamma}\left(f_{\omega_t}^{-\gamma}-1\right).\end{aligned}$$

Therefore, $E_t\left[\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}+d_{1,t+1}}\right]$ can be written as

$$\begin{aligned}E_t\left[\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}+d_{1,t+1}}\right] &= \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}} + E_t\left[\widehat{SDF}_{t,t+1}\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}}\frac{\mathring{P}_{1,t}-\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)}{\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)}\right] \\ &\quad - E_t\left[\beta\left(\frac{d_{1,t+1}}{d_{1,t}}\right)^{-\gamma}\left(f_{\omega_t}^{-\gamma}-1\right)\frac{\Delta \widehat{P}_{1,t+1}}{\mathring{P}_{1,t}\left(\mathcal{F}_{t+1}\right)}\right],\end{aligned} \tag{89}$$

Using Equation (89), the expected equity return can be written as in Equation (53).

To obtain an approximate expression for volatility, we approximate equity return as

$$\widehat{R}_{t,1} = \mathring{R}_{1-t} + \frac{\Delta \widehat{P}_{1,t+1}}{d_{1,t+1} + \mathring{P}_{1,t+1}} - \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}}.$$

The $E_t\left[\widehat{R}_{1,t}^2\right]$ can then be written as

$$\begin{aligned}
E_t\left[\widehat{R}_{1,t}^2\right] &= E_t\left[\mathring{R}_{1,t}^2\right] + E_t\left[\left(\frac{\Delta \widehat{P}_{1,t+1}}{d_{1,t+1} + \mathring{P}_{1,t+1}} - \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right)^2\right] + 2E_t\left[\mathring{R}_{1,t}\left(\frac{\Delta \widehat{P}_{1,t+1}}{d_{1,t+1} + \mathring{P}_{1,t+1}} - \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right)\right] \\
&= E_t\left[\mathring{R}_{1,t}^2\right] + E_t\left[\left(\frac{\Delta \widehat{P}_{1,t+1}}{d_{1,t+1} + \mathring{P}_{1,t+1}} - \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right)^2\right] + 2E_t\left[\mathring{R}_{1,t}\right] E_t\left[\frac{\Delta \widehat{P}_{1,t+1}}{d_{1,t+1} + \mathring{P}_{1,t+1}} - \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right] \\
&+ 2Cov_t\left(\mathring{R}_{1,t}, \left(\frac{\Delta \widehat{P}_{1,t+1}}{d_{1,t+1} + \mathring{P}_{1,t+1}} - \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right)\right),
\end{aligned}$$

and the square of expected return, $E_t\left[\widehat{R}_{1,t}\right]^2$, can be written as

$$E_t\left[\widehat{R}_{1,t}\right]^2 = E_t\left[\mathring{R}_{1,t}\right]^2 + E_t\left[\frac{\Delta \widehat{P}_{1,t+1}}{d_{1,t+1} + \mathring{P}_{1,t+1}} - \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right]^2 + 2E_t\left[\mathring{R}_{1,t}\right]\left(E_t\left[\frac{\Delta \widehat{P}_{1,t+1}}{d_{1,t+1} + \mathring{P}_{1,t+1}}\right] - \frac{\Delta \widehat{P}_{1,t}}{\mathring{P}_{1,t}}\right)$$

Thus, the equity return volatility can be approximately written as in Equation (47).

C.2 Minimum-Payout Model

For minimum-payout model, the constrained institution satisfies its constraint by holding riskfree asset at time $T-1$, and its allocations can be written as

$$\tilde{\theta}^0_{1,T-1} = \begin{cases} 0, & \text{if } \omega_{1,T-1} > \frac{c_T^{min}}{d_{1,T,d}} \\ c_T^{min} - \frac{\omega_{1,T-1} d_{1,T,u} - c_T^{min}}{d_{1,T,u} - d_{1,T,d}} d_{1,T,d}, & \text{if } \frac{c_T^{min}}{d_{1,T,u}} < \omega_{1,T-1} \leq \frac{c_T^{min}}{d_{1,T,d}} \\ c_T^{min}, & \text{if } \omega_{1,T-1} \leq \frac{c_T^{min}}{d_{1,T,u}} \end{cases} \tag{90}$$

$$\tilde{\theta}^1_{1,T-1} = \begin{cases} \omega_{1,T-1}, & \text{if } \omega_{1,T-1} > \frac{c_T^{min}}{d_{1,T,d}} \\ \frac{\omega_{1,T-1} d_{1,T,u} - c_T^{min}}{d_{1,T,u} - d_{1,T,d}}, & \text{if } \frac{c_T^{min}}{d_{1,T,u}} < \omega_{1,T-1} \leq \frac{c_T^{min}}{d_{1,T,d}} \\ 0, & \text{if } \omega_{1,T-1} \leq \frac{c_T^{min}}{d_{1,T,u}} \end{cases} \tag{91}$$

And for the unconstrained institution, the optimal allocations can be written using the market clearing conditions.

The constrained institution's payouts at time T are given by

$$\tilde{c}_{1,T} = \tilde{\theta}^0_{1,T-1} + \tilde{\theta}^1_{1,T-1} d_{1,T} = \omega_{1,T} d_{1,T}, \tag{92}$$

where

$$\tilde{\omega}_{1,T} = \begin{cases} \omega_{1,T-1}, & \text{if } \omega_{1,T-1} > \frac{c_T^{min}}{d_{1,T,d}} \\ \omega_T^{min} + \frac{\omega_{1,T-1} d_{1,T,u} - c_T^{min}}{d_{1,T,u} - d_{1,T,d}} \left(1 - \frac{d_{1,T,d}}{d_{1,T}}\right), & \text{if } \frac{c_T^{min}}{d_{1,T,u}} < \omega_{1,T-1} \leq \frac{c_T^{min}}{d_{1,T,d}} \\ \omega_T^{min}, & \text{if } \omega_{1,T-1} \leq \frac{c_T^{min}}{d_{1,T,u}}. \end{cases} \tag{93}$$

And the unconstrained institution's consumption is given by

$$\tilde{c}_{0,T} = d_{1,T} - c_{1,T} = \frac{1-\omega_{1,T}}{1-\omega_{1,T-1}} \mathring{c}_{0,T} = \mathring{c}_{0,T} \tilde{f}_{\omega_T}, \tag{94}$$

where

$$\tilde{f}_{\omega_T} = \frac{1-\tilde{\omega}_{1,T}}{1-\omega_{1,T-1}}, \tag{95}$$

which again has a two-factor structure as in the case of the funding-ratio constraint.

The Lagrange multipliers for the payout constraint can be solved using kernel conditions, which equate investor's subjective prices for traded assets

$$\beta E_{T-1}\left[\frac{\tilde{c}_{0,T}^{-\gamma}}{c_{0,T-1}^{-\gamma}}d_{n,T}\right] = \beta E_{T-1}\left[\frac{\tilde{c}_{1,T}^{-\gamma}+\tilde{\lambda}_T^{mw}}{c_{1,T-1}^{-\gamma}}d_{n,T}\right]. \tag{96}$$

Solving the two kernel conditions for $n = \{0, 1\}$, we get

$$E_{T-1}\left[\tilde{\lambda}_T^{mw}\right] = \frac{c_{1,T-1}^{-\gamma}}{c_{0,T-1}^{-\gamma}}E_{T-1}\left[\tilde{c}_{0,T}^{-\gamma}\right] \tag{97}$$

$$E_{T-1}\left[\tilde{\lambda}_T^{mw}d_{1,T}\right] = \frac{c_{1,T-1}^{-\gamma}}{c_{0,T-1}^{-\gamma}}E_{T-1}\left[\tilde{c}_{0,T}^{-\gamma}d_{1,T}\right] - E_{T-1}\left[\tilde{c}_{1,T}^{-\gamma}d_{1,T}\right]. \tag{98}$$

As in the case of funding-ratio constraint, for only a single minimum-payout, the payouts across states for $t < T$ remain unaffected by the constraint and are given by Equation (70). Similarly, allocations for $t < T - 1$ can also be obtained using Equation (78).

Thus, the log of SDF between time 0 and T is given by

$$\tilde{m}_{0,T} = T\log\beta - \gamma\log\left(\frac{1-\tilde{\omega}_{1,T}}{1-\tilde{\omega}_{1,0}}\right) - \gamma\log\left(\frac{d_{1,T}}{d_{1,0}}\right) \tag{99}$$

C.3 Surplus-Payout Utility Model

For surplus-payout utility model the optimal allocations at time $T - 1$ can be obtained by solving kernel conditions for the two assets, which yield

$$\breve{\theta}_{1,T-1}^{0} = c_T^{min}\left(1-\omega_{1,T-1}\right) \tag{100}$$

$$\breve{\theta}_{1,T-1}^{1} = \omega_{1,T-1}, \tag{101}$$

and the optimal payout at time T is given by

$$\breve{c}_{1,T} = c_T^{min} + (d_{1,T} - c_T^{min})\omega_{1,T-1}. \tag{102}$$

Thus, the log of SDF between time 0 and T is given by

$$\breve{m}_{0,T} = T \log \beta - \gamma \log \left(\frac{d_{1,T} - c_T^{min}}{d_{1,0}} \right), \tag{103}$$

which preserves a linear one-factor structure of CCAPM.

D Numerical Solution

For numerical results, we solve this system numerically using the method developed in Dumas and Lyasoff (2012). We shift first order condition for fund-withdrawals, budget constraint, and complementary slackness condition for minimum-withdrawal constraint by one period in time. The resulting system of equations is then simultaneously solved for fund-withdrawals at time t and portfolio choices at time $t-1$, in terms of the unconstrained institution's endowment-share at time $t-1$, which serves as our state variable.

After shifting Equations 9, 10, and 13 by one period in time, the resulting system of equations

$$\begin{aligned}
\lambda^{bc}_{m,t+1,\eta} &= u'_m(c_{m,t+1,\eta}) + \lambda^{mw}_{m,t+1,\eta} \\
c_{m,t+1,\eta} + F_{m,t+1,\eta} &= \theta^0_{m,t,\xi} + \theta^1_{m,t,\xi}(d_{1,t+1,\eta} + P_{1,t+1,\eta}) \\
\frac{\beta E_t\left[\lambda^{bc}_{0,t+1,\eta}\right]}{\lambda^{bc}_{0,t,\xi} - \lambda^{fr}_{0,t,\xi}} &= \frac{\beta E_t\left[\lambda^{bc}_{1,t+1,\eta}\right]}{\lambda^{bc}_{1,t,\xi} - \lambda^{fr}_{1,t,\xi}} \\
\frac{\beta E_t\left[\lambda^{bc}_{0,t+1,\eta}(P_{1,t+1,\eta} + d_{1,t+1,\eta})\right]}{\lambda^{bc}_{0,t,\xi} - \lambda^{fr}_{0,t,\xi}} &= \frac{\beta E_t\left[\lambda^{bc}_{1,t+1,\eta}(P_{1,t+1,\eta} + d_{1,t+1,\eta})\right]}{\lambda^{bc}_{1,t,\xi} - \lambda^{fr}_{1,t,\xi}} \\
\lambda^{mw}_{m,t+1,\eta}(c_{m,t+1,\eta} - c^{min}_{m,t+1,\eta}) &= 0 \\
\lambda^{fr}_{m,t,\xi}\left[\theta^0_{m,t,\xi}P_{0,1} + \theta^1_{m,t,\xi}P_{1,t,\xi} - \phi^{min}_m F^{min}_{t,\xi}\right] &= 0 \\
\theta^0_{0,t,\xi} + \theta^0_{1,t,\xi} &= 0 \\
\theta^1_{0,t,\xi} + \theta^1_{1,t,\xi} &= 1
\end{aligned} \tag{104}$$

where ξ denotes a node (state of the world) at time t, ξ^+ denotes the set of nodes at time $t+1$ that can be reached from ξ, $\eta \in \xi^+$, and $F_{m,t,\xi}$ is the mth investor's exiting wealth at time t, which is computed recursively as

$$F_{m,t,\xi} = \frac{\beta E_t\left[(\lambda^{bc}_{m,t+1,\eta})(F_{m,t+1,\eta} + c_{m,t+1,\eta})\right]}{\lambda^{bc}_{m,t,\xi} - \lambda^{fr}_{1,t,\xi}} \tag{105}$$

starting with $F_{m,T,\eta} = 0$.

For numerical convenience, time t consumption of mth investor in a given state of the world, $c_{m,t,\xi}$, can be written in terms of the investor's share of aggregate consumption, $\omega_{m,t,\xi}$ defined as

$$\omega_{m,t,\xi} = \frac{c_{m,t,\xi}}{d_{t,\xi}} \tag{106}$$

where $d_{t,\xi}$ is the aggregate dividend, which in equilibrium is equal to the aggregate fund-withdrawals

$$\sum_{m=0}^{1} c_{m,t,\xi} = d_{t,\xi}$$
$$\Rightarrow \sum_{m=0}^{1} \omega_{m,t,\xi} = 1. \tag{107}$$

We take $\omega_{0,t,\xi}$, endowment share of the constrained institution, as our state variable.

Then:

- We start at the terminal time T, where asset prices, $P_{n,T}$, and portfolio allocations, $\theta^n_{m,T}$, are known to be zero for all assets in all states of the world;
- Express portfolio allocations at node ξ at time $T-1$, $\theta^n_{m,T-1,\xi}$, constraint multipliers, $\lambda^{bc}_{m,T-1,\xi}$, $\lambda^{mw}_{m,T,\eta}$, $\lambda^{fr}_{m,T-1,\xi}$, and fund-withdrawals at nodes η at time T, $c_{m,T,\eta}$, in terms of the time-$T-1$ endogenous state variable, $\omega_{0,T-1,\xi}$, time-T dividends, $d_{n,T,\eta}$, and the probabilities of time-T dividends, $p_{T,\eta}$;
- Determine asset prices, $P_{n,T-1,\xi}$, and exiting wealths, $F_{m,T-1,\xi}$, to the two institutions using

$$P_{0,T-1,\xi} = \frac{\beta E_{T-1}\left[\lambda^{bc}_{m,T,\eta}\right]}{\lambda^{bc}_{m,T-1,\xi} - \lambda^{fr}_{m,T-1,\xi}} \tag{108}$$
$$P_{n,T-1,\xi} = \frac{\beta E_{T-1}\left[(\lambda^{bc}_{m,T,\eta})(P_{n,T,\eta} + d_{n,T,\eta})\right]}{\lambda^{bc}_{m,T-1,\xi} - \lambda^{fr}_{m,T-1,\xi}} \tag{109}$$
$$F_{m,T-1,\xi} = \frac{\beta E_{t}\left[(\lambda^{bc}_{m,T,\eta})(F_{m,T,\eta} + c_{m,T,\eta})\right]}{\lambda^{bc}_{m,T-1,\xi} - \lambda^{fr}_{m,T-1,\xi}} \tag{110}$$

  as a function of $\omega_{0,T-1,\xi}$;

- Repeat this procedure for all time-$(T-1)$ nodes, ξ, and the corresponding time-T nodes, $\eta \in \xi^+$, and determine $P_{n,T-1,\xi}$ and $F_{m,T-1,\xi}$ as a function of $\omega_{0,T-1,\xi}$;
- Interpolate asset price, $P_{n,T-1,\xi}$, and exiting wealth, $F_{m,T-1,\xi}$, functions over all values of $omega_{0,t,\xi}$ for all nodes ξ;
- Repeat this procedure for all times $t = T-1, \cdots, 0$;
- Moving backward in this way, we obtain all future portfolio allocations, $\theta^n_{m,0,\xi}$, fund-withdrawals, $c_{m,1,\eta}$, and asset prices, $P_{n,0,\xi}$, and invested wealths of the two institutions, $F_{m,0,\xi}$, in terms of time-0 endowment share of the unconstrained institution, $\omega_{0,0,\xi}$. Then, we use exogenously specified initial portfolio positions of the two institutions to solve for time 0 endowment shares, $\omega m, 0, \xi$, using

$$\omega_{m,0,\xi} d_{0,\xi} + F_{m,0,\xi} - \sum_{n=1}^{1} \theta^n_{m,-1}(P_{n,0,\xi} + d_{n,0,\xi}) = 0 \tag{111}$$

$$\lambda^{mw}_{1,1,\eta}(c_{1,1,\eta} - c^{min}_{1,1,\eta}) = 0 \tag{112}$$

$$\lambda^{fr}_{1,0,\xi}\left[\theta^0_{m,0,\xi} P_{0,1} + \theta^1_{m,0,\xi} P_{1,0,\xi} - \phi^{min}_m F^{min}_{0,\xi}\right] = \tag{113}$$

where $\theta^n_{m,-1}$ denotes initial portfolio positions of investors prior to any trades at time 0.

To handle inequality constraints, we first solve the system of equations given in Equation (104) for the unconstrained case (with multipliers for all the inequality constraints set to zero), and check if the resulting solution satisfies the constraint.

That is, we solve the following system of equations

$$
\begin{aligned}
\lambda^{bc}_{m,t+1,\eta} &= u'_m(c_{m,t+1,\eta}) \\
c_{m,t+1,\eta} + F_{m,t+1,\eta} &= \theta^0_{m,t,\xi} + \theta^1_{m,t,\xi}(d_{1,t+1,\eta} + P_{1,t+1,\eta}) \\
\frac{\beta E_t\left[\lambda^{bc}_{0,t+1,\eta}\right]}{\lambda^{bc}_{0,t,\xi}} &= \frac{\beta E_t\left[\lambda^{bc}_{1,t+1,\eta}\right]}{\lambda^{bc}_{1,t,\xi}} \\
\frac{\beta E_t\left[\lambda^{bc}_{0,t+1,\eta}(P_{1,t+1,\eta} + d_{1,t+1,\eta})\right]}{\lambda^{bc}_{0,t,\xi}} &= \frac{\beta E_t\left[\lambda^{bc}_{1,t+1,\eta}(P_{1,t+1,\eta} + d_{1,t+1,\eta})\right]}{\lambda^{bc}_{1,t,\xi}} \\
\theta^0_{0,t,\xi} + \theta^0_{1,t,\xi} &= 0 \\
\theta^1_{0,t,\xi} + \theta^1_{1,t,\xi} &= 1,
\end{aligned}
\tag{114}
$$

for $\{c_{m,t+1,\eta}, \theta^n_{m,t,\xi}, P_{n,1,\xi}\}$ and test if the resulting solution satisfies the imposed inequality constraint

$$
\begin{aligned}
c_{m,t+1,\eta} &\geq c^{min}_{m,t+1,\eta} \\
\theta^0_{m,t,\xi} P_{0,1,\xi} + \theta^1_{m,t,\xi} P_{1,t,\xi} - \phi^{min}_m F^{min}_{t,\xi} &\geq 0.
\end{aligned}
$$

If the constraint is satisfied, the complete market solution is kept as the solution for the constrained problem. If the constraint is not satisfied, the constraint is made binding by adding the constraint equation to the system of equations that are to be solved. For instance, in the case of funding-ratio constraint, if the constraint is not satisfied, we solve the following system of equations

$$
\begin{aligned}
\lambda^{bc}_{m,t+1,\eta} &= u'_m(c_{m,t+1,\eta}) \\
c_{m,t+1,\eta} + F_{m,t+1,\eta} &= \theta^0_{m,t,\xi} + \theta^1_{m,t,\xi}(d_{1,t+1,\eta} + P_{1,t+1,\eta}) \\
\frac{\beta E_t\left[\lambda^{bc}_{0,t+1,\eta}\right]}{\lambda^{bc}_{0,t,\xi} - \lambda^{fr}_{0,t,\xi}} &= \frac{\beta E_t\left[\lambda^{bc}_{1,t+1,\eta}\right]}{\lambda^{bc}_{1,t,\xi} - \lambda^{fr}_{1,t,\xi}} \\
\frac{\beta E_t\left[\lambda^{bc}_{0,t+1,\eta}(P_{1,t+1,\eta} + d_{1,t+1,\eta})\right]}{\lambda^{bc}_{0,t,\xi} - \lambda^{fr}_{0,t,\xi}} &= \frac{\beta E_t\left[\lambda^{bc}_{1,t+1,\eta}(P_{1,t+1,\eta} + d_{1,t+1,\eta})\right]}{\lambda^{bc}_{1,t,\xi} - \lambda^{fr}_{1,t,\xi}} \\
\theta^0_{m,t,\xi} P_{0,1} + \theta^1_{m,t,\xi} P_{1,t,\xi} - \phi^{min}_m F^{min}_{t,\xi} &= 0 \\
\theta^0_{0,t,\xi} + \theta^0_{1,t,\xi} &= 0 \\
\theta^1_{0,t,\xi} + \theta^1_{1,t,\xi} &= 1,
\end{aligned}
\tag{115}
$$

for $\{c_{m,t+1,\eta}, \theta^n_{m,t,\xi}, P_{n,1,\xi}, \lambda^{fr}_{1,t,\xi}\}$.

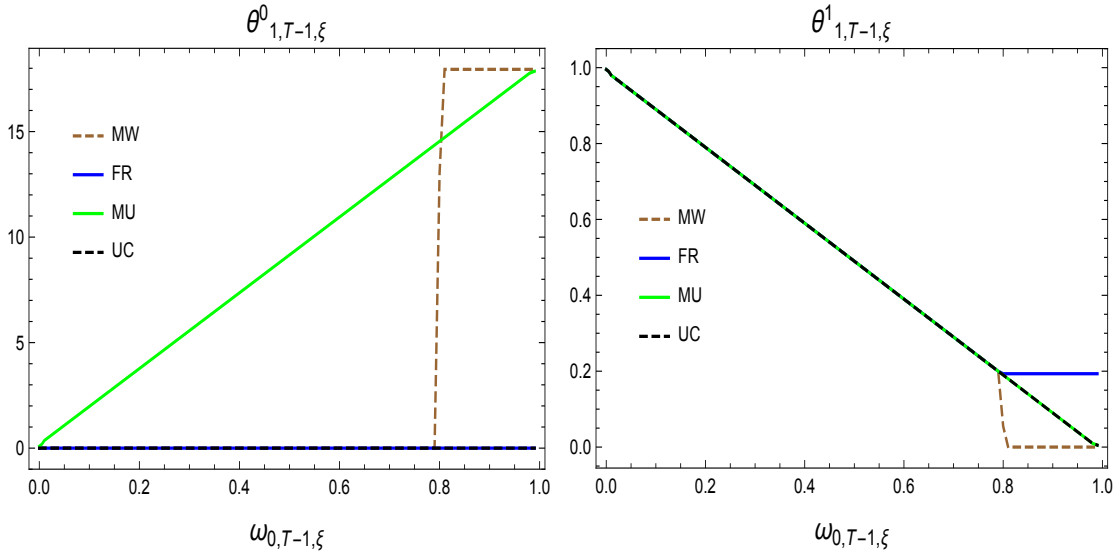
Similarly, in the case of minimum-withdrawals, if the constraint is not satisfied, we solve

$$\begin{aligned}
\lambda^{bc}_{m,t+1,\eta} &= u'_m(c_{m,t+1,\eta}) + \lambda^{mw}_{m,t+1,\eta} \\
c_{m,t+1,\eta} + F_{m,t+1,\eta} &= \theta^0_{m,t,\xi} + \theta^1_{m,t,\xi}(d_{1,t+1,\eta} + P_{1,t+1,\eta}) \\
\frac{\beta E_t\left[\lambda^{bc}_{0,t+1,\eta}\right]}{\lambda^{bc}_{0,t,\xi}} &= \frac{\beta E_t\left[\lambda^{bc}_{1,t+1,\eta}\right]}{\lambda^{bc}_{1,t,\xi}} \\
\frac{\beta E_t\left[\lambda^{bc}_{0,t+1,\eta}(P_{1,t+1,\eta} + d_{1,t+1,\eta})\right]}{\lambda^{bc}_{0,t,\xi}} &= \frac{\beta E_t\left[\lambda^{bc}_{1,t+1,\eta}(P_{1,t+1,\eta} + d_{1,t+1,\eta})\right]}{\lambda^{bc}_{1,t,\xi}} \\
\lambda^{mw}_{m,t+1,\eta}(c_{m,t+1,\eta} - c^{min}_{m,t+1,\eta}) &= 0 \\
\theta^0_{0,t,\xi} + \theta^0_{1,t,\xi} &= 0 \\
\theta^1_{0,t,\xi} + \theta^1_{1,t,\xi} &= 1
\end{aligned} \tag{116}$$

for $\{c_{m,t+1,\eta}, \theta^n_{m,t,\xi}, P_{n,1,\xi}, \lambda^{mw}_{1,t+1,\eta}\}$.

Figure 1: Constrained institution's asset allocations at time $T-1$ and payouts at time T.

(a) Bond (left panel) and equity (right panel) allocations at a single node ξ, which corresponds to the lowest aggregate dividend realisation, at time $T-1$, as a function of the unconstrained institution's share of aggregate endowment at time $T-1$. FR and UC curves always overlap in the left panel, and MU and UC curves always overlap in the right panel.



(b) The left panels shows payouts at a single node η^-, which corresponds to a negative innovation in the aggregate dividend starting from node ξ at time $T-1$, and the right panel shows payouts across nodes at time T, for an initial endowment share of 0.85 for the unconstrained institution. Withdrawals for MW model (brown curve) overlap with c^{min} (red curve) in the constrained region, and overlap with FR and UC models (blue and black curves) in the unconstrained region, making it hard to distinguish.

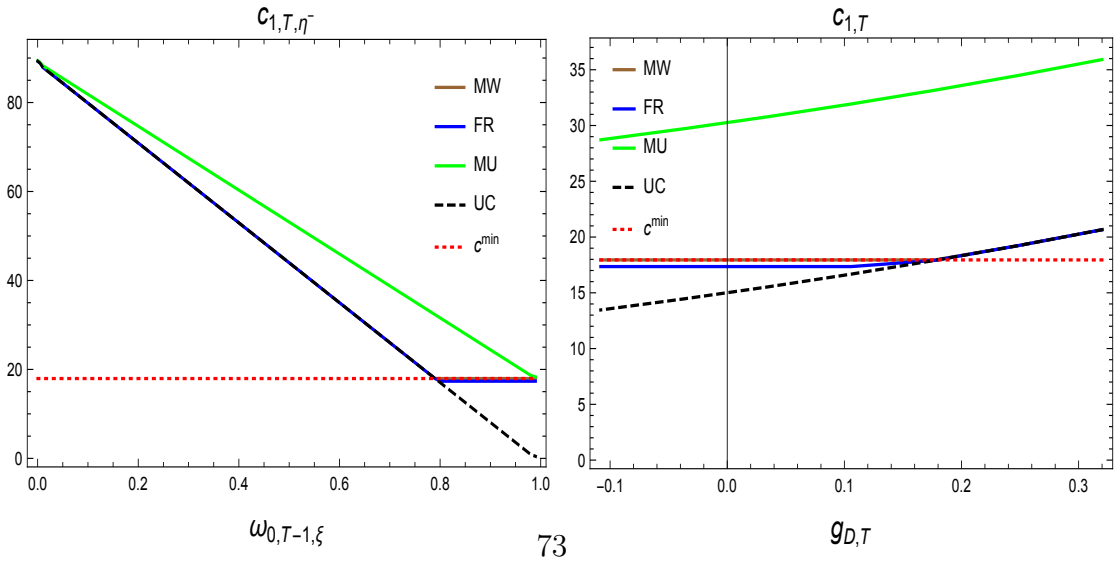
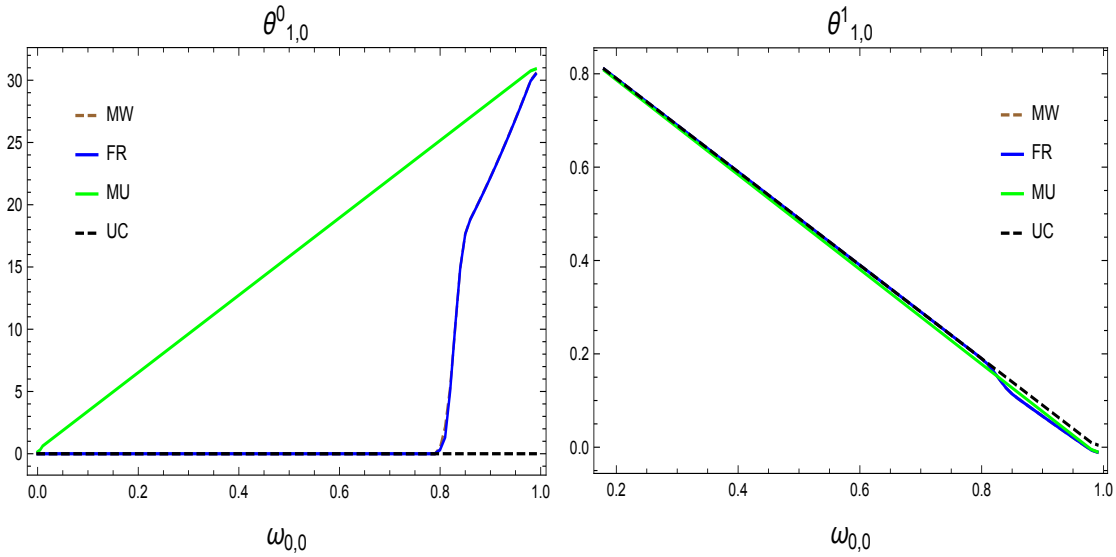


Figure 2: Constrained institution's allocations, payouts, and invested wealth at time 0. At time 0, MW and FR models (brown and blue curves) behave almost identically, making the two curves almost indistinguishable.

(a) Constrained institution's bond (left panel) and equity (right panel) allocations.



(b) Constrained institution's invested wealth (left panel), and payout/invested wealth ratio (cay) at time 0. The payout/wealth ratio decreases by upto about 80%.

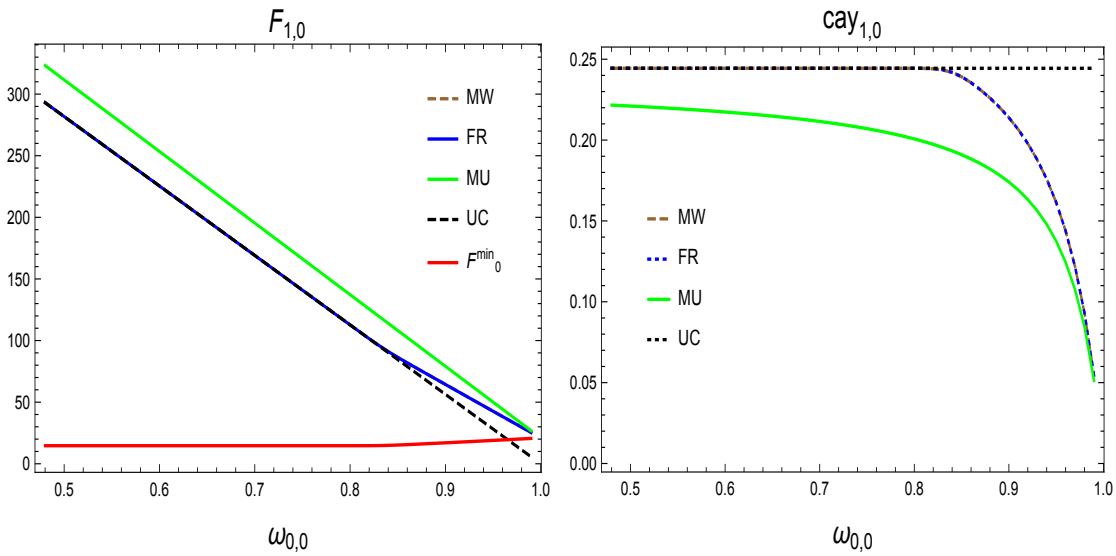


Figure 3: Comparison of the SDF for minimum-payout constraint, funding-ratio constraint, and surplus-payout utility examples. The left panel shows the SDF for the lowest aggregate dividend realization at time T for all possible values of the initial endowment share, $\omega_{0,0}$. The right Panel shows the log of SDF for different realizations of the aggregate dividend at time T, for an initial endowment share of $\omega_{0,0} = 0.85$, except for the dotted blue curve, which corresponds to an initial endowment share of 0.98. $g_{D,T}$ denotes the growth rate of aggregate dividend between time 0 and T. The SDF increases by upto about 45% for the funding-ratio model. The SDF behaves almost identically for FR and MW models, making the two curves almost indistinguishable.

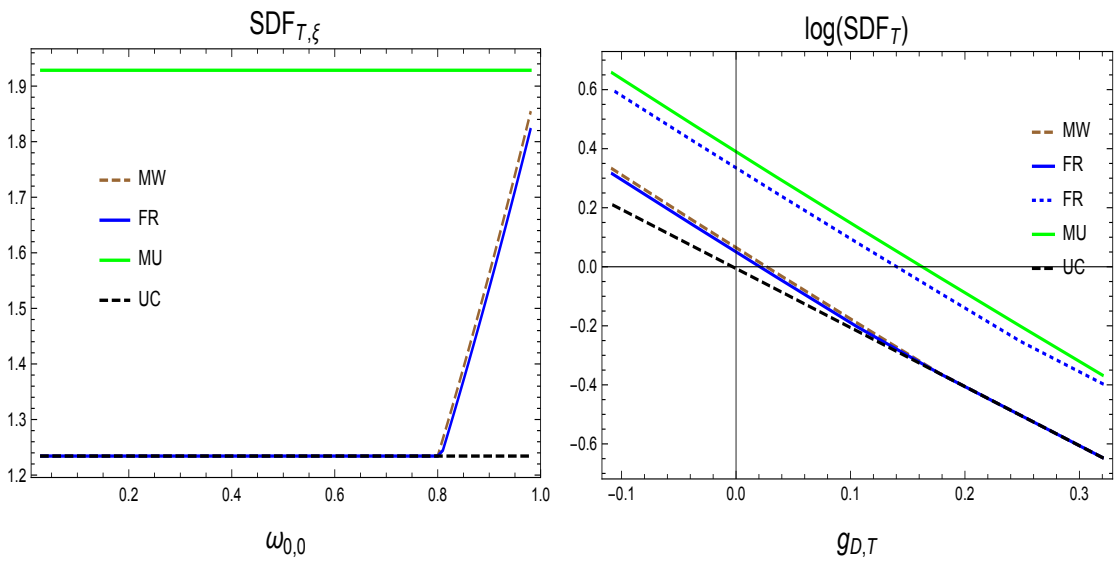
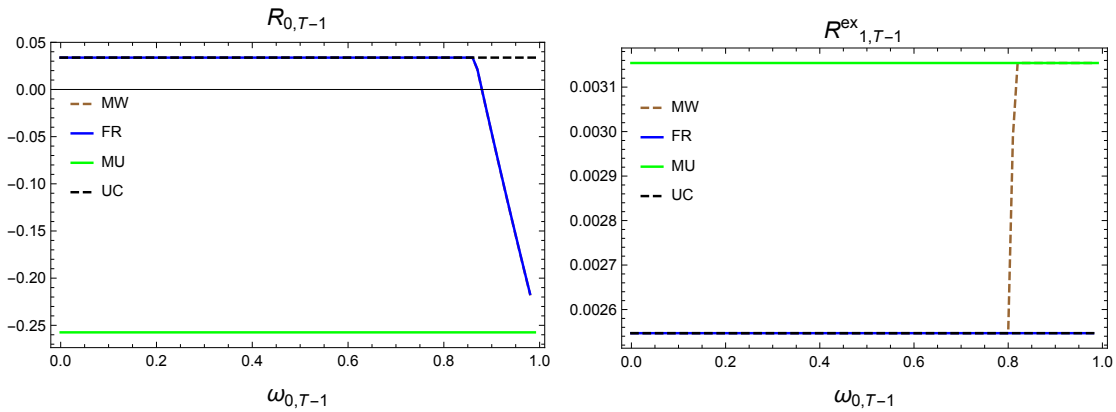
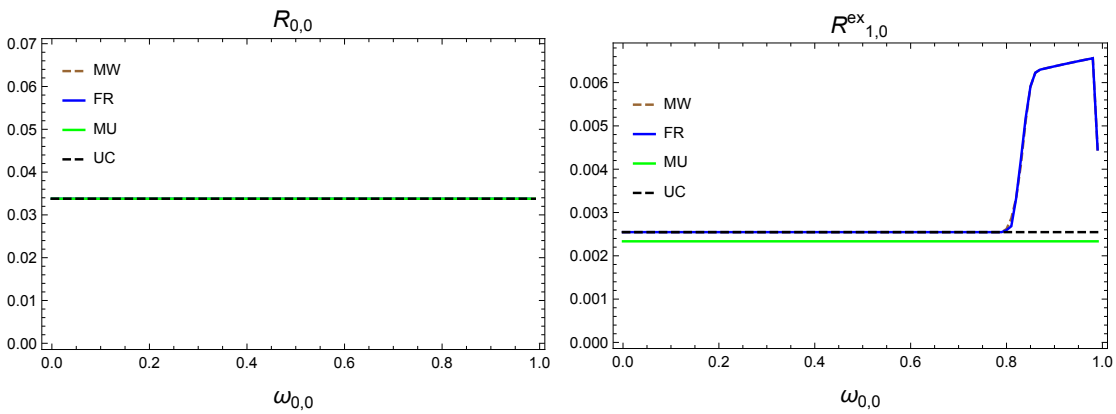


Figure 4: Expected bond return, equity premium, equity premium volatility, and Sharpe ratio. Time-$T-1$ variables are plotted for the lowest aggregate dividend realization at time $T-1$, as a function of time-$T-1$ endowment share. Time-0 quantities are plotted as a function of time-0 endowment share. For the sake of simplicity, we denote expected returns simply as $R_{n,t}$, and drop the expectation operator.

(a) Expected bond return (left panel) and equity premium (right panel) at time $T-1$. The bond return decreases by upto about 160%, while equity premium remains unchanged for the funding-ratio model.



(b) Expected bond return (left panel) and equity premium (right panel) at time 0. At time 0 equity premium increases by about 650%, while bond return remains unchanged for the funding-ratio model.



(c) Equity premium volatility (left panel) and Sharpe ratio (right panel) at time 0. The equity premium volatility decreases by upto about 8%, while the Sharpe ratio increases by upto about 275% for the funding-ratio model. The equity price increases by upto about 10% for the funding-ratio model.

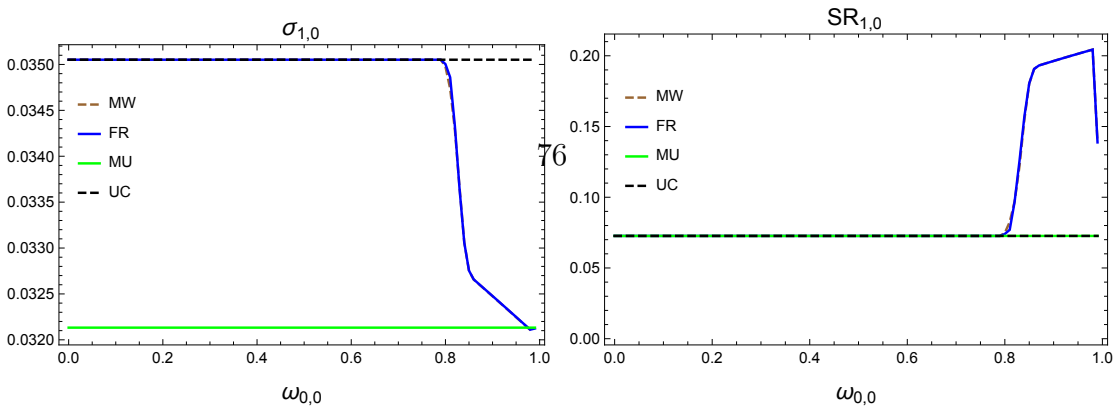
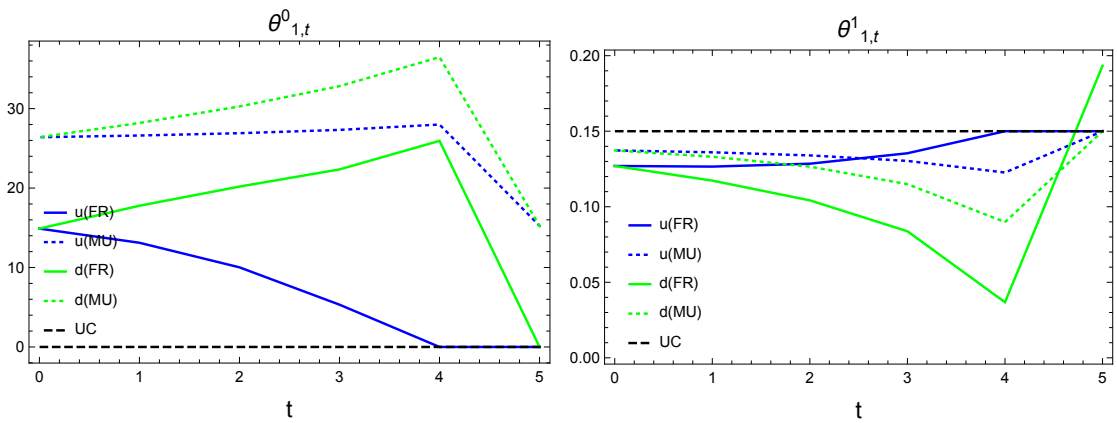


Figure 5: Time evolution of the constrained institution's asset allocations and equity price. Blue (green) curve corresponds to the best (worst) path of the aggregate dividend with all positive (negative) innovations. Solid (dashed) curves correspond to funding-ratio (surplus-payout utility) models, and the dashed black curve corresponds to the unconstrained model. The initial endowment share of the unconstrained institution is assumed to be 0.85 for all curves.

(a) Evolution of the constrained institution's bond (left panel) and equity (right panel) allocation.



(b) Evolution of the equity price for funding-ratio (left panel) and surplus utility (right panel) examples. The equity price is plotted in units of the equity price in the unconstrained case.

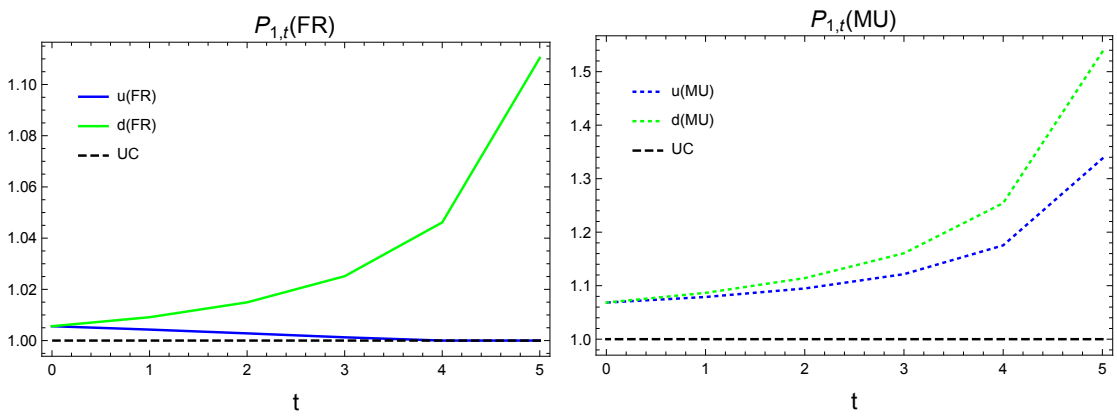
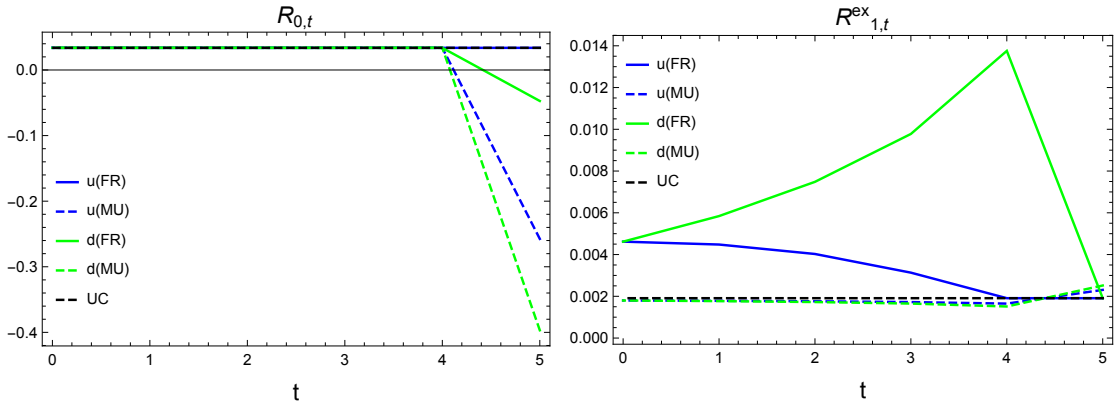


Figure 6: Time evolution of bond return, equity premium, volatility, Sharpe ratio. Blue (green) curve corresponds to the best (worst) path of the aggregate dividend with all positive (negative) innovations. Solid (dashed) curves correspond to funding-ratio (surplus-payout utility) models, and the dashed black curve corresponds the unconstrained model. As all quantities are path-independent in the unconstrained model, we only show them along the worst path for the unconstrained model. The initial endowment share is assumed to be 0.85 for all curves. For the sake of simplicity, we denote expected returns simply as $R_{n,t}$, and drop the expectation operator.

(a) Evolution of bond return (left panel) and equity premium (right panel). Bond return decreases by upto about 200%, while equity premium increases by upto about 600% along the worst possible path in the funding-ratio model.



(b) Evolution of equity premium volatility (left panel) and Sharpe ratio (right panel). The equity premium volatility decreases by upto about 25%, while the Sharpe ratio increases by upto about 900%.

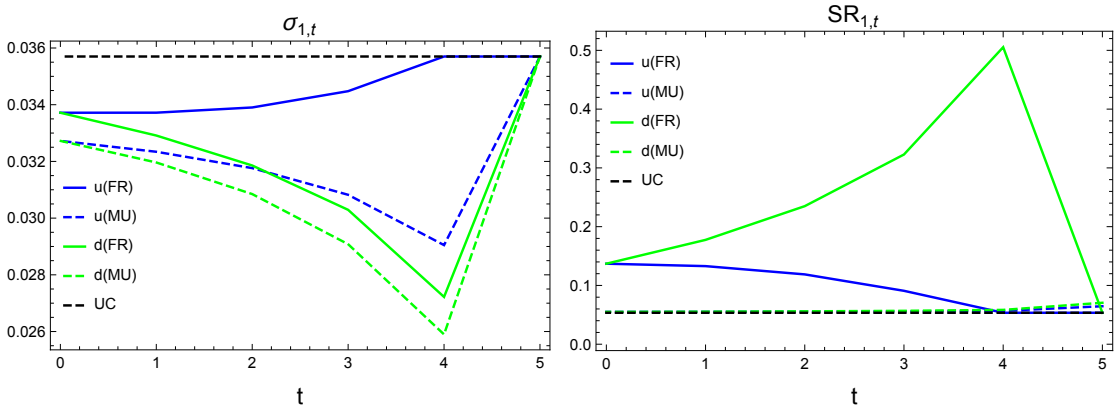
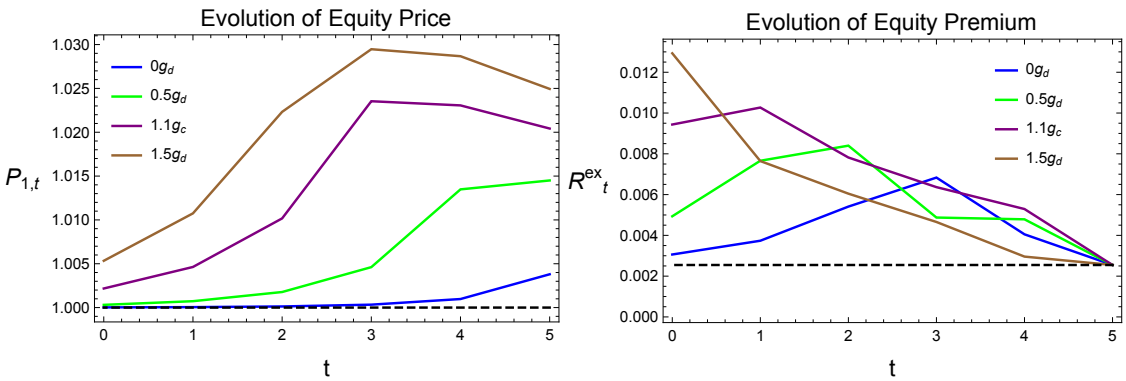
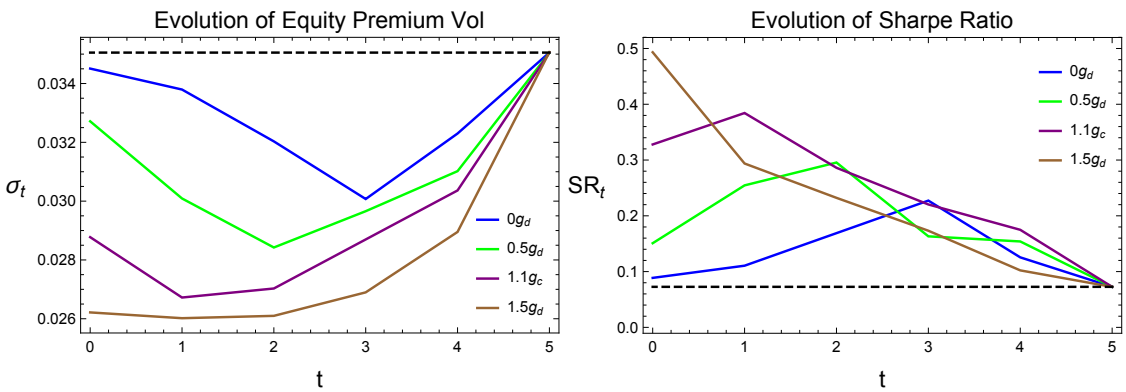


Figure 7: Plots for time-varying minimum-payouts. Minimum-payouts are assumed to grow at a constant rate (Equation (20)), and the results are shown for four different growth rates of minimum-payouts— $0g_d$, $0.5g_d$, $1.1g_d$, $1.5g_d$, where g_d is the expected growth rate of aggregate dividend. The dashed black line corresponds to the unconstrained case.

(a) Evolution of equity price in units of equity price in the unconstrained setting (left panel), and equity premium (right panel) along the worst possible path of the aggregate dividend, starting from an initial endowment share of 0.85 for the unconstrained institution. For the sake of simplicity, we denote expected equity premium simply as $R^{ex}_{1,t}$, and drop the expectation operator.



(b) Evolution of equity premium volatility (left panel), and equity Sharpe ratio (right panel) along the worst possible path of the aggregate dividend, starting from an initial endowment share of 0.85 for the unconstrained institution.



(c) Term structures of riskfree rates for two different values of the unconstrained institution's endowment share, which is 0.87 for the left panel and 0.90 for the right panel.

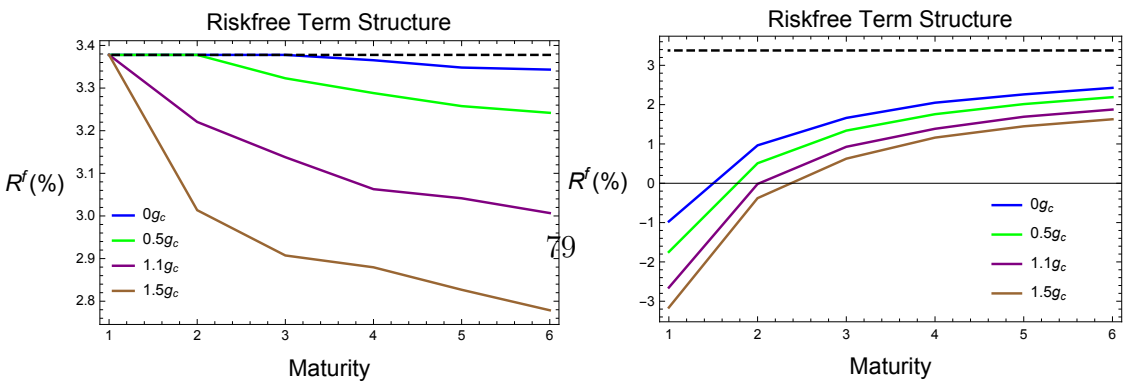
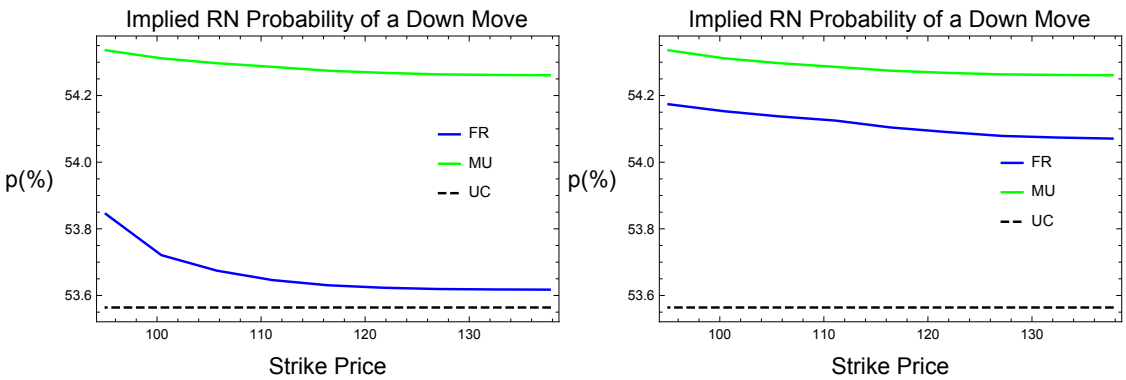


Figure 8: Implied risk-neutral probability plots. The dashed black line corresponds to the unconstrained case.

(a) Implied risk-neutral probability of a down move as a function of strike price for two different values of the unconstrained institution's endowment share. The left panel corresponds to an initial endowment share of 0.84 and and the right panel corresponds to an initial endowment share of 0.87 for the unconstrained institution.



(b) The left panel shows the slope of implied probability curve with the equity premium. The right panel shows the implied probabilities of put options with the same level of 'moneyness' as a function of maturity for different profiles of minimum-payouts. "ST" (short term) curve corresponds to the case when the constrained institution only has non-zero minimum-payouts at dates 2 and 3. In the right panel, the strike price for options at all maturities is set equal to the maximum dividend at the corresponding maturity.

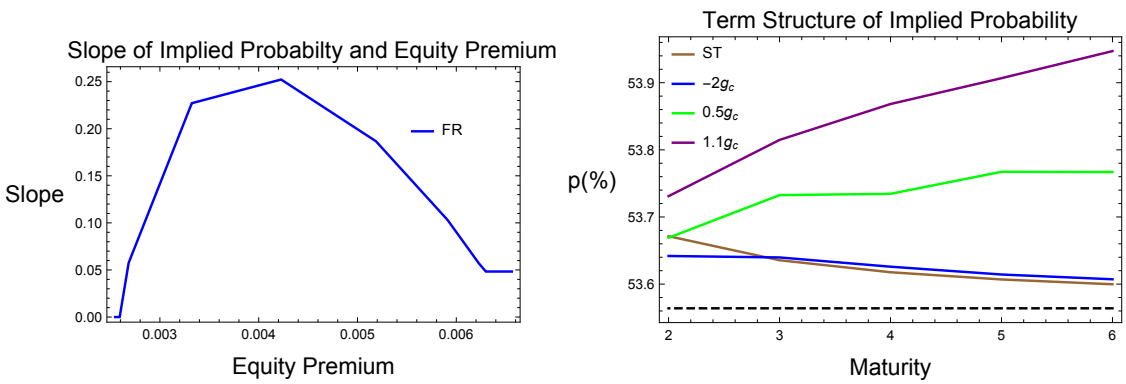
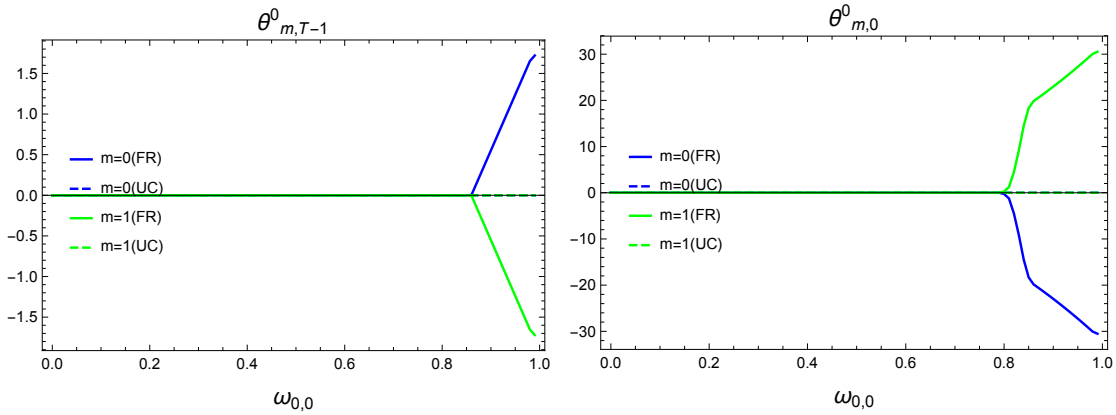
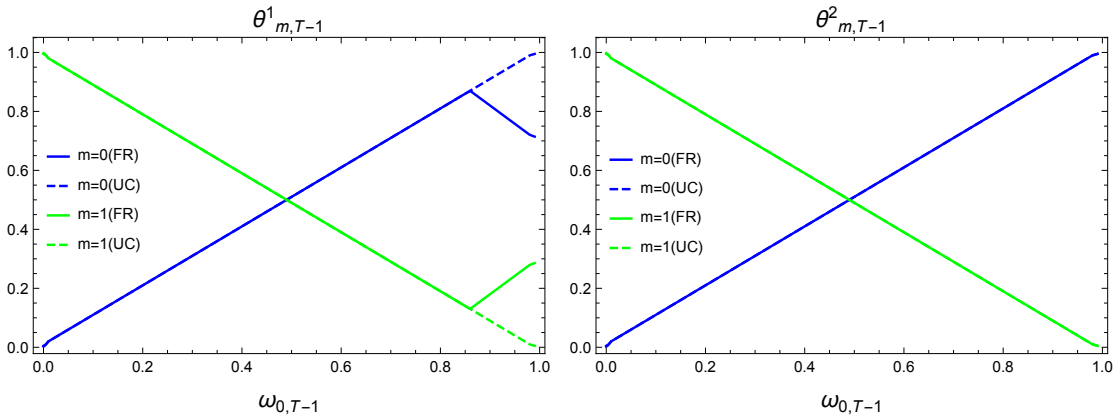


Figure 9: The two institutions' asset allocations for the funding-ratio model with two risky assets, at time $T-1$, and time 0. Solid (dashed) lines correspond to the funding-ratio (unconstrained) model.

(a) Allocations to the riskfree asset at time $T-1$ (left panel), and time 0 (right). Allocations to the riskfree asset are always zero in the unconstrained model, making dashed lines overlap each other.



(b) Allocations to the lower-risk (left panel) and the higher-risk (right panel) assets at time $T-1$. Allocations to the higher-risk asset do not deviate from their unconstrained level, making solid lines overlap dashed lines.



(c) Allocations to the lower-risk (left panel) and higher-risk (right panel) assets at time 0. Allocations to the higher-risk asset do not deviate from their unconstrained level, making solid lines overlap dashed lines.

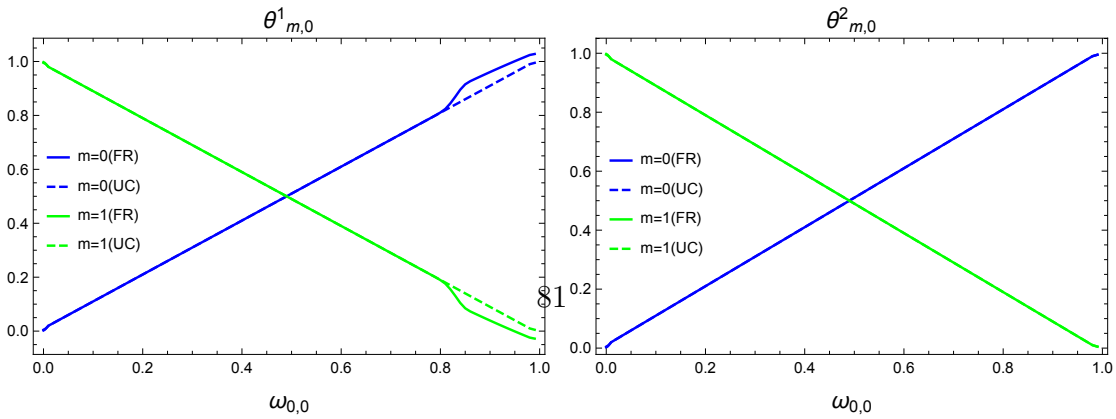
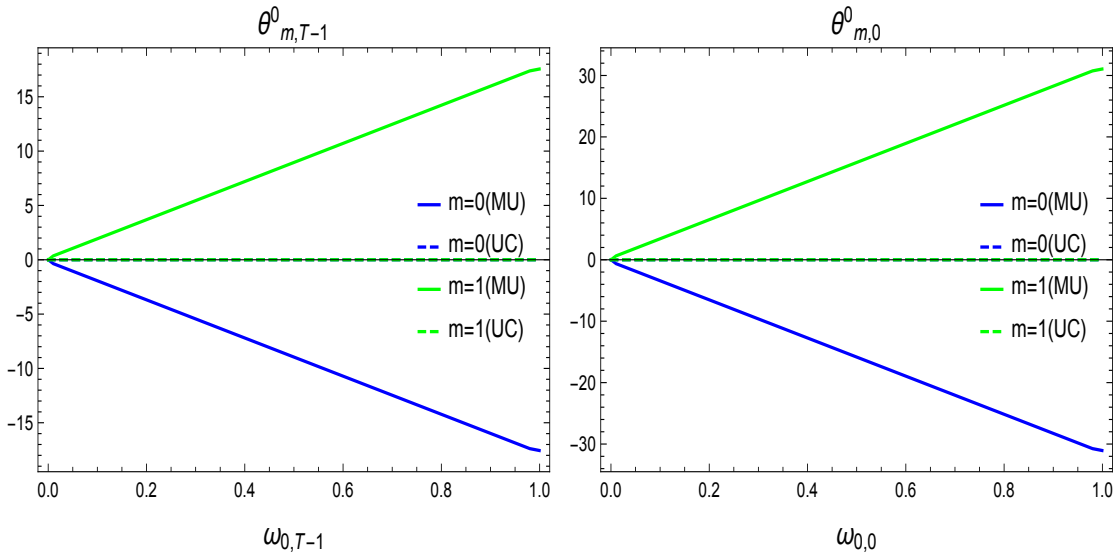


Figure 10: The two institutions' asset allocations for the surplus-payout utility model with two risky assets. Solid (dashed) lines correspond to the surplus-payout utility (unconstrained) model.

(a) Bond allocations at time $T-1$ (left panel) and time 0 (right panel) for the surplus-payout utility model. Allocations to the riskfree asset are always zero in the unconstrained model, making dashed lines overlap each other.



(b) Allocations to the lower-risk risky asset (left panel) and the higher-risk risky asset (right panel) at time $T-1$ for surplus-payout utility model. Allocations to the risky asset do not deviate from their unconstrained level, making solid lines overlap dashed lines.

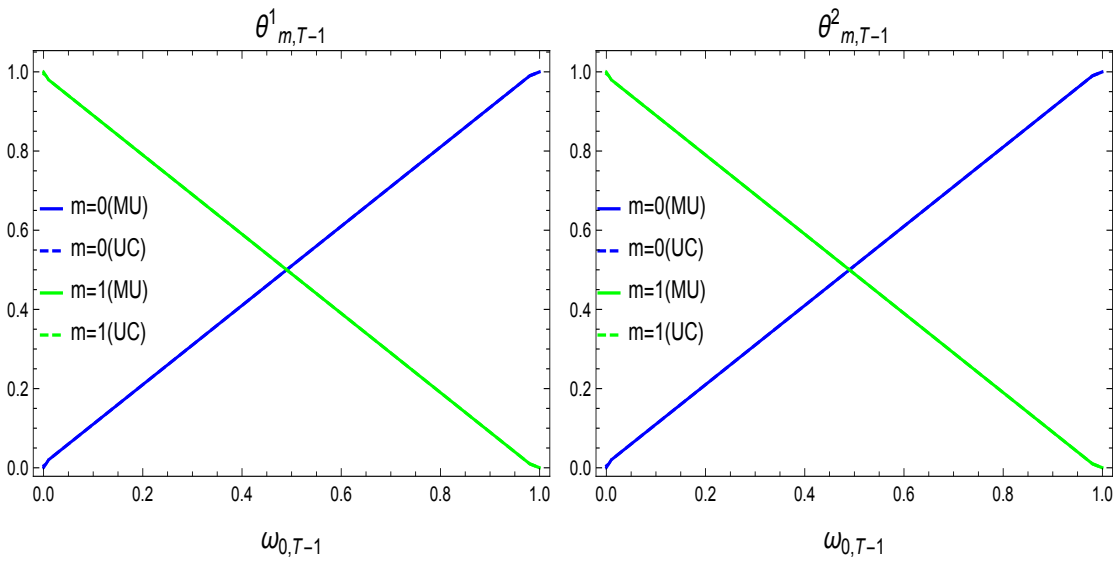


Figure 11: Difference between the price/dividend ratio of the two risky assets at time 0, for the funding-ratio model.

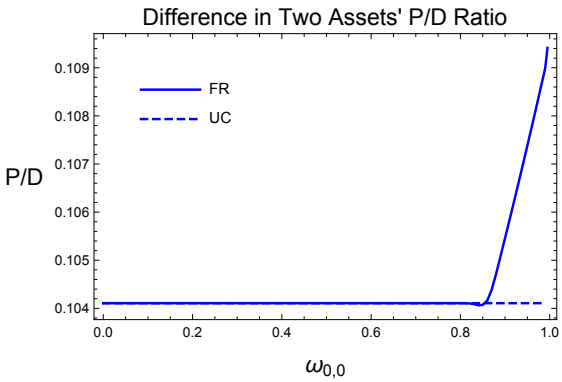
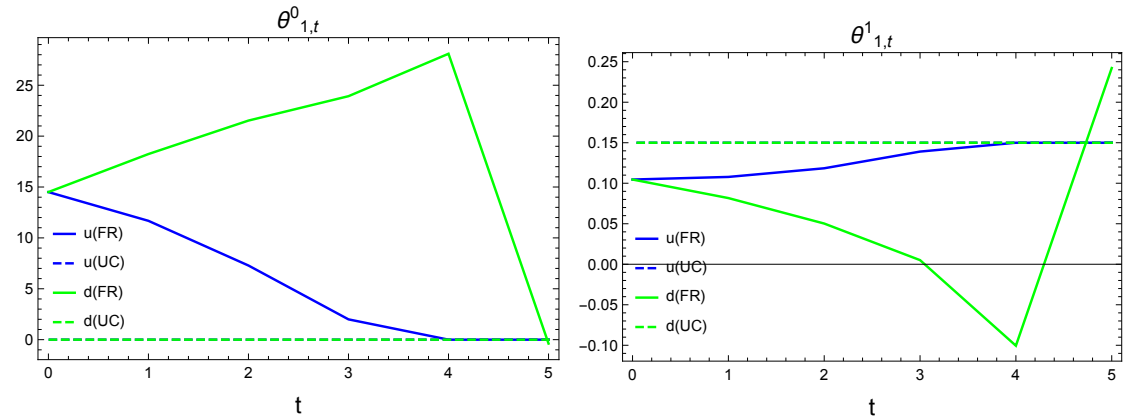
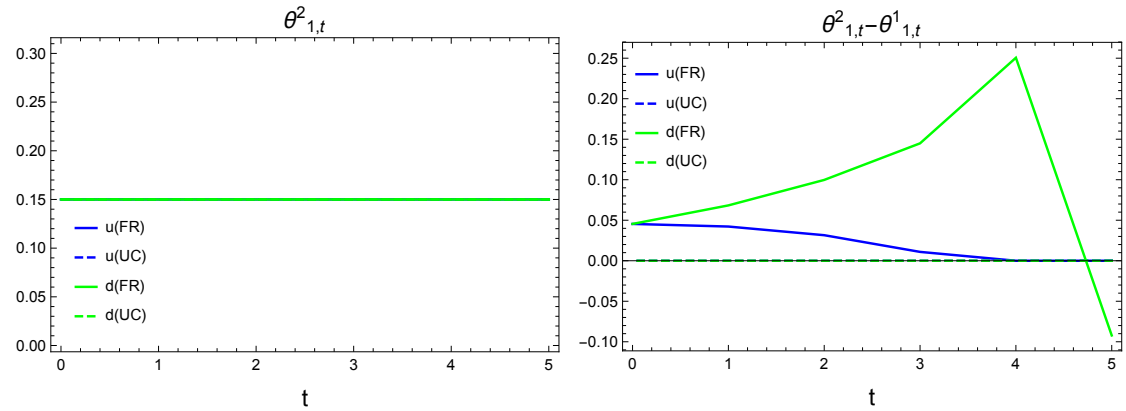


Figure 12: Time evolution plots for the funding-ratio model with two risky assets. Solid (dashed) lines correspond to constrained (unconstrained) model. All curves are drawn for an initial endowment share of 0.85 for the unconstrained institution. For the sake of simplicity, we denote expected excess returns simply as $R^{ex}_{n,t}$, and drop the expectation operator.

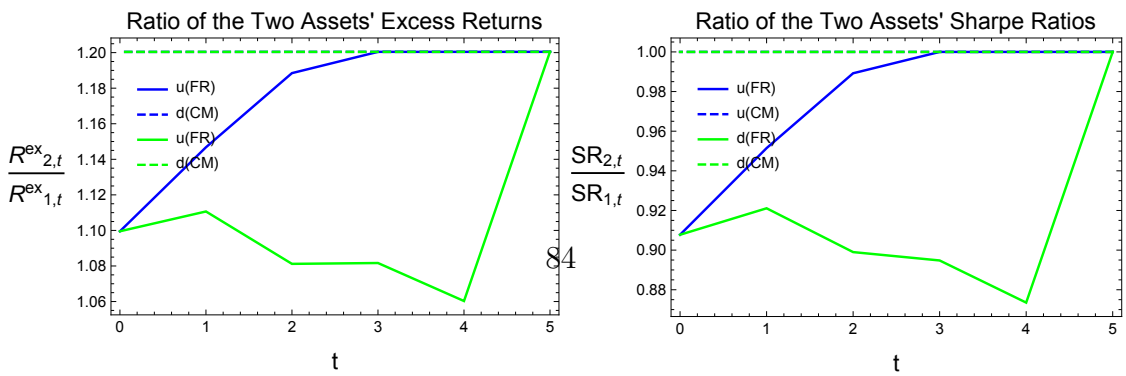
(a) Time evolution of the constrained institution's allocations to the riskfree (left panel) and the lower-risk risky asset (right panel).



(b) Time evolution of the constrained institution's allocations to the higher-risk risky asset (left panel) and the difference between allocations to higher- and lower-risk risky assets (right panel). Allocations to the more-risky asset remain unchanged from the unconstrained case, making all curves overlap each other in the left panel.



(c) Time evolution of the ratio of expected excess returns (left panel) and Sharpe ratios (right panel) of the two risky assets.



References

Basak, S. (1995). A general equilibrium model of portfolio insurance. *8*(4), 309–341.

Basak, S. and A. Shapiro (2001). Value-at-risk-based risk management: optimal policies and asset prices. *Review of Financial studies 14*(2), 371–405.

Bates, D. S. (2000). Post-'87 crash fears in the s&p 500 futures option market. *Journal of Econometrics 94*(1), 181–238.

BIS (2011). Basel iii: A global regulatory framework for more resilient banks and banking systems. *Bank of International Settlement*.

Blome, S., K. Fachiner, D. Franzen, and G. Scheuenstuhl (2007). Pension fund regulation and risk management: Results from an alm optimisation exercise. chapter 4 in: D. lindeman and j. yermo. *Protecting Pensions: Policy Analysis and Examples from OECD Countries*.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics 7*, 265–296.

Brennan, M. J. and E. S. Schwartz (1989). Portfolio insurance and financial market equilibrium. *Journal of Business*, 455–472.

Carney, M. (2016). Uncertainty, the economy and policy. *Speech by Mark Carney, Governor of the Bank of England*.

Chan, Y. L. and L. Kogan (2001). Catching up with the Joneses: Heterogeneous preferences and the dynamics of asset prices. Technical report, National Bureau of Economic Research.

Cuoco, D., H. He, and S. Isaenko (2008). Optimal dynamic trading strategies with risk limits. *Operations Research 56*(2), 358–368.

Cuoco, D. and R. Kaniel (2011, August). Equilibrium prices in the presence of delegated portfolio management. *Journal of Financial Economics 101*(2), 264–296.

Detemple, J. and A. Serrat (2003). Dynamic equilibrium with liquidity constraints. *Review of Financial Studies 16*(2), 597–629.

Dumas, B. and A. Lyasoff (2012). Incomplete-Market Equilibria Solved Recursively on an Event Tree. *The Journal of Finance LXVII*(5), 1897–1941.

Dybvig, P. H. and Z. Qin (2016). How to squander your endowment: Pitfalls and remedies. *AFA 2016 meeting*.

El Karoui, N. and M. Jeanblanc-Picqué (1998). Optimization of consumption with labor income. *Finance and Stochastics 2*(4), 409–440.

Friedman, B. M. (1995). Economic Implications of Changing Share Ownership.

Gompers, P. A. and A. Metrick (2001). Institutional investors and equity prices. *The Quarterly Journal of Economics*.

Gromb, D. and D. Vayanos (2010). Limits of Arbitrage : The State of the Theory.

Grossman, S. J. and Z. Zhou (1996). Equilibrium analysis of portfolio insurance. *The Journal of Finance 51*(4), 1379–1403.

Haldane, A. G. (2014). The age of asset management? *Speech at the London Business School 4*.

Harvey, C. R. (1988). The real term structure and consumption growth. *Journal of Financial Economics 22*(2), 305–333.

He, Z. and A. Krishnamurthy (2013). Intermediary asset pricing. *American Economic Review 103*(2), 732–70.

Kogan, L., S. A. Ross, J. Wang, and M. M. Westerfield (2006). The price impact and survival of irrational traders. *The Journal of Finance 61*(1), 195–229.

Mehra, R. and E. C. Prescott (1985). The equity premium: A puzzle. *Journal of monetary Economics 15*(2), 145–161.

Mitton, T. and K. Vorkink (2007). Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies 20*(4), 1255–1288.

Pan, J. (2002, January). The jump-risk premia implicit in options: evidence from an integrated time-series study. *63*(1), 3–50.

Panageas, S. (2011). Pension Design in the Presence of Systemic Risk. *Working Paper*.

Rime, B. (2001). Capital requirements and bank behaviour: Empirical evidence for switzerland. *Journal of Banking & Finance 25*(4), 789–805.

Yan, S. (2011). Jump risk, stock returns, and slope of implied volatility smile. *Journal of Financial Economics 99*(1), 216–233.