Financial Vulnerability and Monetary Policy*

Tobias Adrian Fernando Duarte

This version: September 2022

Abstract

We present a microfounded New Keynesian model that features financial vulnerability. Financial intermediaries' occasionally binding value at risk constraints give rise to variation in the pricing of risk that generates time varying conditional moments of output. The conditional mean and volatility of the output gap are negatively related: during times of easy financial conditions—when the price of risk is low—growth tends to be high, and risk tends to be low. Monetary policy affects output directly via the intertemporal substitution of savings, and also via the pricing of risk that relates to the tightness of the value at risk constraints. The optimal monetary policy rule always depends on financial vulnerability in addition to the output gap, inflation, and the natural rate. We show that a classic Taylor rule exacerbates deviations of the output gap from its target value of zero relative to an optimal interest rate rule that includes vulnerability. The model provides a microfoundation for optimal monetary policy that takes financial vulnerability into account.

Keywords: optimal monetary policy, macro-finance, financial vulnerability, nonlinear DSGE model, continuous time, closed form
JEL classification: G10, G12, E52

*Adrian is with the International Monetary Fund. Duarte is with Brown University. The authors thank Ricardo Caballero, John Campbell, Richard Crump, Gauti Eggertson, Bill English, Jordi Gali, Marc Giannoni, James Hamilton, Nellie Liang, Alp Simsek, Lars Svensson, Andrea Tambalotti, Ivan Werning, Anna Kovner, Nina Boyarchenko, and seminar and conference participants at the Massachusetts Institute of Technology, the Federal Reserve Bank of New York, International Monetary Fund, National Bureau of Economic Research, the European Central Bank, Yale University, University of Tokyo, Kyoto University, CEMLA, and the Bank of Japan for helpful comments. The views expressed in this paper are those of the authors and do not necessarily represent those of the International Monetary Fund, the Federal Reserve Bank of New York, or the Federal Reserve System.

1 Introduction

Macro-financial linkages feature prominently in monetary policy. The ease of financial conditions—empirical metrics for the pricing of risk—are commonly referred to in monetary policy statements. Financial vulnerability assessments, such as metrics of risk taking in the financial sector, are considered by monetary policy committees. And recent policy frameworks of major central banks reference financial stability, though only few central banks have an explicit financial stability mandate for monetary policy.

In fact, prudential policy (both macro and micro) is usually seen as the primary tool to ensure financial stability, while monetary policy is focused on macroeconomic fluctuations of inflation and real activity. However, prudential policy is unlikely to entirely eliminate the amplification of fundamental shocks due to macro-financial linkages. Even when risks of financial crises are small, macro-financial linkages can have first order impacts on macroeconomic outcomes, particularly on the downside. For example, the deleveraging of financial intermediaries can magnify falling asset prices and amplify adverse macroeconomic shocks. In addition, macroprudential tools may be imperfect, potentially making their use less attractive. To date, the literature has not converged on a macro-financial modeling framework that captures inflation, output, and financial variables in a tractable and parsimonious yet empirically realistic way that is fruitful for the analysis of monetary policy.

In this paper, we explicitly model the link of financial vulnerability to downside risks of GDP, and ask to what extent optimal monetary policy should take such downside risks into account. We define financial vulnerability as the GDP-at-Risk due to macro-financial linkages. GDP-at-Risk is a welfare oriented summary metric that captures macro-financial amplification mechanisms. We present a parsimonious macroeconomic framework for incorporating financial vulnerability in monetary policy, and derive the optimal monetary policy that we show depends no only on inflation and output (gaps), but also on GDP-at-Risk.

Our starting point is the standard New Keynesian (NK) model of Woodford (2003) and Galí (2008). Households have risk averse utility over a consumption basket of differentiated goods, and supply labor to firms that produce the differentiated goods. Good producing firms have labor as the only input for production, are monopolistically competitive in the differentiated goods market, and are subject to nominal price rigidities a la Calvo. They issue equity shares that pay their profits as dividends.

The main point of departure from the standard NK model is the existence of banks that are subject to a value-at-risk (VaR) constraint. The VaR constraint can be interpreted as the optimal contract for a risk-shifting agency problem, as in Adrian & Shin (2013a), or as capturing other real-world elements such as risk-management constraints and financial regulation, including stress-tests. The role of banks is to channel the savings of households into the purchase of equity shares of good producers, which we assume households cannot directly hold (for example, because they lack the necessary information, expertise, technology, or relationships to do so). Despite the restriction on their holdings of good producers' equity shares, households are allowed to trade a full set of state-contingent Arrow-Debreu securities with banks. Therefore, households face complete markets and can trade derivative securities that replicate the payoff of the good producers' stock that they are not allowed to hold directly. However, because of the banks' VaR constraint, the price of such derivative security, and in general of all financial assets, will turn out to be distorted from the point of view of the household, providing scope for welfare-improving policy actions.

Using techniques from continuous-time financial economics, we solve for the full stochastic non-linear equilibrium in closed form, and characterize it through the conditional means and conditional volatilities of endogenous variables as a function of state variables. To provide a more transparent explanation of how the model works, and to make it easier build upon and compare our model to existing NK models, we also derive a simpler four-equation reduced form approximation to our model.

Compared to the standard three-equation NK model, our model has two new features. The first new feature is an equation that describes the evolution of a new state variable, GDP-at-Risk (the "fourth equation"). GDP-at-Risk is defined as the 5^{th} conditional quantile of the one-year-ahead GDP distribution. It summarizes the tail risks to aggregate economic activity that arise from the amplification of financial shocks created by the VaR constraint of banks. GDP-at-Risk is thus the magnitude of downside risk to GDP due to macro-financial linkages.

The second new feature is that the IS curve (derived from the Euler equation) has an endogenously time-varying risk premium. The widely used first-order approximation in discrete time NK models has a risk premium that is identically zero by construction. Preserving time variation in the risk premium using the standard methodology would require a third-order approximation with all of its associated challenges. Instead, we use a novel approximation technique that consists in linearizing *first and second moments* of the fully nonlinear continuous-time stochastic processes, which preserves the time variation in the risk premium while remaining just as tractable (because stochastic processes are linear).

The two new elements we introduce – the GDP-at-Risk state variable and the time-varying risk premium – interact to create inefficient fluctuations in the real economy. A binding VaR constraint amplifies the responsiveness of banks' optimal portfolio to economic and financial fluctuations. When the risk premium increases, banks invest more in risky assets not only because the risky assets are themselves more attractive – an effect present even in the absence of a VaR constraint – but also because their risk-taking capacity has increased as the VaR constraint is relaxed. In the short-term, realized returns on financial assets are high, volatility is low, and employment and output increase. However, the banks' additional investment in risky assets induced by the relaxation of the VaR constraint is financed by debt. The resulting increase in banks' leverage is thus also magnified by the VaR constraint. With higher financial vulnerability due to higher leverage, future gains and losses are also magnified. Households hold financial assets and trade with banks in complete markets, so they inherit the increased amplification of higher leverage through wealth effects and the pricing of risk. This increased amplification ultimately translates into larger expected fluctuations in consumption and output in the medium run. The higher volatility of output expected in the future implies larger downside risks as captured by GDP-at-Risk. When GDP-at-Risk increases, so does the risk premium, because in equilibrium investors require additional compensation for taking the additional tail risk. In turn, the higher risk premium leads to a higher GDP-at-Risk, and so on. The same logic applies to negative shocks, which result in faster and deeper deleveraging, and larger declines in consumption and output, than without the VaR constraint. Thus, the dynamics of the model generate the volatility paradox, a term coined by Brunnermeier & Sannikov (2014a) to describe the idea that that times of easy financial conditions associated with low downside risks in the short term tend to be followed by the buildup of tail risks in the medium term. This intertemporal tradeoff is one of the key new considerations in the setting of optimal policy offered by our model.

Our modeling approach is motivated by the empirical evidence that financial conditions forecast tail risks of macroeconomic outcomes. Estrella & Hardouvelis (1991) and Estrella & Mishkin (1998) show that the term spread, an indicator of the pricing of interest rate risk, forecasts recessions. Gilchrist & Zakrajšek (2012), López-Salido et al. (2016), and Krishnamurthy & Muir (2016) find that credit spreads forecast downside risks to GDP growth. More generally, Adrian et al. (2019) document for the United States that financial conditions are strong forecasters of downside risks to GDP growth, while other non-financial variables such as lagged GDP growth and inflation, are not. More precisely, they find that the upper quantiles of the one- and two-year-ahead GDP growth distribution are roughly constant, while the lower conditional quantiles – such as GDP-at-Risk – respond strongly to financial conditions. They explain this empirical regularity by showing that there is a negative conditional (on financial conditions) correlation between the mean and volatility of the GDP distribution rather than, say, time variation in higher moments. Deteriorating financial conditions give rise to a decline in the conditional mean and an increase in the conditional volatility of GDP, and vice-versa for improving financial conditions. The low-mean high-volatility states thus produce a high GDP-at-Risk. In addition, the conditional negative correlation between the mean and the volatility of output imply that the unconditional distribution of GDP is highly skewed to the left as a function of financial conditions.[1] Adrian et al. (2018) show that these empirical patterns also hold for a panel of emerging and developing countries. Using granular instrumental variables, Adrian et al. (2022) provide evidence that the relationship from loose financial conditions to future downside risks is causal. In addition, they study the term structure of GaR, that is, how financial conditions predict the tails of the GDP growth distribution at different horizons. They find that when initial financial conditions are loose, downside risks are lower in the near term but increase in later quarters – another manifestation of the volatility paradox.

In this paper, we verify that the same patterns that Adrian et al. (2019) found for GDP growth and Adrian et al. (2022) found for the term structure of GaR are also present in the output gap, which is more welfare-relevant in NK models. We also show the new result that inflation, the other important determinant of welfare in NK models, does *not* follow the same patterns. Indeed, financial conditions forecast neither the mean nor the volatility of inflation, providing important restrictions for models that attempt to replicate these empirical patterns.

Correspondingly, we show that our NK model with banks and financial vulnerabilities generates these same patterns for the output gap and inflation that we document empirically.

The tractability of the solution together with the empirically validated connection between financial conditions and the tails of the output gap distribution allow us to study the "risk return tradeoff of monetary policy". The central bank is assumed to minimize a standard loss function with squared deviations of the output gap and inflation from their target of zero as arguments. We highlight that this inflation and output gap "dual mandate" loss function does not have financial stability as part of its mandate and still represents a second-order approximation to the true welfare loss function in our NK model with banks. We solve for the optimal policy with commitment in closed form using dynamic programming, as our new approximation technique preserves the tractability of the linear-quadratic framework while retaining the time variation dynamics of tail risks. The risk-return tradeoff for monetary policy is that changes in interest rates move the conditional mean and conditional volatility of the output gap in opposite directions and cannot both be zero simultaneously. Higher interest rates increase expected consumption growth through the standard intertemporal substitution

[1] When macro-financial linkages are generated by financial vulnerabilities, they can generate downside risks to growth and thus pose threats to financial stability. Bernanke & Gertler (1989) and Bernanke & Blinder (1992) argue for the credit channel of monetary policy where financial frictions shift credit demand or supply curves, thus generating an amplification mechanism that is transmitted via changes in the pricing of risk. Changes in the pricing of risk are generally caused by deeper frictions linked to leverage in the financial and non-financial sectors and the degree of maturity transformation (see Brunnermeier et al. (2013) or Adrian et al. (2015) for surveys on the role of vulnerabilities for macro-financial linkages). Financial vulnerabilities can cause changes in downside risks to GDP, and the pricing of risk tends to reflect financial vulnerabilities. Recent literature on monetary policy has started to explore the role of financial conditions and vulnerabilities in monetary policy settings (see Adrian & Shin (2010), Borio & Zhu (2012)). Curdia & Woodford (2010) and Gambacorta & Signoretti (2014) consider Taylor rules that are augmented with financial variables and study their quantitative relevance.

channel of the Euler equation. In addition, and unlike the standard NK model, higher interest rates reduce the conditional volatility of consumption. The reason is that higher interest rates make borrowing more expensive, inducing a reduction in leverage for banks. Lower leverage leads to lower VaR for banks, smaller future amplification in the financial sector, a lower price of risk, and lower financial volatility. And since banks and households trade in complete markets, the lower financial volatility translates into lower volatility of the consumption the household, and hence the real economy.

Our optimal monetary policy rule can be cast in the language of a flexible inflation targeting framework, such as the one in Svensson (1999), Rudebusch & Svensson (1999), Svensson (2002), and Giannoni & Woodford (2012). Relative to the standard New Keynesian model, there are two important differences. First, vulnerability enters the optimal rule. Most standard inflation forecasting frameworks advocate conditioning monetary policy on financial vulnerability to the extent that it helps forecast inflation or the output gap. In our framework, financial vulnerability does forecast inflation and the output gap. However, optimal monetary policy depends on financial vulnerability not only because its forecasting ability, but also because financial vulnerability endogenously affects the future volatilities of inflation and the output gap. Optimal monetary policy therefore depends on financial vulnerability even keeping expectations of inflation and the output gap fixed. Second, the optimal response of interest rates to changes in inflation and the output gap depend on the bank's structural parameters since they take into account the endogenous effect that changes in inflation and the output gap have on financial vulnerability. Optimal monetary policy can also be expressed as an augmented Taylor rule. The nominal interest rate not only depends on inflation and output, but also on financial vulnerability. The optimal coefficients on output and inflation are different from the standard ones because they take the parameters that govern vulnerability into account.

Our model's optimal policy captures the intuition that in recent years monetary policy has explicitly taken into account and influenced financial conditions (see Dudley (2015, 2017); Yellen (2016)). A deterioration of financial conditions corresponds to an increase in tail risk, as conditional GDP volatility rises, while the conditional growth forecast deteriorates. As a result of such an increase in financial vulnerability, i.e. an increase in the downside risk to GDP growth, monetary policy is relatively easier than under the classic Taylor rule. This results in a concurrent lowering of vulnerability, and hence in less severe left skewness of GDP.

While we do not explicitly consider macroprudential policy, we show in reduced form that the stance of prudential policy can change the tradeoffs faced by monetary policy. When prudential policy is appropriately designed, vulnerabilities are mitigated, improving the tradeoffs for the monetary policy authority. Perfect macroprudential policy would eliminate the need for monetary policy to condition on vulnerability.

2 Financial Vulnerability

2.1 Data

We use data at the quarterly frequency for the period 1971:Q1 to 2022:Q2. Our measure of financial conditions x_t is the National Financial Conditions Index (NFCI) of the Federal Reserve Bank of Chicago. The NFCI aggregates 105 financial market, money market, credit supply, and shadow bank indicators to compute a single index using the filtering methodology of Stock & Watson (1998). It is normalized to have mean zero and standard deviation one, and higher values representing tighter financial conditions. We construct the output gap y_t as the log-difference between real GDP from the U.S. Bureau of Economic Analysis (BEA), and the estimate of real potential output from the Congressional Budget Office. Inflation π_t is year-over-year core PCE inflation from the BEA. Cumulative output gap and inflation over h quarters is

$$y_{t+h}^{(h)} = \sum_{j=1}^{h} \Delta y_{t+j},$$

$$\pi_{t+h}^{(h)} = \sum_{j=1}^{h} \Delta \pi_{t+j}.$$

2.2 Financial Conditions

Adrian et al. (2019) and Adrian et al. (2022) document that the time variation in the conditional distribution of GDP growth associated with financial conditions is driven by time variation in the first and second moments of the conditional distribution rather than changes in higher moments or in the overall shape of the distribution. In this paper, we use a conditionally heteroskedastic linear model to estimate the conditional first and second moments of the cumulative output gap between quarters t and $t+h$ as a function of time $t-1$ output gap, inflation and financial conditions

$$y_{t+h}^{(h)} = \gamma_0^y + \gamma_1^y y_{t-1} + \gamma_2^y \pi_{t-1} + \gamma_3^y x_{t-1} + \sigma_{t+h}^y \varepsilon_{t+h}^y, \tag{1}$$

$$\ln\left(\sigma_{t+h}^y\right) = \delta_0^y + \delta_1^y y_{t-1} + \delta_2^y \pi_{t-1} + \delta_3^y x_{t-1}, \tag{2}$$

and similarly for cumulative inflation

$$\pi_{t+h}^{(h)} = \gamma_0^\pi + \gamma_1^\pi y_{t-1} + \gamma_2^\pi \pi_{t-1} + \gamma_3^\pi x_{t-1} + \sigma_{t+h}^\pi \varepsilon_{t+h}^\pi, \tag{3}$$

$$\ln\left(\sigma_{t+h}^\pi\right) = \delta_0^\pi + \delta_1^\pi y_{t-1} + \delta_2^\pi \pi_{t-1} + \delta_3^\pi x_{t-1}. \tag{4}$$

In equations (1)-(4), ε_t^y and ε_t^π are *iid* standard normal random variables. We estimate the coefficients γ and δ via maximum likelihood and report them in Table 1 for the case $h = 1$. Columns 1, 2, 3 and 4 give the coefficients for equations (1), (2), (3) and (4), respectively. The numbers in brackets are t-statistics computed with White (1980) standard errors.

Table 1: **Conditional Mean and Volatility Estimates**

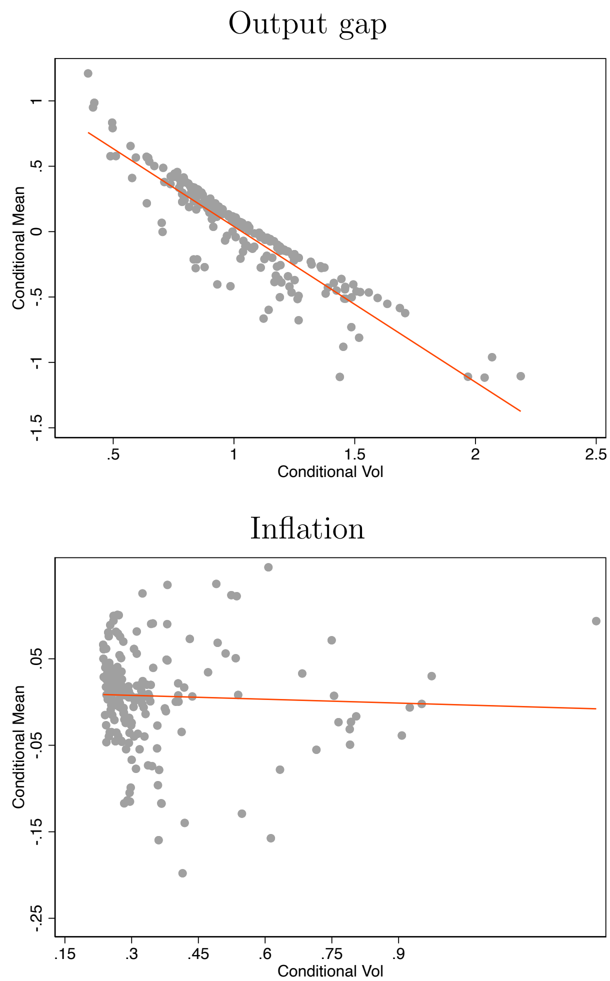
	Output gap t to $t+h$		Inflation t to $t+h$	
	Mean eq.	Vol eq.	Mean eq.	Vol eq.
	(1)	(2)	(3)	(4)
Financial Conditions	-0.339***	0.362**	0.0504	0.542***
	[-3.747]	[2.264]	[0.997]	[3.013]
Output gap	-0.137***	0.220**	0.0144	-0.0443
	[-4.433]	[2.025]	[1.376]	[-0.392]
Inflation	0.0839**	-0.179	-0.0278*	0.0816
	[2.088]	[-1.385]	[-1.747]	[0.819]
Constant	-0.491***	0.921	0.120**	-2.613***
	[-2.702]	[1.153]	[2.339]	[-7.020]
Observations	204	204	204	204

*** p<0.01, ** p<0.05, * p<0.1

The first row of the first two columns of Table 1 show that financial conditions predict both the first and second moments of the output gap (columns (1) and (2)) with coefficients that are statistically significant and of opposite signs. Tighter financial conditions (a higher NFCI) are associated with a lower conditional mean and a higher conditional volatility of the distribution of the future output gap. The conditional mean of inflation, on the other hand, is not significantly affected by financial conditions (row 1, column 3), while the volatility of inflation is (row 1, column 4).

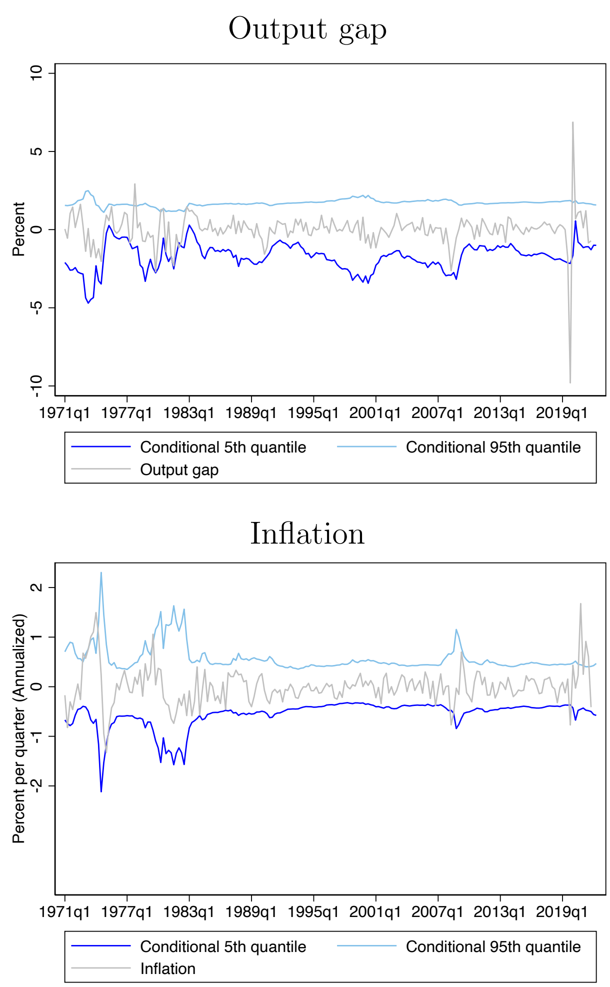
The scatterplot in Figure 1 shows the relation between the conditional mean and conditional volatility of the output gap in Panel (a) and of inflation in Panel (b). Each point in the figure corresponds to a quarter in our sample, with the coordinates of the point given by the estimated value of the conditional mean and conditional volatility in that quarter. For the output gap in Panel (a), mean and volatility are strongly negatively correlated, while for inflation in Panel (b) there is little relation between the two.

Figure 1: **Estimated Conditional Mean and Conditional Volatility of One Quarter Ahead Output Gap and PCE Inflation**. The figure reports estimates from equations (1), (2), (3), and (4). Panel (a) plots the output gap mean against the output gap volatility, panel (b) plots inflation mean against inflation volatility.



The difference in how financial conditions jointly forecast the conditional mean and volatility of the output gap and inflation imply that the tails of their respective conditional distributions behave very differently. In Panel (a) of Figure 2, we plot the time series of the realized values of y_{t+1} (in grey), together with the 5th and 95th quantiles conditional on financial conditions. The 95^{th} quantile (light blue line) is quite stable, almost constant as a function of financial conditions. In contrast, the 5^{th} quantile (blue line) moves significantly over time and responds strongly to financial conditions. We can explain this behavior using the intuition from Figure (1). When financial conditions are tight, the distribution of the output gap shifts down (has lower mean) and, at the same time, widens (has higher standard deviation). Both of these changes increase the probability of lower output gap outcomes (while they offset each other for higher values). As a result, the lower tail of the distribution increases substantially. If financial conditions continue to tighten, the low-mean high-volatility distribution can create higher probabilities of ever-decreasing values. When financial conditions are loose, the conditional distribution has high mean and low volatility. The low volatility during these high-mean episodes explains why the higher quantiles are roughly constant. As the mean increases, the volatility approaches zero, effectively putting a "ceiling" on the distribution. Panel (b) shows the analogous plot for inflation. The 5th and 95th quantiles show comparable degrees of variability and are almost symmetric. Despite the well-documented presence of significant heteroskedasticity in inflation, the distribution of inflation *conditional on financial conditions* appears almost homoskedastic. When financial conditions tighten, inflation is more volatile, but its mean remains stable, causing a similar change in upper and lower quantiles of the distribution. One implication of Figure 2 is that the downside risks to the real economy connected to financial conditions are more closely associated to the output gap than to inflation, an important empirical pattern that can be used to distinguish between different models of financial amplification.

Figure 2: **Estimated Conditional Distribution of One Quarter Ahead output Gap and PCE Inflation**. The figure reports estimates from equations (1), (2), (3), and (4). Panel (a) shows the actual output gap, the conditional mean of output gap, and the 5th and 95th quantile. Panel (b) shows the actual PCE inflation, the conditional mean of inflation, and the 5th and 95th quantile.



In Figure (3), we show results when we estimate the coefficients in equations (1)-(4) for different values of h. Panel (a) shows the coefficient on financial conditions in the mean equation for the output gap, γ_3^y from equation (1), as a function of h. Panel (b) shows the coefficient on financial conditions in the volatility equation for the output gap, δ_3^y from equation (2), as a function of h. For $h = 0$, we set the values of γ_3^y and δ_3^y to zero. For $h = 1$, the values in the figure are the same as the ones reported in columns (1) and (2) of Table 1.

As h increases, the signs of both coefficients flip, resulting in the volatility paradox alluded to in the Introduction. Looser initial financial conditions are associated with a higher mean and a lower volatility for the cumulative output gap distribution in the short term, for horizons between one to two years[2]. However, for longer horizons of around 5 years, the initial looser financial conditions are associated with a *lower* mean and a *higher* volatility.

Panels (c) and (d) of Figure (3) show that there is no volatility paradox in inflation. The coefficient on financial conditions in both the mean and volatility equations for inflation are either positive or statistically insignificant throughout at all horizons.

[2]Because we are considering the *cumulative* gap (rather than the gap at a single point in time), a higher mean at an horizon of, say, $h = 8$, means that the output gap is expected to be higher than today two years from now, which implies an expected increase in the output gap between t and $t + 8$ (rather than just a higher output gap for the single future period $t + 8$)

Figure 3: **Estimated Coefficients on Financial Conditions.** The figure reports estimates for $\gamma_3^y, \delta_3^y, \gamma_3^\pi, \delta_3^\pi$ from equations (1), (2), (3), and (4) as a function of the prediction horizon h.

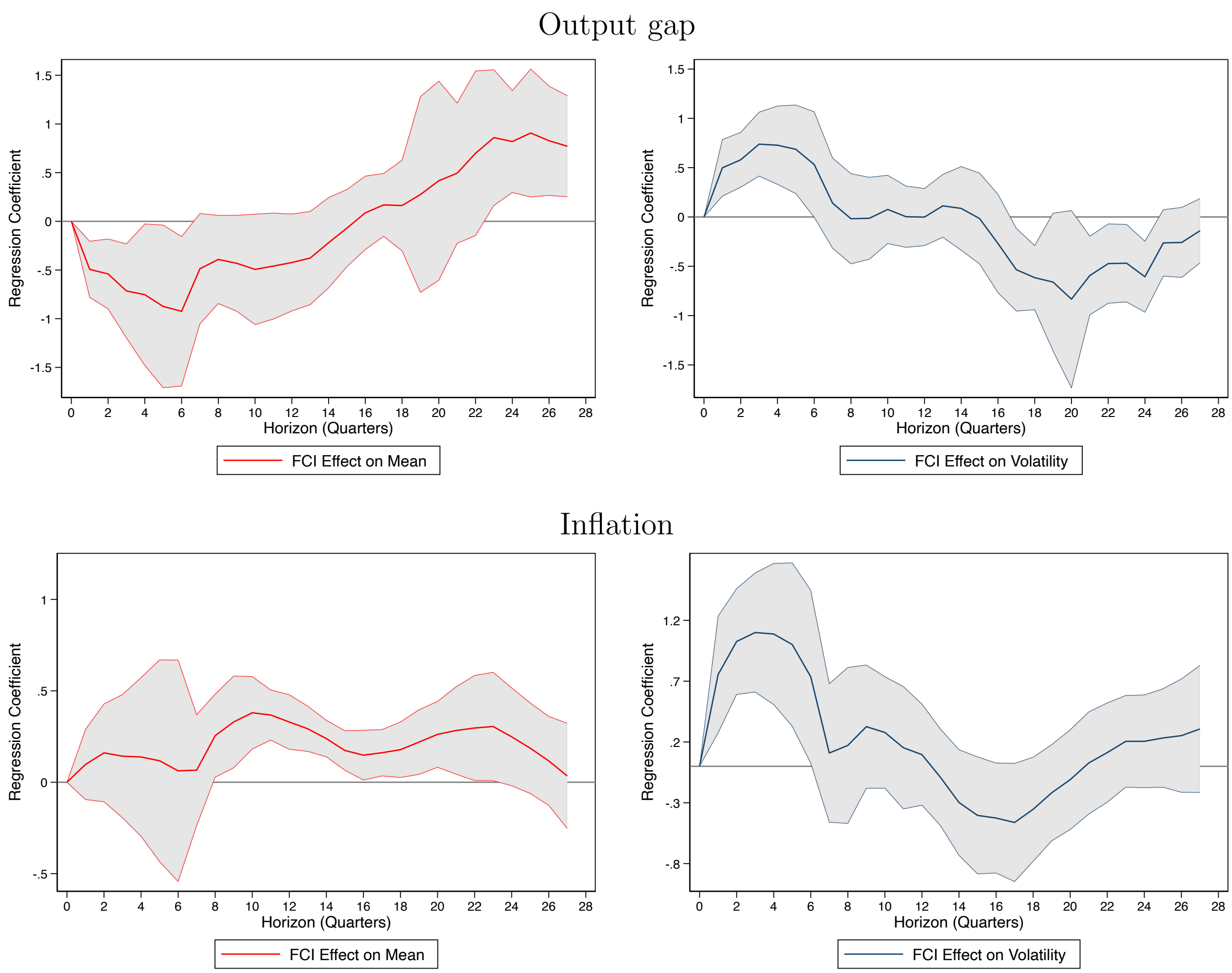
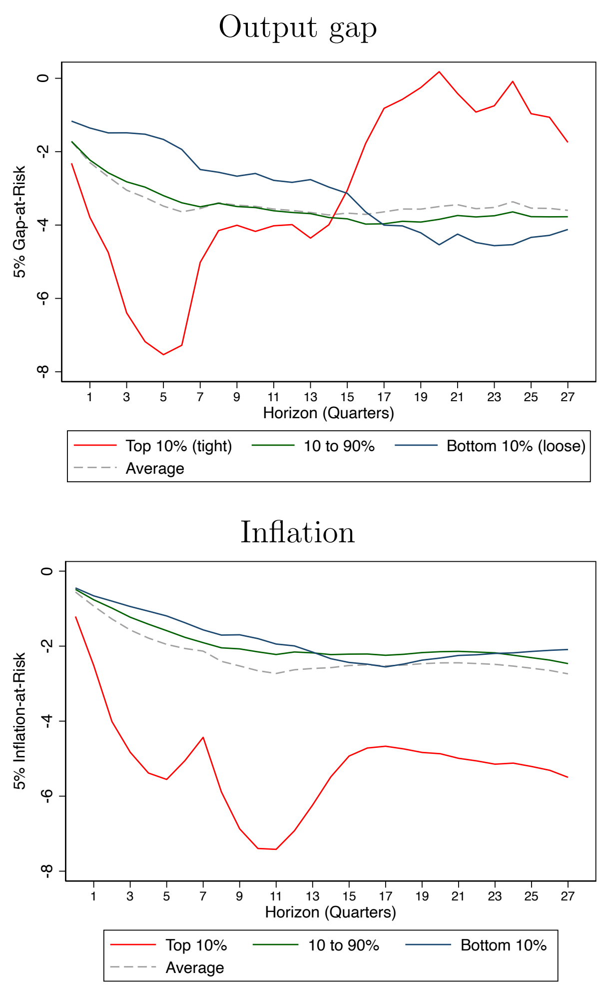


Figure (4) offers another way to understand the volatility paradox. Panel (a) plots the term structure of the 5^{th} conditional quantile of the cumulative output gap distribution, or "GDP-at-Risk", for different initial financial conditions. To construct Figure (4), we first compute "GDP-at-Risk" for every period and for all values of h that we consider (using equations (1)-(4) and the corresponding estimated coefficients). We then divide all time periods into three groups according to whether the NFCI in that quarter was below the 10th quantile, between the 10th and 90th quantile, or above the 90th quantile, of the unconditional NFCI distribution. Last, we average the values of "GDP-at-Risk" within each of the three groups and plot them as a function of h. When financial conditions are initially tight (below the 10th quantile of their distribution), the red line shows that tail risks to the output gap increase ("GDP-at-Risk" goes down) at short horizons but decrease at long horizons. When financial conditions are above their 10th quantile, the green and blue lines show that "GDP-at-Risk" steadily declines with h. The volatility paradox occurs when the GDP-at-Risk term structure lines cross each other.

For example, the red line crosses the blue line at around the 4-year horizon. This means that when financial conditions are initially loose, downside risks to the output gap are smaller in the short run than when financial conditions are initially tight (the blue line is above the red line before $h = 16$) but higher in the medium run (the blue line is below the red line after $h = 16$).

Figure 4: **Term structure of GDP and Inflation-at-Risk.** The figure shows the 5th conditional quantile of the cumulative output gap distribution in Panel (a) and cumulative inflation in Panel (b), as a function of the prediction horizon h and the initial level of financial conditions. The red lines correspond to tight financial conditions with the NFCI below its 10th unconditional quantile; the green lines correspond to intermediate financial conditions with the NFCI between the 10th and 90th quantile of its unconditional distribution; and the blue line corresponds to loose financial conditions with the NFCI above the 90th quantile of its unconditional distribution. The volatility paradox is reflected in the crossing of the term structure lines for GDP-at-Risk, and is absent for inflation.



Taken together, these results suggest that the lower quantiles of the output gap distributions are related to the link between first and second moments of the GDP gap distribution, conditional on financial conditions. Next, we develop a model that can reproduce the empirical regularities in this section by adding a banking sector with a value-at-risk constraint to a baseline New Keynesian model.

3 Environment

Time is continuous with $t \in [0, \infty)$. There is a continuum of intermediate goods indexed by $i \in [0, 1]$ and a single final good. Output $Y_t(i)$ of each intermediate good of type i can be produced using labor $N_t(i)$ through the constant returns to scale technology

$$Y_t(i) = AN_t(i), \tag{5}$$

where $A > 0$ is the constant aggregate level of TFP. The final good Y_t is produced with the technology

$$Y_t = \left(\int_0^1 Y_t(i)^{1-\frac{1}{\varepsilon}} \, di \right)^{\frac{\varepsilon}{\varepsilon-1}}. \tag{6}$$

where $\varepsilon > 1$ is the constant elasticity of substitution across different goods.

There is a continuum of mass one, identical, atomistic, and infinitely lived households who rank consumption streams C_t and labor streams N_t according to

$$\mathbb{E}_0 \int_0^\infty e^{-\beta t} \left(\frac{C_t^{1-\gamma}}{1-\gamma} - \frac{N_t^{1+\xi}}{1+\xi} \right) dt, \tag{7}$$

where $\beta > 0$ is a time-preference parameter, $\gamma > 0$ is the coefficient of relative risk aversion, and $\xi > 0$ is the inverse of the Frisch elasticity of labor supply. The variable C_t represents a consumption index given by

$$C_t \equiv \left(\int_0^1 C_t(i)^{1-\frac{1}{\varepsilon}} \, di \right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{8}$$

where $C_t(i)$ is the quantity of differentiated good i consumed by the household. The variable N_t is the labor supplied by the household to all firms and given by

$$N_t \equiv \int_0^1 N_t(i) \, di, \tag{9}$$

where $N_t(i)$ is the amount of labor supplied to firm i at time t.

The resource constraint of the economy is

$$Y_t(i) = C_t(i), \tag{10}$$

which implies $Y_t = C_t$.

4 The Efficient Allocation (First Best)

The efficient allocation (first best) is obtained by choosing paths of $C_t\,(i)$, $N_t\,(i)$ and $Y_t(i)$ to maximize the utility of the representative household in equation (7), subject to the structure of the economy's physical environment described by equations (5), (8), (9), and (10). The solution to this problem is

$$C_t^{FB}\,(i) = C_t^{FB}\,(j) = C_t^{FB} \tag{11}$$

$$Y_t^{FB}\,(i) = Y_t^{FB}\,(j) = Y_t^{FB} \tag{12}$$

$$N_t^{FB}\,(i) = N_t^{FB}\,(j) = N_t^{FB} \tag{13}$$

for all i and j, with aggregate consumption, output, and employment given by

$$C_t^{FB} = Y_t^{FB} = A^{\frac{1-\gamma}{\gamma+\xi}+1} \tag{14}$$

$$N_t^{FB} = A^{\frac{1-\gamma}{\gamma+\xi}} \tag{15}$$

5 The Decentralized Economy

5.1 Overview

There are four types of agents: firms that produce goods, households, banks, and the government.

Good-producing firms are exactly as in the standard New Keynesian (NK) model, except that they are financed by banks instead of households. There are two kinds: intermediate, and final good producers. Final good producers buy intermediate goods from the intermediate good producers and sell the final good to the households. They sell in a perfectly competitive environment and make zero profits, so their ownership and corporate capital structure are irrelevant for equilibrium outcomes. Intermediate good producers hire labor from the household as the sole input of production in a perfectly competitive and frictionless labor market. Intermediate good producers sell intermediate goods to final good producers in a monopolistically competitive market and set good prices á la Calvo. This structure gives rise to a standard New Keynesian Phillips Curve (NKPC). Intermediate good producers are financed by issuing stocks (equity shares) that pay profits as dividends. The assumption of equity-only financing is without loss of generality since the Modigliani-Miller theorem holds for these firms.

Households are as in the standard NK model with two differences. First, they cannot hold the stocks issued by the intermediate good producers. Second, they can trade a full set of sate-contingent Arrow-Debreu (A-D) securities with the banks without any frictions and, as in the standard NK model, also among themselves.

Banks are intermediation firms that channel households' savings into the financing (that is, the purchasing of the stock) of good producers. Banks are the only agents that have the necessary information, expertise, technology, or relationships to do so.[3] Banks are allowed to trade a full set of A-D securities with the households and among themselves without frictions. Banks finance themselves through the trading of the A-D securities, so their financing can take any state-contingent form, and their capital structure is not a priori restricted. For example, banks could choose to finance themselves with any mix of stocks, short- and long-term debt, and hybrid securities like convertible bonds. Adjusting the capital structure can be done frictionlessly. In particular, the issuance and repurchase of stock are costless.

Banks maximize a risk averse objective function over total distributions to shareholders (dividends plus net stock issuance) by choosing the amount of total distributions and a portfolio of investments in the financial assets available to them, subject to a standard budget constraint and a value-at-risk (VaR) constraint on their wealth[4]. The bank's objective function is subject to preference shocks that are driven by a one-dimensional Brownian motion B_t, which is the only shock in the economy.

The ability of banks and households to trade a full set of A-D securities implies that households face complete markets despite the restriction on their holdings of good producers' stocks. Households can trade derivative securities that replicate the payoff of the good producers' stock that they are not allowed to hold directly. However, because of the banks' preference shocks and VaR constraint, the price of such derivative security, and in general of all financial assets, will turn out to be distorted from the point of view of the household.

Government spending is zero. The government provides a proportional subsidy labor income that is financed each period by lump-sum taxes levied on the household. The subsidy is such that it eliminates the distortions arising from the monopolistic market power of intermediate good producers. The central bank sets the short-term (instantaneous) nominal interest rate by paying interest on base money in the cashless limit, as in Woodford (2003).

The social welfare in the decentralized economy can differ from its first-best level due to the presence of three frictions: nominal price rigidities of intermediate good producers, preference shocks to banks and the VaR constraint of banks. The distortions due to the market power

[3]We use the label of "bank" in our model as a short-hand for for any firm that conducts intermediation between households and firms.

Of course, in practice, only a fraction of all firm financing is intermediated. Allowing households to provide a share of financing to firms without any additional frictions is straightforward. However, doing so provides no additional insight in our setup (although it would change the quantitative magnitude of some of our results). See Brunnermeier & Sannikov (2014b) and Adrian & Boyarchenko (2012) for examples of households that finance firms along with specialized intermediaries, but with additional frictions that make households less efficient at the task.

[4]Banks' wealth, net worth and equity capital are all the same.

of monopolistic firms are corrected by the government subsidy, so they are not a source of inefficiency. Because the first-best levels of output, consumption and employment are constant, all fluctuations in these variables are inefficient. Monetary policy cannot induce the first-best. The first-best would require banks to fully insure households against all risk in all states of the world so that their consumption is constant. But there always exist states of the world in which the risk is too high for banks to insure households and satisfy the VaR-constraint at the same time. Monetary policy can shift risk to banks only up to the point allowed by the VaR-constraint. If any one of the three frictions that cause inefficiency were removed, then it would be possible for monetary policy to induce the efficient allocation.

5.2 Financial Assets

As mentioned above, the financial assets in the economy are a complete set of A-D securities, and the stocks of the intermediate good producers. Because the only source of uncertainty in this model is a one-dimensional Brownian motion B_t, there is a single A-D security. In addition, a riskless real bond and any single risky asset (that is, exposed to B_t) span all possible payoffs of the economy, including that of the A-D security. We can therefore assume without loss of generality that banks are financed by issuing risky stocks and riskless short-term debt only; we refer to total distributions as dividends from now on. We also correspondingly assume, again without loss of generality, that households trade the stocks of banks and the riskless bond only. It is then possible to back out the prices, payoffs, and implied portfolio positions in terms of the original A-D security by using no-arbitrage relations and market clearing conditions.

We framed the original discussion of the market structure in terms of "a full set of A-D securities" (rather than the single A-D security that ends up being all that is needed) to emphasize two aspects of our setup. First, our results straightforwardly generalize to a setup with multiple shocks without requiring any additional market-incompleteness friction.[5] The decentralized equilibrium is inefficient despite markets being complete, with one or with many shocks. Second, the capital structure of banks is unrestricted and can be changed costlessly and instantaneously. That banks issue stocks and riskless debt only is a consequence of the single-shock assumption. In a multiple-shock model, the universe of securities available to banks to finance themselves would be larger. In other words, our results do not rely on a specific capital structure for banks.

Without loss of generality, we group all of the stocks of banks into a single banking sector stock, and all intermediate good producer stocks into a single intermediate good producer sector stock. The banking sector stock pays the aggregate dividends of all banks and the

[5] Coletti , Duarte, Feunou, Meh, Zhang (2020).

producer sector stock pays as dividends the aggregate profits of all producers. Thus, all in all, we characterize the economy in terms of three financial assets: a real riskless bond, the stock of banks and the stock of the intermediate good producers. Households can only hold and trade the riskless bond and the stock of banks, while banks can trade and hold all three assets. Even though banks find one of the stocks redundant (it can be replicated by the other stock and the bond), the restriction on households that they cannot hold the stock of good producers implies that banks must hold both stocks in equilibrium. The bond has a net supply of zero while both of the stocks are in positive net supply normalized to one.

The riskless bond has a price $S_{0,t}$ in real terms that evolves according to $dS_{0,t} = S_{0,t}R_t dt$, where R_t is the equilibrium real short-term interest rate. The two risky assets are indexed by $j \in \{banks, goods\}$. Each stock has real price S_{jt} that satisfies

$$\frac{dS_{jt}}{S_{jt}} = \tilde{\mu}_{jt}dt + \sigma_{jt}dB_t, \tag{16}$$

where $\tilde{\mu}_{jt}$ is the real expected return (including dividends) and σ_{jt} is the volatility.[6] We define real expected excess returns — the risk premium — by $\mu_{jt} \equiv \tilde{\mu}_{jt} - R_t$, and the market price of risk by $\eta_t \equiv \mu_{jt}/\sigma_{jt}$. We note that η_t is not indexed by j since, in equilibrium, the absence of arbitrage requires that the price of risk is the same for both stocks, that is, $\eta_t = \mu_{banks,t}/\sigma_{banks,t} = \mu_{goods,t}/\sigma_{goods,t}$. The market price of risk is a measure of risk-adjusted expected excess returns and captures the compensation that agents require in equilibrium to hold the risk associated with the Brownian motion B_t. More precisely, the market price of risk is the equilibrium risk premium earned per unit of exposure to B_t.

Last, we define the real state price density (SPD) Q_t as the solution to

$$dQ_t \equiv -Q_t R_t dt - Q_t \eta_t dB_t, \tag{17}$$

$$Q_0 \equiv 1. \tag{18}$$

No arbitrage implies that the price of an asset that pays a stream of cash flows $\{D_s\}_{s\geq t}$ is $\mathbb{E}_t\left[\int_t^\infty (Q_s/Q_t)D_s ds\right]$. The ratio Q_s/Q_t will turn out to be the equilibrium stochastic discount factor (SDF) of the representative household.

[6]With some abuse of language, we refer to S_t as the "price" of the stock instead of using the more precise "gain process" terminology.

Although at this point we do not know whether in equilibrium $\sigma_t \neq 0$ a.s. (which is required for markets to be complete), we nevertheless proceed as if were true, since it does turn out to be true anyway (which is easy to verify once the model is solved without having assumed $\sigma_t \neq 0$ in the first place).

5.3 Banks

5.3.1 Optimization Problem

There is a continuum of mass one, identical, atomistic, and infinitely lived banks. The representative bank solves a Merton portfolio problem augmented by a VaR constraint and preference shocks, with all variables expressed in real terms:

$$\max_{\{\theta_s, f_s\}_{s\geq t}} \mathbb{E}_t \left[\int_t^\infty e^{-\beta(s-t)} e^{\zeta_s} \log\left(f_s X_s\right) ds \right] \tag{19}$$

$$s.t.$$

$$\frac{dX_t}{X_t} = \left(R_t - f_t + \theta_t \mu_t\right) dt + \theta_t \sigma_t dB_t, \tag{20}$$

$$VaR_{\tau,\alpha}\left(t, \theta_t, f_t\right) \leq a_V X_t, \tag{21}$$

$$d\zeta_t = -\frac{1}{2} g_t^2 dt - g_t dB_t, \tag{22}$$

$$dg_t = -\kappa_g \left(g_t - m_g\right) dt + \sigma_g dB_t. \tag{23}$$

The variable X_t is the wealth of the bank, e^{ζ_t} is a preference shock driven by the exogenous stochastic processes in equations (22)-(23), f_t is the share of wealth paid out as dividends to shareholders, and θ_t is the share of wealth invested in the two risky stocks. Equation (20) is the budget constraint. Equation (21) is a VaR constraint. In its maximization problem, the bank takes X_t and the paths $\{R_s, \mu_s, \sigma_s, \zeta_s, g_s\}_{s\geq t}$ as given.

Because one of the stocks can always be replicated by the other stock and the bond, the portfolio choice of the bank only determines the allocation of wealth between the bond and a portfolio of the two stocks, but not the allocation of wealth to each of the two stocks individually. Therefore, μ_t and σ_t in the portfolio choice problem of the bank should be interpreted as the drfit and volatility of the protfolio of stocks and not as vectors of drifts and volatilities that contain the drifts and volatilities of each stock. Similarly, θ_t is the share of wealth invested in the risky portfolio (a single number) and not a vector of portfolio weights for each stock. Given θ_t, the amount that the bank invests in each of the two stocks individually is determined by the market clearing condition that banks hold the entire supply of good producers' stock.

The budget constraint in (20) states that wealth changes are equal to changes in the value of the bank's portfolio minus the dividends paid to shareholders. The bank will be levered most of the time with $\theta_t > 1$ (borrowing can be done by short-selling the bond, which can be also interpreted as issuing riskless demand deposits). When $\theta_t > 1$, the balance sheet is[7]

Assets	Liabilities
risky assets $\theta_t X_t$	debt $(\theta_t - 1) X_t$
	equity X_t

We do not distinguish between book and market values, as they are identical in this model. Log-preferences imply that $X_t > 0$, so the representative bank never undergoes bankruptcy, consistent with its ability to lend and borrow at the riskless rate.

In equation (21), $VaR_{\tau,\alpha}(t, \theta_t, f_t)$ is the value-at-risk of the bank's wealth with horizon $\tau > 0$ at level $\alpha \in (0, 1/2]$. The constant $a_V \in (0, 1)$ defines the VaR limit as a share of wealth X_t. Informally, the constraint says that the bank cannot take too much left tail or downside risk. More formally, $VaR_{\tau,\alpha}(t, \theta_t, f_t)$ is defined as the negative of the α^{th} quantile of the distribution of changes in wealth between t and $t + \tau$, conditional on time-t information, assuming the portfolio weight θ_t, dividends f_t, the interest rate R_t, and the risky asset coefficients μ_t and σ_t remain constant at their time-t levels in the interval $[t, t + \tau]$.[8] For example, if $\tau = 1$ month, $\alpha = 0.05$ and $VaR_{\tau,\alpha}$ equals \$100 million, wealth will drop by \$100 million or more over the next month with only 5% probability. If $a_V = 0.1$, the VaR constraint is satisfied if and only if \$100 million is less than 10% of the bank's wealth at time t. Lower values of a_V represent a more restrictive VaR constraint that prescribes smaller tail losses.

Since $B_{t+\tau} - B_t$ is normally distributed, the quantile function of the normal distribution gives

$$\begin{aligned} VaR_{\tau,\alpha}(t, \theta_t, f_t) &= X_t \left[1 - \exp \left\{ \frac{1}{dt} \mathbb{E}_t [d \log X_t] \tau + \mathcal{N}^{-1}(\alpha) \sqrt{\frac{1}{dt} Var_t (d \log X_t) \tau} \right\} \right] \\ &= X_t \left[1 - \exp \left\{ (R_t - f_t + \theta_t \mu_t - \frac{1}{2} \theta_t^2 \sigma_t^2) \tau + \mathcal{N}^{-1}(\alpha) |\theta_t \sigma_t| \sqrt{\tau} \right\} \right], \end{aligned}$$

where $\mathcal{N}^{-1}(\alpha) < 0$ is the inverse cumulative distribution function of a standard normal random variable.[9]

[7] If the bank is not levered ($\theta_t \le 1$), debt is zero, assets are given by $\max\{0, \theta_t X_t\} + (1 - \theta_t) X_t$ and liabilities by $\min\{0, \theta_t X_t\} + X_t$.

[8] This is not only the standard in the literature (see Cuoco et al. (2008) and Leippold et al. (2006) for discussions), it is also generally the case in practice. One reason to fix variables at their current levels is to try to avoid the possibility of gaming the constraint through promised yet unenforceable changes in future behavior, or through projections for the price processes of assets that are deceitfully biased. Fixing the relevant variables at their current levels (or at historical averages) is also required whenever the VaR constraint is regulatory in nature. Last, fixing variables at their current levels can be thought of as an approximation of a VaR without fxing variables that is very accurate for small τ and deteriorates as τ increases.

[9] For a generic stochastic process $dx_t = m_t dt + v_t dB_t$, we use the notation $\frac{1}{dt} E_t [dx_t]$ or $drift(dx_t)$ to mean $\frac{d}{ds} E_t [x_s]\big|_{s=t} = m_t$. We use $\frac{1}{dt} Var_t (dx_t)$ to mean $\frac{d}{ds} Var_t (x_s)\big|_{s=t} = v_t^2$, $stoch(dx_t)$ to mean v_t, and $Vol_t(dx_t)$

$VaR_{\tau,\alpha}$ is larger when the conditional distribution of wealth has lower mean or higher variance. The mean is lower when dividends are paid at a higher rate or the portfolio (of bonds and risky assets together) has lower expected return. The variance is higher when the bank's portfolio of risky assets is more volatile, which can occur because the underlying risky assets are themselves more volatile (larger $|\sigma_t|$) or because the bank invests a larger share of wealth in risky assets ($|\theta_t|$ is higher).

An important consideration when we later study monetary policy is how the risk-free rate R_t affects $VaR_{\tau,\alpha}$. When the bank is levered ($\theta_t > 1$), increasing R_t lowers the conditional mean of the wealth distribution ceteris paribus, since the bank has a negative portfolio position in the bond. The "price of leverage" has increased and honoring debt has just become more expensive. This lowering of the conditional mean of the wealth distribution is a budget constraint effect that is present even when there is no VaR constraint. In turn, the lower conditional mean of the wealth distribution implies a higher $VaR_{\tau,\alpha}$. On the other hand, R_t does not affect the conditional variance of the wealth distribution because riskless bonds are not exposed to the Brownian motion B_t. Thus, the direct effect of higher real rates is to tighten the VaR constraint (and to loosen it when real rates are lower). Changes in R_t can also affect $VaR_{\tau,\alpha}$ through general equilibrium effects that we study later on.

The process for g_t in (23) is an Ornstein–Uhlenbeck process— analogous to an $AR(1)$ process in discrete time — with mean-reversion parameter $\kappa_g > 0$, constant volatility $\sigma_g > 0$, and long-run mean m_g. We choose the specific form of the process for ζ_t in (22) so that the preference shock e^{ζ_t} is a change of probability measure (a Radon–Nikodym derivative), which makes the preference shock interpretable as a belief shock. The bank's problem under the "bank measure" defined by e^{ζ_t} is

$$\max_{\{\theta_s, f_s\}_{s\geq t}} \mathbb{E}_t^{bank}\left[\int_t^\infty e^{-\beta(s-t)} \log\left(f_s X_s\right) ds\right] \tag{24}$$

$$s.t.$$

$$\frac{dX_t}{X_t} = \left(R_t - f_t + \theta_t\left(\mu_t - \sigma_t g_t\right)\right) dt + \theta_t \sigma_t dB_t^{bank}, \tag{25}$$

$$VaR_{\tau,\alpha}^{bank}\left(t, \theta_t, f_t\right) \leq a_V X_t, \tag{26}$$

$$dg_t = -\left(\kappa_g + \sigma_g\right)\left(g_t - \frac{\kappa_g m_g}{\kappa_g + \sigma_g}\right) dt + \sigma_g dB_t^{bank}, \tag{27}$$

The conditional expectation $\mathbb{E}_t^{bank}[\cdot]$ and the value-at-risk $VaR_{\tau,\alpha}^{bank}(t, \theta_t, f_t)$ are now computed according to the processes (25) and (27) rather than (20) and (22)-(23) using that, under the bank measure, $dB_t^{bank} = dB_t + g_t dt$ are increments of a standard Brownian motion.

to mean $|v_t|$.

5.3.2 Solution

The solution to the bank problem (24)-(27) is[10]

$$\theta_t = \frac{1}{\gamma_t}\theta_{M,t}, \tag{28}$$

$$f_t = u\left(\gamma_t, \eta_t^{bank}\right) f_M, \tag{29}$$

where $\eta_t^{bank} \equiv \eta_t - g_t$ is the bank's "effective"price of risk under the bank measure, γ_t is defined by $\gamma_t \equiv \max\{1, \hat{\gamma}_t\}$ with $\hat{\gamma}_t$ such that

$$VaR_{\tau,\alpha}^{bank}\left(t, \frac{1}{\hat{\gamma}_t}\theta_{M,t}, u\left(\hat{\gamma}_t, \eta_t^{bank}\right) f_M\right) = a_V X_t, \tag{30}$$

and the function u is defined by

$$u(\gamma, \eta) \equiv 1 + \frac{\sqrt{\tau}\,|\eta|}{\mathcal{N}^{-1}(\alpha)}\left(1 - \frac{1}{\gamma}\right). \tag{31}$$

To understand this solution, we first look at the solution to the bank problem if the VaR constraint were removed:

$$\theta_{M,t} = \frac{\eta_t^{bank}}{\sigma_t} = \frac{\eta_t - g_t}{\sigma_t},$$

$$f_M = \beta.$$

This solution is identical to the familiar solution of a Merton portfolio problem. The optimal share of wealth invested in risky assets, $\theta_{M,t}$, is higher when the price of risk η_t^{bank} is higher, and when volatility σ_t is lower. Given η_t and σ_t, the value of g_t shifts the bank's demand for the risky asset by shifting the bank's perceived price of risk; under the bank measure, the bank has no preference shocks but believes expected returns are $\mu_t - \sigma_t g_t$ rather than μ_t. The optimal portfolio is the mean-variance efficient portfolio. Hedging demand is zero due to log-utility. [11] Also because of log-utility, the income and substitution effects for dividends cancel each other, and dividends are paid at a constant rate equal to the time-preference rate $f_M = \beta$.

We now return to the discussion of the full bank problem that includes the VaR constraint. According to (90), the optimal share of wealth invested in risky assets is the same as in the

[10] The derivation of the solution to the bank's problem can be found in Appendix C.

[11] Hedging demand, sometimes also called "dynamic" or "intertemporal" demand, is the demand for risky assets arising from changes in the investment opportunity set. With log-utility, the wealth effects exactly cancel out the re-allocation effects, giving a hedging demand of zero irrespective of the investment opportunity set.

The remaining demand for risky assets is called the "static" or "myopic" demand. It is the demand for the mean-variance efficient portfolio.

Merton case, $\theta_{M,t}$, but divided by $\gamma_t \geq 1$. We refer to γ_t as the bank's effective risk aversion, since $\theta_{M,t}/\gamma_t$ mirrors the static demand $\theta_{M,t}/\gamma$ (the demand for the mean-variance efficient portfolio) of the optimal portfolio in a standard Merton problem without a VaR constraint but with CRRA utility with coefficient γ. As in the unconstrained Merton case, log-utility continues to produce no hedging demand, which means that the optimal portfolio is still the mean-variance efficient portfolio. Therefore, the VaR constraint distorts the size, but not the composition, of the bank's portfolio.

Equation (30) determines the level of effective risk aversion $\hat{\gamma}_t$ that would make the VaR constraint (26) hold with equality under optimal bank behavior. The actual level of effective risk aversion is $\gamma_t = \max\{1, \hat{\gamma}_t\}$, which takes into account that $\hat{\gamma}_t < 1$ means that the VaR constraint is not binding. The degree of effective risk aversion of the bank can be interpreted as a measure of the tightness of the VaR constraint. Indeed, γ_t is a one-to-one function of the Lagrange multiplier λ_t associated to the VaR constraint:

$$1 + \beta\tau\lambda_t = \frac{1}{u\left(\gamma_t, \eta_t^{bank}\right)} \tag{32}$$

We can therefore interpret the function u in equation (31) as providing a conversion between the Lagrange multiplier of the VaR constraint and effective risk aversion.

For a given η_t^{bank}, when the multiplier λ_t is larger, the constraint is tighter and effective risk aversion is higher. The multiplier is zero and the constraint not binding if and only if $\gamma_t = 1$. When the bank is unconstrained, its effective risk aversion of $\gamma_t = 1$ is equal to its structural risk aversion of 1 given by the log-utility. Since $\gamma_t = \max\{1, \hat{\gamma}_t\} \geq 1$, the bank never invests a larger share of wealth in risky assets than when the VaR constrained is removed from the problem. In this sense, the VaR is thus effective at curbing risk-taking.

For a given Lagrange multiplier λ_t, (32) shows that the bank's effective risk aversion γ_t is decreasing in the market price of risk η_t^{bank}. Keeping the tightness of the VaR constraint constant, higher risk-adjusted excess returns are associated with a lower effective risk aversion. One implication is that a binding VaR constraint amplifies the responsiveness of the bank's optimal portfolio to fluctuations in the price of risk. An increase in η_t induces changes in θ_t for two reasons. First, $\theta_{M,t}$ increases. Second, γ_t decreases. Equation (90) shows that these two changes reinforce each other and create a larger increase in θ_t than if γ_t had remained fixed. The intuition is that when risk-adjusted expected returns increase, the bank invests more in risky assets not only because the risky assets are themselves more attractive — as it would even in the absence of a VaR constraint —but also because its risk capacity has increased as the VaR constraint is relaxed.

Equation (29) gives the optimal share of wealth that the bank pays out as dividends. It is equal to the Merton solution $f_M = \beta$ multiplied by $u\left(\gamma_t, \eta_t^{bank}\right)$, defined in (31). Since $u\left(\gamma_t, \eta_t^{bank}\right) \leq 1$, dividends are paid at a lower rate than in the Merton case, consistent with the bank having become more conservative regarding risk. However, as was the case for the optimal portfolio θ_t, dividends are more responsive to fundamentals vis-a-vis the case without the VaR constraint. Without the VaR constraint, the dividend rate $f_M = \beta$ is constant, so all fluctuations in f_t stem from the presence of the VaR constraint. Everything else equal, dividends are higher when effective risk aversion γ_t is lower and the VaR constraint less binding (lower λ_t). Dividends are also higher when the market price of risk η_t^{bank} is lower since, everything else equal, investing in the risky asset is less attractive, making dividend payouts more attractive by comparison.

5.3.3 The Bank's State Price Density

The bank's SPD implied by its optimization problem is

$$Q_t^{bank} = \frac{1}{\beta}\frac{e^{-\beta t}e^{\zeta_t}}{\lambda_{bc}X_t} + 2\tau\lambda_t\frac{e^{-\beta t}e^{\zeta_t}}{\lambda_{bc}X_t}$$

The first term captures the marginal value of issuing dividends. The second term gives the marginal value of relaxing the VaR constraint. The constant $\lambda_{bc} > 0$ is such that $Q_0^{bank} = 1$ and can be interpreted as the time-0 Lagrange multiplier on the bank's static budget constraint. That the SPD of the bank moves one-for-one with e^{ζ_t} and is proportional to $1/X_t$ follow directly from the preferences of the bank in equation (24).

5.3.4 Some Thoughts on the Bank's Preference Shocks and VaR Constraint

By writing $\mu_t - \sigma_t g_t = \sigma_t(\eta_t - g_t)$ in the drift of the wealth process in (25) and in the corresponding expression inside $VaR_{\tau,\alpha}^{bank}(t, \theta_t, f_t)$ in (25), we see that problem (24)-(27) has the same structure as problem (19)-(23) under the physical measure with four differences: (i) it has no preference shocks, (ii) the price of risk is $\eta_t^{bank} = \eta_t - g_t$, (iii) in (27), the mean-reversion parameter is higher and the long-run mean is lower and (iv) the probability measure is different. We can then interpret the bank's preference shock not just as a generic belief shock, but as one that makes the bank behave as if it did not have any preference shocks but believed that the market price of risk is η_t^{bank}, and that g_t is a less persistent process with a lower mean.

There are other ways to interpret the preference shocks. For example, they can represent shocks to time-preference rates, risk aversion, habits, optimism, information sets, biases or mistakes; they can be fully rational or behavioral. If the preference shock were removed from banks and its inverse included as preference shocks to households, all equilibrium outcomes would be identical, opening up the possibility for additional interpretations, such as consumer "demand" shocks. Some of the many alternatives just mentioned lend themselves to a structural interpretation, while others may be the reduced form of a deeper structure. We do not take a stance regarding which interpretation is correct, but point out that since we are interested in monetary policy, non-structural interpretations are adequate only to the extent that the underlying microfoundation of the preference shock remains unaffected by changes in monetary policy.

The same logic applies to interpretations of the VaR constraint. It can be understood as a literal VaR constraint, or as the reduced-form of other risk-management constraints, other regulations, stress-tests, agency problems, and so on. Again, any interpretation can be applied to our results as long as whatever mechanism gives rise to the VaR constraint remains unchanged by monetary policy.

Adrian & Shin (2013b) offer empirical evidence and a theoretical model that jointly microfounds a VaR constraint and shocks that, like the bank preference shocks, shift its tightness. In their model, the VaR constraint is the optimal incentive-compatible contract between an intermediary and its creditor when faced with a risk-shifting moral hazard problem. The shocks in their model are to the state of the business cycle. Their results apply to a large family of shock distributions determined by extreme-value theory. The shape of the optimal contract is not affected by monetary policy, as would also be arguably the case for private contracts in actuality.[12] Therefore, for our purpose of studying monetary policy, were this our preferred interpretation, there would be no need to embed the microfoundations of the VaR constraint into the bank problem. Our preference shocks can also be directly mapped to the business cycle shocks in Adrian & Shin (2013b) by interpreting them as shocks that affect the business cycle through broad financial conditions (in the case of shocks to banks) or through broad economic conditions (if the shocks were interpreted as demand shocks from households). In our empirical results, we focus on financial conditions, which motivates our choice of having the shocks appear on the bank side of the economy rather than in households, and to interpret them as shocks to financial conditions. The results in Adrian & Shin (2013b) therefore show that, at the very least, there exist coherent and empirically relevant microfoundations for our bank setting.

The reduction in risk induced by the VaR constraint is compatible with the risk-shifting moral hazard interpretation of Adrian & Shin (2013b). In their model, the optimal contract reduces the amount of risk taken by the intermediary to levels closer to what the creditor prefers. In our model, the household performs the role of the creditor, delegating the financing of good producers to the bank. The household can have a CRRA coefficient larger than the

[12] We note that monetary policy can and indeed does change the tightness of the constraint — the numerical value of VaR itself — in both their model and ours. What remains invariant under changes in monetary policy is that the optimal contract places a limit on VaR, that is, has the functional form of a VaR constraint.

bank's ($\gamma \geq 1$). If this indeed the case, the bank would, absent the VaR constraint, take too much risk from the point of view of the household, just as in the risk-shifting problem. The VaR constraint acts as the contract that induces the bank to have effective risk aversion of γ_t rather than the log-utility risk aversion of 1 when choosing its portfolio, reducing the risks it takes from $\theta_{M,t}$ to $\theta_{M,t}/\gamma_t$. The closer γ_t is to the household's γ, the closer the alignment between the household's and the bank's preferred size of investment in the mean-variance efficient portfolio.

5.4 Good-producing firms

Since the structure of good-producing firms is exactly as in the standard New Keynesian (NK), we relegate details to Appendix A. Their monopolistically competitive market structure and nominal price rigidities a la Calvo give rise to a non-linear Phillips curve that can be linearized to

$$d\pi_t = (\beta\pi_t - \kappa y_t)\, dt \tag{33}$$

where π_t is inflation, y_t is log-output, and $\kappa > 0$ is a reduced-form parameter that depends on the structural parameters of the model and captures, among other things, the degree of price stickiness.

Intermediate good producers are financed by issuing stocks that are bought by banks. The real profits for the producer of intermediate good i are

$$D_{t,goods}(i) = \frac{P_t(i)\, Y_t(i)}{P_t} - MC_t Y_t(i)$$

Aggregating across firms gives the aggregate profits for the sector, which are paid out as dividends to shareholders

$$\begin{aligned} D_{t,goods} &= \int_0^1 D_{t,goods}(i)\, di \\ &= \frac{1}{P_t}\int_0^1 P_t(i)\, Y_t(i)\, di - MC_t \int_0^1 Y_t(i)\, di \\ &= \frac{1}{P_t}(P_t Y_t) - MC_t (v_t Y_t) \\ &= (1 - v_t MC_t)\, Y_t \end{aligned} \tag{34}$$

where

$$v_t \equiv \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{-\varepsilon} di$$

is a measure of the aggregate loss of output due to price distortions.

5.5 Households

There is a continuum of mass one, identical, atomistic, and infinitely lived households. The representative household solves

$$\max_{\{C_s,N_s,\omega_s\}_{s\geq t}} \mathbb{E}_t \int_t^\infty e^{-\beta(s-t)} \left(\frac{C_s^{1-\gamma}}{1-\gamma} - \frac{N_s^{1+\xi}}{1+\xi} \right) ds, \tag{35}$$

s.t.

$$\frac{dF_t}{F_t} = \left(R_t + \omega_t \mu_{banks,t} - \frac{1}{F_t} \left(C_t - (1-s_t) \frac{W_t}{P_t} N_t + T_t \right) \right) dt + \omega_t \sigma_{banks,t} dB_t, \tag{36}$$

$$\lim_{s\to\infty} \mathbb{E}_t \left[Q_s F_s \right] = 0. \tag{37}$$

The household maximizes utility by choosing the path of final good consumption C_t, supply of labor N_t, and share of wealth (portfolio weight) on the risky asset ω_t, subject to the budget constraint (36), and a solvency constraint (transversality condition). The variable F_t is the household's real financial wealth, R_t is the riskless real rate, $\mu_{banks,t}$ and $\sigma_{banks,t}$ are the is the expected returns and volatility of the bank stock, s_t is a labor tax (or subsidy, if negative), T_t are lump-sum taxes, and Q_t is the real state-price density (SPD). In its maximization, the household takes $\left\{P_t, R_t, W_t, \mu_{banks,t}, \sigma_{banks,t}, s_t, T_t, Q_t\right\}_{s\geq t}$ and F_t as given.

The dynamic flow budget constraint in equation (36) states that changes in the household's financial wealth must be equal to the payoff on financial assets (which can be negative), minus nominal consumption expenditures, plus after-tax nominal labor income, plus lump-sum taxes. The transversality condition in equation (37) is a no-Ponzi condition for the household.

The FOC for consumption and labor are

$$Q_t = e^{-\beta t} C_t^{-\gamma}, \tag{38}$$

$$C_t^{\gamma} N_t^{\xi} = (1-s_t) \frac{W_t}{P_t}. \tag{39}$$

With complete markets, the flow budget constraint (35) and the transversality condition (37) are equivalent to the static budget constraint

$$F_t = \mathbb{E}_t \int_t^\infty \frac{Q_s}{Q_t} \left(C_s - (1-s_s) \frac{W_s}{P_s} N_s + T_s \right) ds \tag{40}$$

Given the optimal choice of consumption and labor, the portfolio weight ω_t is implicitly determined by equation (36).

6 Asset Pricing

6.1 Stock Prices and Portfolio Positions

The prices $S_{banks,t}$ and $S_{goods,t}$ of the two stocks in the economy are equal to the expected present discounted value of their dividends:

$$\begin{aligned} S_{t,goods} &= E_t\left[\int_t^\infty \frac{Q_s}{Q_t} D_{s,goods} ds\right] \\ S_{t,banks} &= E_t\left[\int_t^\infty \frac{Q_s}{Q_t} \left(f_s X_s\right) ds\right] \end{aligned}$$

Since θ_t finances the cash flow $f_t X_t$ and $S_{banks,t}$ pays dividends $f_t X_t$, the absence of arbitrage implies that

$$S_{banks,t} = X_t$$

In addition, equation (34), goods market clearing, and the optimality conditions of the household and firms gives

$$D_{t,goods} = C_s - (1 - s_s) \frac{W_s}{P_s} N_s + T_s$$

which implies that the dividends of the stock of good producers are equal to the representative household's cash flows, and hence

$$S_{t,goods} = F_t$$

Market clearing of stocks and bonds, together with the above relations, allow us to split the share of wealth that the bank invests in its porfolio of risky assets, θ_t, into the portions invested in the stocks of good producers and banks,

$$\begin{aligned} \theta_{goods,t} &= \frac{F_t}{X_t} \\ \theta_{banks,t} &= \theta_t - \frac{F_t}{X_t} \end{aligned}$$

where $\theta_t = \theta_{goods,t} + \theta_{banks,t}$. The bank holds all of the good producer's stock, investing and amount $\theta_{goods,t} X_t = S_{t,goods} = F_t$ on it. We note that $\theta_{goods,t}$ can be written without direct reference to θ_t, which reflects that, in equilibrium, banks must be the ones holding all the good producer's stock, irrespective of optimality or its VaR constraint. The bank also invests $\theta_t X_t - F_t$ on the bank's stock. The remaining wealth, $X_t - (\theta_{goods,t} X_t) - (\theta_t X_t - F_t) = (1 - \theta_t) X_t$, is invested in riskless bonds.

Similarly, the household's portfolio position can be written as

$$\omega_t = 1 + (1 - \theta_t) \frac{X_t}{F_t}$$

The household invests $\omega_t F_t$ in the bank stock, and the rest of its wealth in the riskless bond. When the bank is levered with $\theta_t > 1$, the household must be long (have a positive position on) the riskless bond, since it is in zero net supply. A household with constant levels of consumption and employment —as, for example, in the first best allocation— must have no risk and $\omega_t = 0$. In this case, $\theta_{banks,t} = 1$ and $\theta_{goods,t} = F_t / X_t$.

6.2 The Market Price of Risk

We now find the price of risk η_t, as a function of the exogenous variables ζ_t and g_t. The key relation we use is that, because households and banks trade with each other in a complete market, their state price densities (equivalently, their stochastic discount factors) must agree

$$Q_t = Q_t^{bank}$$

Equating the stochastic parts of $-dQ_t/Q_t$ and $-dQ_t^{bank}/Q_t^{banks}$ gives

$$\eta_t = \frac{\eta_t - g_t}{\gamma_t} + g_t - stoch\left(d\log\left(\frac{1}{\beta} + 2\tau\lambda_t\right)\right) \tag{41}$$

The left-hand side of the equation gives η_t, the household's required compensation for taking on one unit of B_t-risk. The right-hand side gives the bank's required compensation for risk. The first term, $(\eta_t - g_t)/\gamma_t$, gives the bank's required risk compensation for changes in its wealth. Instead of η_t, the bank uses its "perceived" $\eta_t^{bank} = \eta_t - g_t$. Similarly, it uses its effective risk aversion γ_t, rather than its log-utility risk aversion of 1. The second term is the volatility comes from the direct effect of the preference shock e^{ζ_t} in the bank's preferences in equation (19). Becuase e^{ζ_t} appears directly in the preferences of the bank, its volatility of g_t shifts the volatility of bank's SPD one-for-one. The last term gives the risk compensation required from changes in the tightness of the VaR constraint.

When the VaR is not binding we have that $\gamma_t = 1$, $\lambda_t = 0$. In this case the right-hand side of equation (41) equals η_t and hence equation (41) holds for any η_t. The intuition is that if the VaR were never binding, or absent altogether, the bank's pricing of risk would be identical to the household's at all times, and the bank's behavior would impose no restrictions on the equilibrium price of risk. If the VaR were indeed never binding, the bank would be a "pass-through" intermediary, simply relaying the dividends from the good-producing firms to the household without any frictions. In this case, the presence of the bank would have no influence on the economy's equilibrium outcome, recovering the standard textbook NK model.

However, when the VaR is binding, equation (41) fully determines η_t. We look for a Markov equilibrium, that is, an equilibrium in which all equilibrium variables can be written as a deterministic function of the two state variables ζ_t and g_t. In particular, we can write $\eta_t = \eta(\zeta_t, g_t)$ for some function $\eta(\cdot,\cdot)$. Under the Markov assumption, and using the optimality conditions of the bank's problem in equations (90)-(32) and Ito's lemma, equation (41) gives a partial differential equation (PDE) for $\eta(\zeta_t, g_t)$. The resulting PDE is linear and can be solved in closed-form. There are two different cases to consider.

The first case is when the VaR constraint is binding so tightly that the only way to satisfy it is to only hold a completely riskless portfolio, that is, $\theta_t = 0$. We note that although bonds are riskless (are not exposed to the Brownian motion B_t), the interest rate R_t itself is not necessarily constant, and hence the VaR can be binding even when the portfolio consists only of riskless bonds due to the riskiness of R_t. When $\theta_t = 0$, we have that

$$\eta(\zeta_t, g_t) = g_t + \frac{2\beta}{(R_t - A_V)(R_t - A_V + 2\beta)}\left(g\frac{\partial R_t}{\partial \zeta_t} - \sigma_g \frac{\partial R_t}{\partial g_t}\right) \tag{42}$$

The bank's risk compensation required by changes in the tightness of the VaR constraint only depends on R_t because the bank's portfolio consists only of riskless bonds whose payoffs are determined solely by R_t.

The second case is when the VaR constraint does bind but $\theta_t \neq 0$. When this happens, we get

$$\eta(\zeta_t, g_t) = g_t - N - \sqrt{\left(\frac{N}{\sqrt{1 + u(\zeta_t, g_t)}} - N - \frac{\beta}{N}\right)^2 - \frac{\beta^2}{N^2} - 2(R_t - A_V)} \tag{43}$$

where, to shorten notation, we have defined the constants

$$A_V \equiv \frac{1}{\tau}\log(1 - a_V)$$
$$N \equiv \frac{\mathcal{N}^{-1}(\alpha)}{\sqrt{\tau}}$$

and the function

$$u_t \equiv 4\beta \exp\left[K\left(\frac{g_t^2}{2} + \sigma_g \zeta_t\right) - \frac{N}{\sigma_g} g_t\right]$$

In the above definition for u_t, the function $K(\cdot)$ is an arbitrary function to be determined by the PDE's boundary condition. The boundary condition is that the function *eta* from equation (43) pastes continuously with the function η in equation (42). Because the function η in equation

(42) depends only on R_t, $K(\cdot)$ is fully pinned down by the behavior of R_t.

6.3 Monetary Policy

The nominal interest rate i_t under control of the central bank satisfies the Fisher equation $R_t = i_t - \pi_t$. Using this Fisher equation in (42) and (43) explicitly characterizes the effects of monetary policy on the pricing of risk.

There are two channels through which monetary policy influences the price of risk. First, changes in i_t have a direct impact on the value-at-risk of the bank, as discussed in Section 5.3: Keeping inflation constant, higher interest rates tighten the VaR constraint. A tighter VaR constraint increases the bank's effective risk aversion γ_t. The risk compensation that the bank requires to hold risk thus increases. More precisely, when $\theta_t \neq 0$, equation (43) implies that

$$\frac{\partial \eta_t}{\partial i_t} = \frac{1}{g_t - \eta_t - N} > 0.$$

The second channel is a general equilibrium channel. Changes in i_t influence the price of risky assets, both through discount rates and through cash flows. These effects are reflected in equation (43) by u_t. The value of u_t is not influenced by the current value of interest rates, but by the shape of the *interest rate policy rule*. Indeed, the shape of K is determined by the values of η_t where the values of η in equations (42) and (43) paste continuously. The value η_t and the set of (ζ_t, g_t) where the continuous pasting occurs are fully determined by the corresponding values of R_t and its derivatives – and hence also by the values of i_t and its derivatives. In this channel, it is not the current level of interest rates that matter, but the expectation of the future actions of the central bank and the future behavior of inflation, that matter.

7 GDP-at-Risk

To connect the empirical results in Section 2 to the model, just as we defined VaR for the bank's wealth, we can define "GDP-at-Risk" by

$$GaR_t \equiv Y_t \left[1 - \exp\left\{ \frac{1}{dt}\mathbb{E}_t \left[d\log Y_t \right] \tau + \mathcal{N}^{-1}(\alpha) \sqrt{\frac{1}{dt} Var_t \left(d\log Y_t \right) \tau} \right\} \right]$$

Or, in logs,

$$\begin{aligned} V_t &\equiv -\log\left(1-\frac{VaR_{\tau,\alpha}\left(Y_t\right)}{Y_t}\right) \\ &= -\frac{1}{dt}\mathbb{E}_t\left[d\log Y_t\right]\tau - \mathcal{N}^{-1}\left(\alpha\right)\sqrt{\frac{1}{dt}Var_t\left(d\log Y_t\right)\tau} \end{aligned} \tag{44}$$

Differentiating the household's Euler equation in (38) gives

$$d\log Y_t = \frac{1}{\gamma}\left(R_t - \beta + \frac{1}{2}\eta_t^2\right)dt + \frac{1}{\gamma}\eta_t dB_t \tag{45}$$

so that

$$V_t = \frac{1}{\gamma}\left(R_t - \beta + \frac{1}{2}\eta_t^2\right)\tau + \frac{\mathcal{N}^{-1}\left(\alpha\right)\sqrt{\tau}}{\gamma}\eta_t$$

Solving for R_t gives

$$R_t = -\frac{\gamma}{\tau}V_t - \frac{\mathcal{N}^{-1}\left(\alpha\right)}{\sqrt{\tau}}\left|\eta_t\right| + \beta - \frac{1}{2}\eta_t^2 \tag{46}$$

Plugging this equation in the bank's VaR constraint, equation (21), with equality, and using the relation between the bank's and household's SDF in equation (41) gives η_t as a function of g_t and V_t

$$\eta_t = -\frac{\gamma}{\tau g_t}V_t + S\left(g_t\right) \tag{47}$$

where S is a function of g_t only, given by

$$\begin{aligned} S\left(g_t\right) &= -\frac{1}{\tau g_t}\log\left(1-a_V\right) - \frac{\mathcal{N}^{-1}\left(\alpha\right)}{\sqrt{\tau}} + \frac{1}{2}g_t \\ &\quad -\frac{\sigma_g}{\sqrt{\tau}}\left(\frac{\tau\beta}{\mathcal{N}^{-1}\left(\alpha\right)} + \mathcal{N}^{-1}\left(\alpha\right)\right)\frac{\left|\Phi'\left(g_t\right)\right|}{\Phi\left(g_t\right)}\frac{1}{g_t} - \frac{\sigma_g^2}{2g_t}\left(\frac{\Phi'\left(g_t\right)}{\Phi\left(g_t\right)}\right)^2 \end{aligned}$$

Equation (47) is not a new result, it is just a way to express some of the results in the previous section in terms of V_t rather than R_t.

Linearizing equation (47) around $\hat{\eta}$ and plugging into (45) gives

$$d\log Y_t = \frac{1}{\gamma}\left(R_t - \beta + \gamma\hat{\eta}\xi\left(V_t - s_t - \frac{1}{2}\frac{\hat{\eta}}{\xi\gamma}\right)\right)dt + \xi\left(V_t - s_t\right)dZ_t \tag{48}$$

where $\hat{\eta}$ and ξ are linearization constants and s_t is a random variable that depends linearly on

g_t. We therefore have

$$\begin{aligned}\mathbb{E}_t\left[dy_t\right] &= \frac{1}{\gamma}\left(R_t-\beta+\gamma\hat{\eta}\xi\left(V_t-s_t-\frac{1}{2}\frac{\hat{\eta}}{\xi\gamma}\right)\right)\\ \mathbb{V}_t\left[dy_t\right] &= \xi\left(V_t-s_t\right)\end{aligned}$$

Solving for R_t and V_t in (44) to get

$$R_t=\beta-\frac{\gamma}{\tau}\left(\xi\sqrt{\tau}\left(\alpha+\sqrt{\tau}\eta\right)+1\right)V_t+\gamma\xi\left(\frac{\alpha}{\sqrt{\tau}}+\eta\right)s_t+\frac{1}{2}\eta^2 \tag{49}$$

Plug in (49) into (48) to get

$$dy_t=-\frac{\alpha\sqrt{\tau}\xi+1}{\tau}\left(V_t-\frac{\alpha\sqrt{\tau}\xi}{\alpha\sqrt{\tau}\xi+1}s_t\right)dt+\xi\left(V_t-s_t\right)dZ_t \tag{50}$$

Use

$$\begin{aligned}\mathbb{E}_t\left[dy_t\right] &= -\frac{\alpha\sqrt{\tau}\xi+1}{\tau}\left(V_t-\frac{\alpha\sqrt{\tau}\xi}{\alpha\sqrt{\tau}\xi+1}s_t\right)\\ \mathbb{V}_t\left[dy_t\right] &= \xi\left(V_t-s_t\right)\end{aligned}$$

and then eliminating V_t to get

$$\mathbb{E}_t\left[dy_t\right]=-\frac{1+\alpha\sqrt{\tau}\xi}{\tau\xi}\mathbb{V}_t\left[dy_t\right]-\frac{1}{\tau}s_t \tag{51}$$

We have thus obtained the mean-volatility line of Figure 2. Equation (51) also makes clear that the shocks s_t are shifts to vulnerability that shift the mean-volatility line up and down, while all other changes in the economy involve moving along the mean-volatility line. Empirically, the slope is negative and the intercept is positive, which in term of the model parameters implies that

$$\begin{aligned}-\frac{1+\alpha\sqrt{\tau}\xi}{\tau\xi} &< 0\\ -\frac{\overline{s}}{\tau} &> 0\end{aligned}$$

and therefore $\overline{s}<0$ and

$$\begin{aligned}\xi &> 0 \text{ and } 1+\alpha\sqrt{\tau}>0\\ &\text{or}\\ \xi &< 0 \text{ and } 1+\alpha\sqrt{\tau}<0\end{aligned}$$

To match empirical estimates, we set

$$\begin{aligned}\alpha &= -1.645\\ \sqrt{\tau} &= 1\end{aligned}$$

To match the actual slope and intercept

$$\begin{aligned}-\frac{1+\alpha\sqrt{\tau}\xi}{\tau\xi} &= -1.15\\ \bar{s} &= -0.67\tau\end{aligned}$$

which gives

$$\begin{aligned}\xi &= 0.36\\ \bar{s} &= -0.67\end{aligned}$$

We identify $s_t - \bar{s}$ with the residuals of the regression of $\mathbb{V}_t\left[dy_t\right]$ on $\mathbb{V}_t\left[dy_t\right]$. The standard deviation and $AR(1)$ coefficient of these residuals then identify σ_s and κ, respectively. Since

$$\begin{aligned}Std\left(-\frac{1}{\tau}\left(s_t-\bar{s}\right)\right) &= 0.62\\ AR(1) &= 0.12\end{aligned}$$

we get, converting to annualized values

$$\begin{aligned}\kappa &= -\log\left(0.12\right)=2.12\\ \sigma_s &= 0.31\end{aligned}$$

8 Monetary Policy

8.1 Optimal Monetary Policy

The central bank minimizes a quadratic loss function over the output gap and inflation

$$L\left(y_t,\pi_t\right)=\min_{i_t}\mathbb{E}_t\int_t^{\infty}e^{-t\beta}\left(y_t^2+\pi_t^2\right)dt. \tag{52}$$

subject to the dynamics of the economy and with perfect commitment. Minimizing the quadratic loss function is a standard approach in the NK literature, as Rotemberg & Woodford (1997), Rotemberg & Woodford (1999) and Woodford (2003) have shown that aggregate welfare can be approximated by such a loss function. Since banks do not consume, the quadratic approximation in equation (52) is also valid in our model with banks.

We focus on the case with fully fixed prices first. The interest rate R_t can be eliminated from the optimization problem, so that the central bank's problem can be written as

$$L\left(y_t, s_t\right) = \min_{\{V_s\}_{s=t}^{\infty}} \mathbb{E}_t \int_t^{\infty} e^{-s\beta} y_s^2 ds \tag{53}$$

$$s.t.$$

$$V_t = \frac{-\gamma^{-1}\left(R_t-\beta\right)+\alpha\xi s_t\sqrt{\tau}+\hat{\eta}\xi\left(s_t+\frac{1}{2}\frac{\hat{\eta}}{\xi\gamma}\right)\tau}{1+\alpha\xi\sqrt{\tau}+\hat{\eta}\xi\tau} \tag{54}$$

$$dy_t = -\frac{\alpha\sqrt{\tau}\xi+1}{\tau}\left(V_t-\frac{\alpha\sqrt{\tau}\xi}{\alpha\sqrt{\tau}\xi+1}s_t\right)dt+\xi\left(V_t-s_t\right)dZ_t \tag{55}$$

$$ds_t = -\kappa\left(s_t-\overline{s}\right)+\sigma_s dZ_t \tag{56}$$

The central bank thus effectively picks V_t, which is connected to R_t in a one-to-one fashion by

$$R_t = \beta-\frac{\gamma}{\tau}\left(\xi\sqrt{\tau}\left(\alpha+\sqrt{\tau}\eta\right)+1\right)V_t+\gamma\xi\left(\frac{\alpha}{\sqrt{\tau}}+\eta\right)s_t+\frac{1}{2}\eta^2$$

The Hamilton-Jacobi-Bellman (HJB) equation for the central banker's optimization is

$$\begin{aligned} 0 \;=\; & \min_V\left\{y^2-\beta L-\frac{\partial L}{\partial y}\frac{\alpha\sqrt{\tau}\xi+1}{\tau}\left(V-\frac{\alpha\sqrt{\tau}\xi}{\alpha\sqrt{\tau}\xi+1}s\right)+\frac{1}{2}\frac{\partial^2 L}{\partial y^2}\xi^2\left(V-s\right)^2\right\} \\ & -\kappa\left(s-\overline{s}\right)\frac{\partial L}{\partial s}+\frac{1}{2}\frac{\partial^2 L}{\partial s^2}\sigma_s^2 \end{aligned}$$

Intuitively, the HJB takes into account the current value of welfare, as well as the change in welfare associated with changes in the state variables y and s.

The first order condition is

$$0 = -\frac{\partial L}{\partial y}\frac{\alpha\sqrt{\tau}\xi+1}{\tau}+\frac{\partial^2 L}{\partial y^2}\xi^2\left(V-s\right) \tag{57}$$

$$V = \frac{\partial L}{\partial y}\frac{\alpha\sqrt{\tau}\xi+1}{\tau\xi^2}\left(\frac{\partial^2 L}{\partial y^2}\right)^{-1}+s \tag{58}$$

Hence at the optimum, vulnerability is proportional to s, and depends on the first and second derivative of welfare with respect to output. It is also noteworthy that $\frac{\alpha\sqrt{\tau}\xi+1}{\tau\xi^2}$, which defines the slope of output volatility with respect to expected output, appears in the FOC.

We look for a quadratic solution of the form

$$L(y,x) = c_0 + c_1 y + c_2 y^2 + c_3 s + c_4 s^2 + c_5 ys$$

where $c.$ are constants.

Plugging into the HJB, and using

$$\begin{aligned}
\frac{\partial L}{\partial y} &= c_1 + 2c_2 y + c_5 s \\
\frac{\partial^2 L}{\partial y^2} &= 2c_2 \\
\frac{\partial L}{\partial s} &= c_3 + 2c_4 s + c_5 y \\
\frac{\partial^2 L}{\partial s^2} &= 2c_4
\end{aligned}$$

we get the following system of equations on the coefficients $c_0, ..., c_5$

$$\begin{aligned}
\left[y^2\right] &: \quad 0 = \left(-\beta - \frac{1}{\tau^2\xi^2}\left(\alpha\sqrt{\tau}\xi + 1\right)^2\right) c_2 + 1 \\
[ys] &: \quad 0 = \left(-\frac{2}{\tau}\right) c_2 + \left(-\frac{1}{\tau^2\xi^2}\left(2\alpha\sqrt{\tau}\xi + \alpha^2\tau\xi^2 + \beta\tau^2\xi^2 + 1\right)\right) c_5 \\
[y] &: \quad 0 = -\frac{1}{\tau^2\xi^2} c_1 \left(2\alpha\sqrt{\tau}\xi + \alpha^2\tau\xi^2 + \beta\tau^2\xi^2 + 1\right) \\
\left[s^2\right] &: \quad 0 = -\frac{1}{4\tau^2\xi^2 c_2}\left(c_5^2\left(2\alpha\sqrt{\tau}\xi + \alpha^2\tau\xi^2 + 1\right) + 4\tau\xi^2 c_2 c_5 + 4\beta\tau^2\xi^2 c_2 c_4\right) \\
[s] &: \quad 0 = -\frac{1}{2\tau^2\xi^2 c_2}\left(c_1 c_5\left(2\alpha\sqrt{\tau}\xi + \alpha^2\tau\xi^2 + 1\right) + 2\tau\xi^2 c_1 c_2 + 2\beta\tau^2\xi^2 c_2 c_3\right) \\
[const] &: \quad 0 = -\frac{1}{4\tau^2\xi^2 c_2}\left(c_1^2\left(2\alpha\sqrt{\tau}\xi + \alpha^2\tau\xi^2 + 1\right) + 4\beta\tau^2\xi^2 c_0 c_2 - 4\tau^2\xi^2\sigma_s^2 c_2 c_4\right)
\end{aligned}$$

with solution

$$
\begin{aligned}
c_0 &= \frac{\tau^2\xi^4\sigma_s^2\left(\left(\alpha\sqrt{\tau}\xi+1\right)^2+2\beta\tau^2\xi^2\right)}{\beta^2\left(\left(\alpha\sqrt{\tau}\xi+1\right)^2+\beta\tau^2\xi^2\right)^3} > 0 \\
c_1 &= 0 \\
c_2 &= \frac{\tau^2\xi^2}{\tau^2\xi^2\beta+\left(\alpha\sqrt{\tau}\xi+1\right)^2} > 0 \\
c_3 &= 0 \\
c_4 &= \frac{\xi^4\tau^2\left(\left(\alpha\sqrt{\tau}\xi+1\right)^2+2\beta\tau^2\xi^2\right)}{\beta\left(\left(\alpha\sqrt{\tau}\xi+1\right)^2+\beta\tau^2\xi^2\right)^3} > 0 \\
c_5 &= -\frac{2\tau^3\xi^4}{\left(\left(\alpha\sqrt{\tau}\xi+1\right)^2+\beta\tau^2\xi^2\right)^2} < 0
\end{aligned}
$$

To pick the optimal initial conditions, we minimize L with respect to y_0 taking s_0 as given

$$
\begin{aligned}
L\left(y_0, s_0\right) &= c_0 + c_1 y_0 + c_2 y_0^2 + c_3 s_0 + c_4 s_0^2 + c_5 y_0 s_0 \\
FOC &: \frac{\partial L}{\partial y_0} = 0 \\
SOC &: \frac{\partial^2 L}{\partial y_0^2} > 0
\end{aligned}
$$

The FOC and SOC can be solved to get

$$
\begin{aligned}
y_0^* &= -\left(\frac{c_1}{2c_2}+\frac{c_5}{2c_2}s_0\right) \\
&= \frac{\tau\xi^2 s_0}{\left(\alpha\sqrt{\tau}\xi+1\right)^2+\beta\tau^2\xi^2} \\
c_2 &> 0
\end{aligned}
$$

The optimal policy in terms of V_t is given by plugging in the optimal solution into the FOC in equation (57):

$$
V = \frac{\left(\alpha\sqrt{\tau}\xi+1\right)}{\tau\xi^2}y + \left(1-\frac{\left(\alpha\sqrt{\tau}\xi+1\right)}{\left(\alpha\sqrt{\tau}\xi+1\right)^2+\beta\tau^2\xi^2}\right)s \tag{59}
$$

This can be viewed as a "flexible inflation targeting rule" (see Svensson (1999), Svensson (2002) and Rudebusch & Svensson (1999)) or, more generally, as a linear optimal targeting criterion (Giannoni & Woodford (2012)). Even though vulnerability and its shocks, V_t and s_t, are not target variables, i.e., they do not appear in the loss function equation (52), they still enter the inflation targeting rule, the first-order condition given by equation (59). There are no independent target values for V_t and s_t that the central bank hopes to achieve. The reason V_t and s_t enter the targeting rule is that they forecast the conditional mean and variance of y_t even after controlling for the information already contained in the mean of y_t itself (more generally, in the means of y_t and π_t when a Phillips Curve is included). This is consistent with the empirical results in Table 1 and with the findings in Adrian et al. (2019), who show that financial conditions are excellent predictors of the tail of the GDP distribution in a way that non-financial variables are not. Alternatively, equation (59) can be interpreted as a traditional flexible inflation targeting rule in which the targets for inflation and/or output are time-varying and depend on V_t and s_t. It also important to note that even if a central bank decided not to condition its actions on V_t and s_t, the tradeoff between inflation and output –reflected in the coefficients of the rule in equation (59)– now depends on γ and ξ, the parameters that dictate the strength of the mean-variance tradeoff of output.

Using the optimal solution in the process for the output gap in equation (55), we then find that

$$\begin{aligned} dy_t &= -\frac{\alpha\sqrt{\tau}\xi+1}{\tau}\left(V_t-\frac{\alpha\sqrt{\tau}\xi}{\alpha\sqrt{\tau}\xi+1}s_t\right)dt+\xi\left(V_t-s_t\right)dZ_t \\ &= -\left(\frac{\left(\alpha\sqrt{\tau}\xi+1\right)^2}{\tau^2\xi^2}y_t+\frac{\beta\tau\xi^2}{\left(\alpha\sqrt{\tau}\xi+1\right)^2+\beta\tau^2\xi^2}s_t\right)dt \\ &\quad+\left(\frac{\left(\alpha\sqrt{\tau}\xi+1\right)}{\tau\xi}y_t-\frac{\xi\left(\alpha\sqrt{\tau}\xi+1\right)}{\left(\alpha\sqrt{\tau}\xi+1\right)^2+\beta\tau^2\xi^2}s_t\right)dZ_t \end{aligned}$$

Recalling that

$$\mathbb{E}_t\left[dy_t\right]=-\frac{1+\alpha\sqrt{\tau}\xi}{\tau\xi}\mathbb{V}_t\left[dy_t\right]-\frac{1}{\tau}s_t \tag{60}$$

And defining the slope as

$$M\equiv-\frac{1+\alpha\sqrt{\tau}\xi}{\tau\xi}$$

we get

$$V=-\frac{M}{\xi}y+\left(1+\frac{M}{\tau\xi\left(M^2+\beta\right)}\right)s$$

and

$$dy_t=-\left(M^2\times y_t+\frac{\beta/\tau}{M^2+\beta}\times s_t\right)dt-\left(M\times y_t-\frac{M/\tau}{M^2+\beta}\times s_t\right)dZ_t \tag{61}$$

The last equation shows that the magnitude of the tradeoff between stabilizing the mean and the variance of the output gap is given by the slope M of the mean-volatility line in Figure 2.

We can also express monetary policy as an interest rate rule. Using the FOC for V, the optimal interest rate is

$$
\begin{aligned}
R_t &= \beta - \frac{\gamma}{\tau}\left(\xi\sqrt{\tau}\left(\alpha+\sqrt{\tau}\eta\right)+1\right)V_t + \gamma\xi\left(\frac{\alpha}{\sqrt{\tau}}+\eta\right)s_t + \frac{1}{2}\eta^2 \\
&= \beta + \frac{1}{2}\eta^2 - \gamma\xi\left(\frac{\alpha}{\sqrt{\tau}}+\frac{1}{\tau}+\eta\right)V_t + \gamma\xi\left(\frac{\alpha}{\sqrt{\tau}}+\eta\right)s_t \\
&= \gamma M\left(\eta+\frac{\alpha}{\sqrt{\tau}}\right)\frac{M^2+\beta}{\beta+M^2+M/\tau\xi}y_t - \gamma\xi\left(\left(\eta+\frac{\alpha}{\sqrt{\tau}}\right)\frac{M/\tau\xi}{\beta+M^2+M/\tau\xi}+\frac{1}{\tau}\right)V_t \\
&\quad + \left(\frac{1}{2}\eta^2+\beta\right)
\end{aligned}
$$

The optimal interest rule can thus be viewed as an augmented Taylor rule. In addition to the output gap y and the equilibrium rate of interest r (and inflation π_t in the more general case), the level of vulnerability V enters the optimal rule. As before, the coefficients on y (and π in the more general case) depend on the parameters that define vulnerability ξ and γ and thus monetary policy is different from the typical NK model without vulnerabilities not only because vulnerability enters the augmented Taylor rule directly, but also because the presence of vulnerabilities alter the optimal response of interest rates to changes in output and inflation.

8.2 Alternative Monetary Policy Rules

In general, the central bank might follow other monetary policy rules. We consider alternative linear rules that do not explicitly condition on vulnerability or its shocks:

$$i_t = \psi_0 + \psi_y y_t \tag{62}$$

We show that even after picking the coefficients ψ_0, ψ_y in an optimal way, the rule in equation (62) implies quantitatively large welfare losses compared to the optimal monetary policy found in the last section. To find the coefficients ψ_0, ψ_y that minimize welfare losses, we solve

$$\min_{(\psi_0,\psi_y)} L(y_0, s_0) \tag{63}$$

$$s.t.$$

$$dy_t = \frac{1}{\gamma}\left(i_t - r + \gamma\hat{\eta}\xi\left(V_t - s_t - \frac{1}{2}\frac{\hat{\eta}}{\xi\gamma}\right)\right)dt + \xi\left(V_t - s_t\right)dZ_t \tag{64}$$

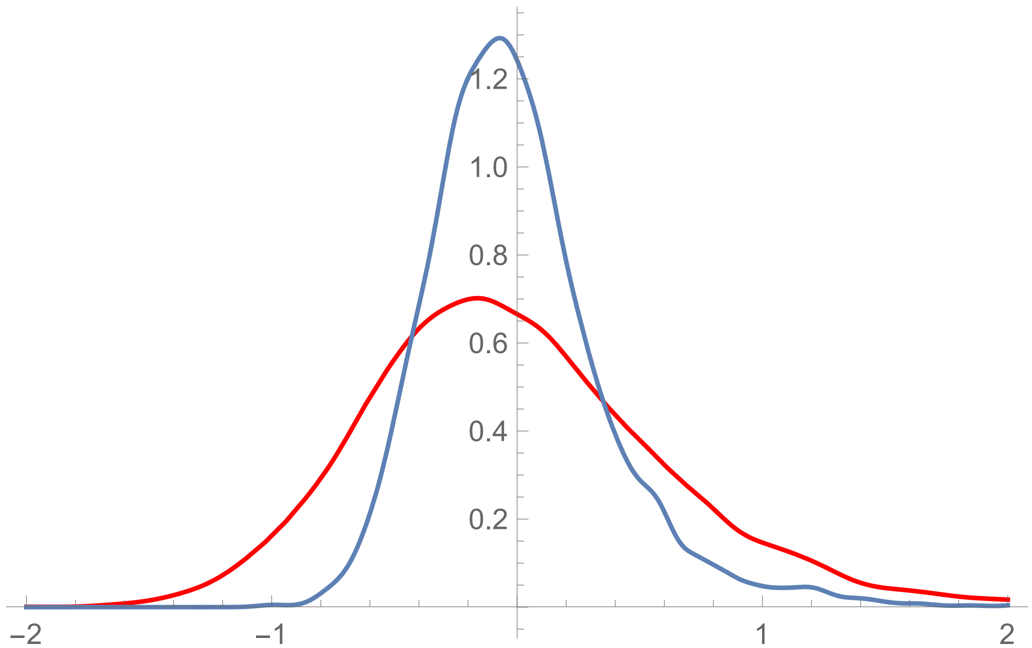
$$i_t = \psi_0 + \psi_y y_t \tag{65}$$

$$V_t = -\mathbb{E}_t\left[dy_t\right]\tau - \alpha\mathbb{V}_t\left[dy_t\right]\sqrt{\tau} \tag{66}$$

$$ds_t = -\kappa\left(s_t - \overline{s}\right) + \sigma_s dZ_t \tag{67}$$

Figure 5 shows the steady-state distribution of the output gap y_t using the optimal policy rule that explicitly takes vulnerability into account (using equation (61)), and the Taylor-type rule that does not condition on vulnerability V_t, given by equation (62) with coefficients found by solving (63)-(67). Intuitively, shocks to vulnerability s contain information about the conditional distribution of the output gap that the policy maker should take into account in setting optimal policy. For a given level of the output gap, a higher vulnerability –a larger VaR of output– calls for higher interest rates. Higher interest rates induce the private sector to save more and consume less, thus shifting the conditional future distribution of y_t upwards by shifting its conditional mean upwards. Given the link between the expected mean and the expected volatility of output induced by the presence of vulnerability, a higher conditional mean induces a lower volatility of y_t. Together, higher mean and lower volatility mean lower vulnerability – lower VaR for output. For the suboptimal Taylor rule that ignores vulnerability, interest rates remain unchanged when, for a given level of y_t, V_t changes. Compared to the optimal rule, when V_t increases but i_t remains unchanged, the conditional mean of output is lower and its conditional volatility is higher. Over time, more frequent visit to states of lower mean and higher volatility create an unconditional distribution that is more negatively skewed. When instead V_t decreases, the optimal rule and the suboptimal Taylor rule produce similar right tails for the unconditional distribution of output. The reason is that lower V_t induces both higher mean and lower volatility of output. Therefore, even though the changes in mean and volatility of y_t are different for the two different rules, the actual differences in outcomes for y_t are small because the lower volatility minimizes all fluctuations.

Figure 5: **Probability Density Functions of the Output under the Optimal Policy Rule and a Standard Taylor Rule**. The figure shows the PDFs using the optimal policy rule and the standard Taylor rule. The standard Taylor rule coefficients are calculated for the economy assuming that the policy maker is ignoring the presence of financial vulnerability.



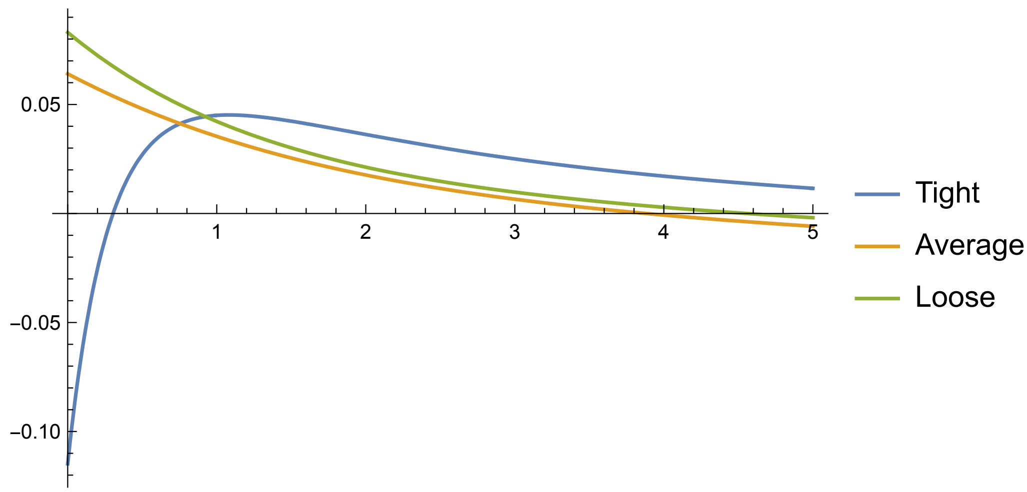
9 The volatility paradox

Under the optimal policy (and under a variety of other linear policies, including the standard Taylor rule considered earlier), the model generates a term structure for GDP-at-Risk V_t that exhibits the volatility paradox (term structure lines cross) analogous to the one in Figure 4. Compared to Figure 4, the model-based Figure 6 shows a crossing at the 1 year horizon, rather than at the 4 year horizon, an indication that despite capturing the qualitative empirical pattern, more elements may be required to achieve a quantitatively more accurate match.

10 Conclusion

The degree to which financial stability considerations should be incorporated in the conduct of monetary policy has long been debated, see Adrian & Liang (2016) for an overview. In this paper, we extend the basic, two equation New Keynesian model to incorporate a notion of financial vulnerability. Shocks to risk premia impact aggregate demand via the Euler equation. The shocks to risk premia are assumed to impact the volatility of output, which is motivated from the empirical observation by Adrian et al. (2019) that financial conditions forecast both the mean and the volatility of output. Importantly, our framework reproduces the stylized fact that

Figure 6: **The volatility paradox in the model.** This plot shows the term-structure of GDP-at-Risk under the optimal monetary policy rule for three different initial values of g_t that represent the lowest 10th percentile, the average, and the highest 90th percentile of its distribution.



the conditional mean and the conditional volatility of output are strongly negatively correlated, giving rise to a sharply negatively skewed unconditional output distribution. Vulnerability thus captures movements in the conditional GDP distribution that correspond to the downside risk of growth.

We further assume that the central bank minimizes the expected discounted sum of squared output gaps and squared inflation, which is standard in the literature. This is therefore a central bank that is subject to a dual mandate, without an independent financial stability objective. Despite that narrow objective function, the optimal flexible inflation targeting rule conditions on the level of vulnerability. Intuitively, all variables that provide information about the conditional distribution of GDP should be taken into account in setting optimal monetary policy. This translates into an augmented Taylor rule, where financial vulnerability—as measured by output gap tail risk as a function of financial variables—is an input into the Taylor rule. Furthermore, the magnitude of the Taylor rule coefficients on output gap and inflation depend on the parameters that determine vulnerability.

The striking result from our setup is that the central bank should **always** condition monetary policy on financial vulnerability. Relative to earlier literature that has made similar arguments (e.g. Curdia & Woodford (2010), Cúrdia & Woodford (2016) and Gambacorta & Signoretti (2014), our modeling approach is deeply rooted in empirical observations which capture macoreconomic shocks of the 2008 crisis very well. Through the negative correlation between conditional mean and conditional variance, our setup captures nonlinearity in macro dynamics in a tractable linear-quadratic setting. The implications of our results for the conduct of monetary policy are in line with the arguments or Adrian & Shin (2010) and Borio & Zhu

(2012).

References

Adrian, T. & Boyarchenko, N. (2012). *Intermediary Leverage Cycles and Financial Stability*. Technical report, Becker Friedman Institute for Research In Economics.

Adrian, T., Boyarchenko, N., & Giannone, D. (2019). Vulnerable growth. *American Economic Review*, 109(4), 1263–89.

Adrian, T., Covitz, D., & Liang, N. (2015). Financial stability monitoring. *Annual Review of Financial Economics*, 7, 357–395.

Adrian, T., Duarte, F., Grinberg, F., & Mancini-Griffoli, T. (2018). *Advancing the Frontiers of Monetary Policy*, chapter Monetary Policy and Financial Conditions: A Cross-Country Study, (pp. 83–105). Number 7. International Monetary Fund.

Adrian, T., Grinberg, F., Liang, N., Malik, S., & Yu, J. (2022). The term structure of growth-at-risk. *American Economic Journal: Macroeconomics*, 14(3), 283–323.

Adrian, T. & Liang, N. (2016). Monetary policy, financial conditions, and financial stability. *International Journal of Central Banking*, forthcoming.

Adrian, T. & Shin, H. (2010). Financial intermediaries and monetary economics. *Handbook of Monetary Economics*, 3, 601–650.

Adrian, T. & Shin, H. S. (2013a). Procyclical Leverage and Value-at-Risk. *The Review of Financial Studies*, 27(2), 373–403.

Adrian, T. & Shin, H. S. (2013b). Procyclical leverage and value-at-risk. *The Review of Financial Studies*, 27(2), 373–403.

Bernanke, B. & Blinder, A. (1992). The federal funds rate and the channels of monetary transmission. *American Economic Review*, (pp. 901–921).

Bernanke, B. & Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *American Economic Review*, (pp. 14–31).

Borio, C. & Zhu, H. (2012). Capital regulation, risk-taking and monetary policy: A missing link in the transmission mechanism? *Journal of Financial Stability*, 8(4), 236–251.

Brunnermeier, M. K., Eisenbach, T. M., & Sannikov, Y. (2013). Macroeconomics with financial frictions: A survey. In *Advances in Economics and Econometrics: Volume 2, Applied Economics: Tenth World Congress*, volume 50 (pp.$\tilde{3}$).: Cambridge University Press.

Brunnermeier, M. K. & Sannikov, Y. (2014a). A macroeconomic model with a financial sector. *American Economic Review*, 104(2), 379–421.

Brunnermeier, M. K. & Sannikov, Y. (2014b). A macroeconomic model with a financial sector. *American Economic Review*, 104(2), 379–421.

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

Curdia, V. & Woodford, M. (2010). Credit spreads and monetary policy. *Journal of Money, Credit and Banking*, 42(1), 3–35.

Cúrdia, V. & Woodford, M. (2016). Credit frictions and optimal monetary policy. *Journal of Monetary Economics*, 84, 30–65.

Dudley, W. C. (2015). Remarks at the fed at a crossroads: Where to go next?, brookings institution, washington, d.c. https://www.newyorkfed.org/newsevents/speeches/2015/dud151015.

Dudley, W. C. (2017). The importance of financial conditions in the conduct of monetary policy, remarks at the university of south florida sarasota-manatee, sarasota, florida. https://www.newyorkfed.org/newsevents/speeches/2017/dud170330.

Estrella, A. & Hardouvelis, G. A. (1991). The term structure as a predictor of real economic activity. *Journal of Finance*, 46(2), 555–576.

Estrella, A. & Mishkin, F. S. (1998). Predicting us recessions: Financial variables as leading indicators. *Review of Economics and Statistics*, 80(1), 45–61.

Galí, J. (2008). *Monetary Policy, Inflation, and the Business Cycle: an Introduction to the New Keynesian Framework and its Applications*. Princeton University Press.

Gambacorta, L. & Signoretti, F. M. (2014). Should monetary policy lean against the wind?: An analysis based on a dsge model with banking. *Journal of Economic Dynamics and Control*, 43, 146–174.

Giannoni, M. & Woodford, M. (2012). *Optimal target criteria for stabilization policy*. Staff Reports 535, Federal Reserve Bank of New York.

Gilchrist, S. & Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4), 1692–1720.

Huang, C.-F. & Pages, H. (1992). Optimal consumption and portfolio policies with an infinite horizon: Existence and convergence. *Ann. Appl. Probab.*, 2(1), 36–64.

Krishnamurthy, A. & Muir, T. (2016). How credit cycles across a financial crisis. *Unpublished paper, Stanford University (June)*.

Leippold, M., Trojani, F., & Vanini, P. (2006). Equilibrium impact of value-at-risk regulation. *Journal of Economic Dynamics and Control*, 30(8), 1277 – 1313.

López-Salido, D., Stein, J. C., & Zakrajšek, E. (2016). Credit-market sentiment and the business cycle. *National Bureau of Economic Research*, 21879.

Rotemberg, J. & Woodford, M. (1997). An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual 1997, Volume 12*, (pp. 297–361).

Rotemberg, J. J. & Woodford, M. (1999). Interest rate rules in an estimated sticky price model. *Monetary Policy Rules*, (pp. 57–126).

Rudebusch, G. & Svensson, L. E. (1999). Policy rules for inflation targeting. *Monetary Policy Rules*, (pp. 203–262).

Stock, J. H. & Watson, M. W. (1998). Diffusion indexes. *National Bureau of Economic Research*, 6702.

Svensson, L. E. (1999). Inflation targeting as a monetary policy rule. *Journal of Monetary Economics*, 43(3), 607–654.

Svensson, L. E. (2002). Inflation targeting: Should it be modeled as an instrument rule or a targeting rule? *European Economic Review*, 46(4), 771–780.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, (pp. 817–838).

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Yellen, J. (2016). The federal reserve's monetary policy toolkit: Past, present, and future, speech at "designing resilient monetary policy frameworks for the future," a symposium sponsored by the federal reserve bank of kansas city, jackson hole, wyoming. `https://www.federalreserve.gov/newsevents/speech/yellen20160826a.htm`.

A Appendix: Details for Good Producing Firms

A.1 Final Good Producers

There is continuum of measure one of final good producers. Firms in the final good sector produce a homogeneous final good, Y_t, using intermediate goods $Y_t(i)$ of different varieties indexed by $i \in [0, 1]$. The production function for each final good producer is

$$Y_t = \left(\int_0^1 Y_t(i)^{\frac{\varepsilon-1}{\varepsilon}} \, di \right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{68}$$

where $\varepsilon > 1$ is the constant elasticity of substitution for differentiated goods (which is equal to the elasticity of substitution across goods for consumers).

The representative final good producer chooses inputs $Y_t(i)$ to maximize real profits

$$Y_t - \frac{1}{P_t} \int_0^1 P_t(i) \, Y_t(i) \, di,$$

where the first term is real revenue and the second term represents real costs. Because final good producers are competitive, they take $P_t(i)$ and P_t as given. Because of constant returns and competition, the size of any one final good firm is indeterminate. However, their input demand is determined by the following cost minimization problem

$$\begin{aligned} & \min_{Y_t(i)} \int_0^1 P_t(i) Y_t(i) \, di, \\ & s.t \\ Y_t \leq & \left(\int_0^1 Y_t(i)^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}. \end{aligned}$$

The cost minimization yields the demand for intermediate good i

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon} Y_t. \tag{69}$$

A.2 Intermediate Good Producers

There is continuum of mass one of monopolistically competitive atomistic firms indexed by $i \in [0, 1]$. Each firm faces a demand curve given by equation (69). Firms use labor $N_t(i)$ to produce output according to the technology

$$Y_t(i) = AN_t(i). \tag{70}$$

Labor is hired in a competitive market with perfect mobility.

Firms set prices according to Calvo staggered pricing. Receiving the signal that allows a firm to change its price is independent of the last time the firm received the signal and across firms. The probability density of receiving a signal after an amount of time h has elapsed since the last signal is $\delta e^{-\delta h}$, where $\delta > 0$ is the Calvo parameter. Firms that are able to adjust the price choose the price optimally. Firms that cannot change their price adjust output to meet demand at the pre-established price. Both types of firms choose inputs to minimize costs, given output demand.

We characterize first the input choice problem conditional on output. We then characterize the optimal price adjustment and output decisions.

Input Demand and Marginal Cost Firm i chooses $N_t(i)$ to minimize total cost, given by

$$\frac{W_t}{P_t} N_t(i)$$

subject to

$$AN_t(i) - Y_t(i) \geq 0 \tag{71}$$

where W_t/P_t is the real wage and s_t is the government labor tax. Let MC_t denote the Lagrange multiplier with respect to the constraint. Note that MC_t is the firm's real marginal cost (the derivative of total cost with respect to $Y_t(i)$). The FOC with respect to $N_t(i)$ is

$$MC_t = \frac{1}{A}\frac{W_t}{P_t} \tag{72}$$

Since the firm takes W_t/P_t as given, real marginal cost is constant across firms. Equation (71) with equality gives labor demand

$$N_t(i) = \frac{Y_t(i)}{A} \tag{73}$$

Optimal Price Setting Intermediate good producers maximize real expected profits discounted using the household's stochastic discount factor, subject to their production technology (70), the demand curve (69) and the Calvo constraint on price adjustment. A firm that is allowed to change its price at time t picks $P_t(i)$ to maximize

$$\mathbb{E}_t \int_t^\infty SDF_{s,t}\delta e^{-\delta(s-t)} \left(\frac{P_t(i)}{P_s} Y_{s|t}(i) - MC_s Y_{s|t}(i) \right) ds$$

subject to

$$Y_{s|t}(i) = \left(\frac{P_t(i)}{P_s} \right)^{-\varepsilon} Y_s,$$

where $SDF_{s,t}$ is the stochastic discount factor used by the household at time t to discount time-s payoffs and $Y_{s|t}(i)$ is the demand of good i at time s conditional on firm i having changed prices for the last time at time t. In the optimal price setting decision, the firm takes the paths $\left\{SDF_{s,t}, P_s, Y_{s|t}(i), Y_s, MC_s\right\}_{s\geq t}$ as given. The maximization yields an optimal price

$$P_t^*(i) = \frac{\varepsilon}{\varepsilon - 1}\mathbb{E}_t \int_t^\infty \Upsilon_{s,t} MC_s ds \tag{74}$$

where $\varepsilon/(\varepsilon - 1)$ is the gross markup and

$$\Upsilon_{s,t} \equiv \frac{Q_s e^{-\delta s} P_s^\varepsilon Y_s}{\mathbb{E}_t \int_t^\infty Q_s e^{-\delta s} P_s^{\varepsilon-1} Y_s ds}.$$

Equation (74) show that the optimal price is a weighted average of real marginal costs times the markup. Because the optimal price $P_t^*(i)$ depends only on aggregate variables, all firms that are allowed to change the price pick the same optimal price, so we drop the index i from

$P_t^*(i)$ and simply write P_t^*. Firms who are not allowed to reset their prices at time t solve the same problem but instead of picking $P_t(i)$ they keep it constant at its pre-existing level.

By the Calvo price setting assumption, the aggregate price P_t is given by

$$\begin{aligned} P_t^{1-\varepsilon} &= \int_0^1 P_t(i)^{1-\varepsilon} di \\ &= \int_{-\infty}^{t} \delta e^{-\delta(t-s)} P_s^*(i)^{1-\varepsilon} ds \end{aligned} \tag{75}$$

Intuitively, the last expression says that a mass $\delta e^{-\delta(t-s)}$ of firms changed their price to $P_s^*(i)$ at time $t-s$. Differentiating both sides of (75) with respect to time and defining inflation π_t by $\pi_t \equiv (1/dt)(dP_t/P_t)$ we get

$$\pi_t = \frac{\delta}{1-\varepsilon}\left(\left(\frac{P_t^*}{P_t}\right)^{1-\varepsilon} - 1\right). \tag{76}$$

The real profits for the producer of intermediate good i are

$$D_{t,goods}(i) = \frac{P_t(i) Y_t(i)}{P_t} - MC_t Y_t(i). \tag{77}$$

Phillips Curve Differentiating (76) with respect to time gives the non-linear NKPC

$$d\pi_t = \delta\left(\frac{x_{2,t}}{x_{1,t}}\right)^{1-\varepsilon}\left(\frac{\delta}{x_{2,t}}\frac{MC_t}{MC} - \frac{\delta}{x_{1,t}} - \pi_t\right) dt \tag{78}$$

where $x_{1,t}$ and $x_{2,t}$ have dynamics given by

$$\begin{aligned} \frac{dx_{1,t}}{x_{1,t}} &= \left(1 + \frac{1}{x_{1,t}}\right)\delta dt + (1-\varepsilon)\pi_t dt + Q_t Y_t d\left(\frac{1}{Q_t Y_t}\right) \\ \frac{dx_{2,t}}{x_{2,t}} &= \left(1 + \frac{1}{x_{2,t}}\frac{MC_t}{MC}\right)\delta dt - \varepsilon\pi_t dt + Q_t Y_t d\left(\frac{1}{Q_t Y_t}\right) \end{aligned}$$

and satisfy $P_t^*/P_t = x_{2,t}/x_{1,t}$.

B Appendix: Definition of Equilibrium

An equilibrium is a collection of paths $\{N_t, W_t, P_t, C_t, Q_t, F_t, R_t, \omega_t, \mu_t, \sigma_t, Y_t, Y_t(i), P_t(i), P_t, f_t, \theta_t, X_t, \gamma_t, S_t$ such that, for all realizations of the exogenous processes ζ_t and g_t, the following conditions hold:

1. Households optimize

$$
\begin{array}{rcl}
\text{(labor supply)} & : & N_t = \left(\left(1-s_t\right) \frac{W_t}{P_t} C_t^{-\gamma} \right)^{\frac{1}{\xi}} \\
\text{(demand for final goods)} & : & C_t = \left(e^{\beta t} Q_t\right)^{-\frac{1}{\gamma}} \\
\text{(demand for risky asset)} & : & \frac{dF_t}{F_t} = \left(R_t + \omega_t \mu_{banks,t} - \frac{1}{F_t} \left(C_t - \left(1-s_t\right) \frac{W_t}{P_t} N_t + T_t \right) \right) dt + \omega_t \sigma_{banks} \\
\text{(demand for riskless bond)} & : & \left(1-\omega_t\right) F_t \\
\text{(transversality condition)} & : & \lim_{s\to\infty} \mathbb{E}_t \left[Q_s F_s \right] = 0
\end{array}
$$

2. Good producers optimize

   (a) Final good producers

$$
\begin{array}{rcl}
\text{(demand for intermediate good } i\text{)} & : & Y_t\left(i\right) = \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon} Y_t \\
\text{(supply of final goods)} & : & Y_t = \left(\int_0^1 Y_t\left(i\right)^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}
\end{array}
$$

   (b) Intermediate goods producers

$$
\begin{array}{rcl}
\text{(labor demand)} & : & N_t\left(i\right) = \frac{Y_t(i)}{A} \\
\text{(supply of intermediate goods)} & : & Y_t\left(i\right) = A N_t\left(i\right) \\
\text{(price setting)} & : & d\pi_t = \delta \left(\frac{x_{2,t}}{x_{1,t}} \right)^{1-\varepsilon} \left(\frac{\delta\varepsilon}{\varepsilon-1} \frac{1}{A} \frac{W_t}{P_t} \frac{1}{x_{2,t}} - \frac{\delta}{x_{1,t}} - \pi_t \right) dt \\
& & \frac{dx_{1,t}}{x_{1,t}} = \left(1 + \frac{1}{x_{1,t}}\right) \delta dt + \left(1-\varepsilon\right) \pi_t dt + Q_t Y_t d \left(\frac{1}{Q_t Y_t} \right) \\
& & \frac{dx_{2,t}}{x_{2,t}} = \left(1 + \frac{\varepsilon}{\varepsilon-1} \frac{1}{A} \frac{W_t}{P_t} \frac{1}{x_{2,t}}\right) \delta dt - \varepsilon \pi_t dt + Q_t Y_t d \left(\frac{1}{Q_t Y_t} \right)
\end{array}
$$

3. Banks optimize

$$
\begin{array}{rcl}
\text{(dividends)} & : & f_t = u\left(\gamma_t, \eta_t - g_t\right) \beta \\
\text{(portfolio)} & : & \theta_t = \frac{1}{\gamma_t} \frac{\eta_t - g_t}{\sigma_t} \\
\text{(wealth)} & : & \frac{dX_t}{X_t} = \left(R_t - f_t + \theta_t \mu_t\right) dt + \theta_t \sigma_t dB_t \\
\text{(VaR constraint)} & : & \gamma_t = \max\left\{1, \hat{\gamma}_t\right\} \text{ with } \hat{\gamma}_t \text{ such that: } VaR^{bank}_{\tau,\alpha} \left(t, \frac{\eta_t - g_t}{\hat{\gamma}_t \sigma_t}, u\left(\hat{\gamma}_t, \eta_t - g_t\right) \beta \right) = a_V X
\end{array}
$$

4. Markets clear

$$
\begin{array}{rcl}
\text{(intermediate goods)} & : & \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon} Y_t = A_t N_t\left(i\right) \\
\text{(final goods)} & : & C_t = Y_t \\
\text{(labor)} & : & \int_0^1 N_t\left(i\right) di = N_t \\
\text{(risky asset)} & : & \frac{F_t \omega_{banks,t}}{S_{banks,t}} + \frac{X_t \theta_{banks,t}}{S_{banks,t}} = 1 \\
\text{(riskless bond)} & : & \frac{X_t \theta_{goods,t}}{S_{goods,t}} = 1
\end{array}
$$

5. The central bank sets nominal interest rates $i_t \equiv R_t + \pi_t$.

C Solution to the bank problem

The bank problem under the bank measure is

$$
\begin{aligned}
V(X_t) &= \max_{\{\theta_s, f_s\}_{s\geq t}} \mathbb{E}_t^{bank}\left[\int_t^{\infty} e^{-\beta(s-t)} \log(f_s X_s)\, ds\right] && (79)\\
& s.t. \\
\frac{dX_t}{X_t} &= \left(R_t - f_t + \theta_t\left(\mu_t - \sigma_t g_t\right)\right) dt + \theta_t \sigma_t dB_t^{bank}, && (80)\\
VaR_{\tau,\alpha}^{bank}(t, \theta_t, f_t) &\leq a_V X_t, && (81)\\
dg_t &= -\left(\kappa_g + \sigma_g\right) g_t dt + \sigma_g dB_t^{bank}, && (82)
\end{aligned}
$$

with X_t and g_t given. Note that $E_s^{bank}[\cdot]$ and $VaR_{\tau,\alpha}^{bank}$ denote, respectively, conditional expectations and the value-at-risk under the bank measure.

Let

$$U(t, \theta, f, g) \equiv R_t - f + \theta\left(\mu_t - \sigma_t g\right) - \frac{1}{2}\left(\theta \sigma_t\right)^2$$

be the drift of $d \log X_t$. Then, the dynamic budget constraint of the bank in equation (80) has a strong solution given by

$$X_t = X_0 \exp\left\{\int_0^t U(s, \theta_s, f_s, g_s)\, ds + \int_0^t \theta_s \sigma_s dB_s^{bank}\right\}$$

with X_0 given. Projected wealth loss between t and $t+\tau$ when keeping $(R_t, \mu_t, \sigma_t, g_t, f_t, \theta_t)$ constant at their time-t values during the interval $t \in [t, t+\tau]$ is

$$X_t - X_{t+\tau} = X_t\left[1 - \exp\left\{U(t, \theta_t, f_t, g_t)\tau + \theta_t \sigma_t\left(B_{t+\tau}^{bank} - B_t^{bank}\right)\right\}\right]$$

Value-at-risk at level α and horizon τ is defined as the α-percentile of the projected wealth loss, $X_t - X_{t+\tau}$, conditional on time-t information, and is given by

$$VaR_{\tau,\alpha}^{bank}(t, \theta_t, f_t, X_t, g_t) = X_t\left[1 - \exp\left\{U(t, \theta_t, f_t, g_t)\tau + \mathcal{N}^{-1}(\alpha)\left|\theta_t \sigma_t\right|\sqrt{\tau}\right\}\right]$$

where $\mathcal{N}^{-1}$ is the inverse cumulative distribution function of a standard normal distribution. Define

$$g_V(t, \theta, f, g) \equiv -U(t, \theta, f, g)\tau - \mathcal{N}^{-1}(\alpha)\left|\theta \sigma_t\right|\sqrt{\tau}$$

Then

$$VaR^{bank}_{\tau,\alpha}\left(t,\theta,f,x,g\right) \leq xa_V$$

if, and only if,

$$g_V\left(t,\theta,f,g\right) \leq \log\frac{1}{1-a_V}.$$

Since $\alpha \in (0,1/2]$, we have that $\mathcal{N}^{-1}\left(\alpha\right) \leq 0$ and that $g_V\left(t,\theta,f,g\right)$ is convex in (θ,f).

Log-utility allows us to transform the bank's optimization problem into a non-stochastic problem. Indeed, the objective function can be written as

$$\begin{aligned}
\int_0^\infty e^{-\beta t}\log\left(f_tX_t\right)dt &= \int_0^\infty e^{-\beta t}\log\left(X_t\right)dt + \int_0^\infty e^{-\beta t}\log\left(f_t\right)dt \\
&= \int_0^\infty e^{-\beta t}\log\left(X_0\right)dt + \int_0^\infty e^{-\beta t}\left\{\int_0^t U\left(s,\theta_s,f_s,g_s\right)ds + \int_0^t \theta_s\sigma_s dB^{bank}_s\right\}dt + \int_0 \\
&= \log\left(X_0\right)\int_0^\infty e^{-\beta t}dt + \int_0^\infty e^{-\beta t}\log\left(f_t\right)dt + \int_0^\infty\int_0^t e^{-\beta t}U\left(s,\theta_s,f_s,g_s\right)dsdt + \int_0 \\
&= \log\left(X_0\right)\int_0^\infty e^{-\beta t}dt + \int_0^\infty e^{-\beta t}\log\left(f_t\right)dt + \int_0^\infty\int_s^\infty e^{-\beta t}U\left(s,\theta_s,f_s,g_s\right)dtds + \int_0 \\
&= \log\left(X_0\right)\int_0^\infty e^{-\beta t}dt + \int_0^\infty e^{-\beta t}\log\left(f_t\right)dt + \int_0^\infty U\left(s,\theta_s,f_s,g_s\right)\left[\int_s^\infty e^{-\beta t}dt\right]ds +
\end{aligned}$$

where the change in the order of integration follows from Fubini's theorem. We assume all the usual regularity conditions. In particular, we assume that

$$\int_0^\infty \left|\sigma_t^{-1}\mu_t\right|^2 dt < \infty \tag{84}$$

Under the regularity condition in equation (84), the stochastic part of the bank's objective function in (83) is a martingale and not just a local martingale, so

$$E^{bank}_0\int_0^\infty\int_0^t e^{-\beta t}\theta_s\sigma_s dB^{bank}_s dt = 0$$

Therefore, taking expectations in (83) gives

$$\begin{aligned}
E^{bank}_0\int_0^\infty e^{-\beta t}\log\left(f_tX_t\right)dt &= \log\left(X_0\right)\int_0^\infty e^{-\beta t}dt + E^{bank}_0\int_0^\infty e^{-\beta t}\log\left(f_t\right)dt + E^{bank}_0\int_0^\infty U\left(s,\theta_s,f_s,g_s\right) \\
&= \frac{\log\left(X_0\right)}{\beta} + E^{bank}_0\int_0^\infty e^{-\beta t}\log\left(f_t\right)dt + \frac{1}{\beta}E^{bank}_0\int_0^\infty U\left(s,\theta_s,f_s,g_s\right)e^{-\beta s}ds \\
&= \frac{\log\left(X_0\right)}{\beta} + E^{bank}_0\int_0^\infty e^{-\beta t}\left(\log\left(f_t\right) + \frac{1}{\beta}U\left(t,\theta_t,f_t,g_t\right)\right)dt
\end{aligned}$$

Thus, to maximize

$$E_0^{bank} \int_0^{\infty} e^{-\beta t} \log\left(f_t X_t\right) dt$$

over the constrained set, it suffices to maximize

$$h\left(t, \theta_t, f_t, g_t\right) \equiv \log\left(f_t\right) + \frac{1}{\beta} U\left(t, \theta_t, f_t, g_t\right)$$

pathwise over the constrained set. For a fixed path, at time t the bank then solves

$$\begin{aligned} & \max_{\theta_t, f_t} h\left(t, \theta_t, f_t, g_t\right) \\ & s.t. \\ g_V\left(t, \theta_t, f_t, g_t\right) \quad \leq \quad & \log \frac{1}{1 - a_V} \end{aligned} \tag{85}$$

The function $h\left(t, \theta_t, f_t, g_t\right)$ is concave in $\left(\theta_t, f_t\right)$ and maximized over $\left(\theta_t, f_t\right)$ by

$$\begin{aligned} f_t &= f_{M,t} \\ \theta_t &= \theta_{M,t} \end{aligned}$$

when the VaR constraint is not binding, where we derive $f_{M,t}, \theta_{M,t}$ using the FOC

$$\begin{aligned} [f_t] \quad &: \quad 0 = \frac{\partial}{\partial f_t} h\left(t, \theta_t, f_{M,t}, g_t\right) \\ &: \quad 0 = \frac{1}{f_{M,t}} - \frac{1}{\beta} \\ &: \quad f_{M,t} = \beta \\ [\theta_t] \quad &: \quad 0 = \nabla_\theta h\left(t, \theta_{M,t}, f_t, g_t\right) \\ &: \quad 0 = \frac{1}{\beta}\left(\mu_t - \sigma_t g_t - \sigma_t^2 \theta_{M,t}\right) \\ &: \quad \theta_{M,t} = \left(\sigma_t\right)^{-1}\left(\sigma_t^{-1} \mu_t - g_t\right) \end{aligned}$$

Using the definition of the market price of risk η_t, we can also write

$$\theta_{M,t} = \sigma_t^{-1}\left(\eta_t - g_t\right)$$

If $\left(\theta_{M,t}, f_{M,t}\right)$ satisfy

$$g_V\left(t, \theta_{M,t}, f_{M,t}, g_t\right) \leq \log \frac{1}{1 - a_V}$$

then $\left(\theta_t^*, f_t^*\right) = \left(\theta_{M,t}, f_{M,t}\right)$ is the solution to the problem with the VaR constraint. Otherwise, because the constraint set is compact and convex, and the objective is continuous, there will be a unique solution $(\theta_t^*, f_t^*) \neq (\theta_{M,t}, f_{M,t})$. Moreover, in this case, $(\theta_t^*, f_t^*) \neq (\theta_{M,t}, f_{M,t})$ must be such that the VaR constraint holds with equality.

If $|\theta\sigma_t| = 0$, then (85) gives

$$f_t = R_t - \frac{1}{\tau}\log(1 - a_V)$$

If $|\theta\sigma_t| \neq 0$, $g_V(t, \theta_t, f_t, g_t)$ is differentiable in θ_t and thus we can solve (85) using the Karush–Kuhn–Tucker conditions. Set up the Lagrangian

$$\mathcal{L} = h(t, \theta_t, f_t, g_t) - \lambda\left(g_V(t, \theta_t, f_t, g_t) - \log\frac{1}{1 - a_V}\right)$$

Direct computation shows that $\nabla g_V(t, \theta_t, f_t, g_t) \neq 0$.[13] Thus, $\lambda \neq 0$ and the FOC is

$$\nabla h(t, \theta_t, f_t, g_t) = \lambda \nabla g_V(t, \theta_t, f_t, g_t) \tag{86}$$

We compute

$$\begin{aligned}
\nabla_{\theta_t} h &= \frac{1}{\beta}\left(\mu_t - \sigma_t g_t - \sigma_t^2\theta_t\right) \\
\nabla_{\theta_t} g_V &= -\left(\mu_t - \sigma_t g_t - \sigma_t^2\theta_t\right)\tau - \mathcal{N}^{-1}(\alpha)\frac{\sigma_t^2\theta_t}{|\theta_t\sigma_t|}\sqrt{\tau} \\
\nabla_{f_t} h &= \frac{1}{f_t} - \frac{1}{\beta} \\
\nabla_{f_t} g_V &= \tau
\end{aligned}$$

so that the FOC become

$$\begin{aligned}
\nabla_{\theta_t} h(t, \theta_t, f_t, g_t) &= \lambda \nabla_{\theta_t} g_V(t, \theta_t, f_t, g_t) \\
\frac{1}{\beta}\left(\mu_t - \sigma_t g_t - \sigma_t^2\theta_t\right) &= \lambda\left(-\left(\mu_t - \sigma_t g_t - \sigma_t^2\theta_t\right)\tau - \mathcal{N}^{-1}(\alpha)\frac{\sigma_t^2\theta_t}{|\theta_t\sigma_t|}\sqrt{\tau}\right) \\
(1 + \beta\tau\lambda)(\mu_t - \sigma_t g_t) &= \left(1 + \left(\tau - \frac{\sqrt{\tau}\mathcal{N}^{-1}(\alpha)}{|\theta_t\sigma_t|}\right)\beta\lambda\right)\sigma_t^2\theta_t
\end{aligned} \tag{87}$$

[13] The gradient is taken with respect to θ_t and f_t. We also use the notation ∇_x to denote $\partial/\partial x$.

and

$$
\begin{aligned}
\nabla_{f_t} h\left(t, \theta_t, f_t, g_t\right) &= \lambda \nabla_{f_t} g_V\left(t, \theta_t, f_t, g_t\right) \\
\frac{1}{f_t} - \frac{1}{\beta} &= \lambda \tau \\
f_t &= \frac{\beta}{\beta \lambda \tau + 1}
\end{aligned}
$$

Since $\theta_{M,t} = \sigma_t^{-2}\left(\mu_t - \sigma_t g_t\right)$, equation (87) shows that θ_t is parallel to $\theta_{M,t}$. Then, to solve the maximization problem in (85), all we need is to find λ_1, λ_2 that solve

$$
\begin{aligned}
& \max_{\lambda_1, \lambda_2} h\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right) \\
& s.t. \\
g_V\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right) &\leq \log \frac{1}{1 - a_V}
\end{aligned}
$$

Again, it can be checked that the constraint holds with equality. The Lagrangian is

$$
\mathcal{L} = h\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right) - \psi\left(g_V\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right) - \log \frac{1}{1 - a_V}\right)
$$

The FOC are

$$
\begin{aligned}
\frac{\partial}{\partial \lambda_1} h\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right) &= \psi \frac{\partial}{\partial \lambda_1} g_V\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right) \\
\frac{\partial}{\partial \lambda_2} h\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right) &= \psi \frac{\partial}{\partial \lambda_2} g_V\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right)
\end{aligned}
$$

Computing the derivatives gives

$$
\begin{aligned}
\frac{\partial}{\partial \lambda_1} h\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right) &= \psi \frac{\partial}{\partial \lambda_1} g_V\left(t, \theta, f, g\right) \\
\frac{1}{\beta}\left(\theta_{M,t}\left(\mu_t - \sigma_t g_t\right) - \lambda_1\left(\theta_{M,t} \sigma_t\right)^2\right) &= \psi\left(-\left(\theta_{M,t}\left(\mu_t - \sigma_t g_t\right) - \lambda_1\left(\theta_{M,t} \sigma_t\right)^2\right) \tau - \mathcal{N}^{-1}(\alpha)\left|\theta_{M,t} \sigma_t\right| \sqrt{\tau}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial}{\partial \lambda_2} h\left(t, \lambda_1 \theta_{M,t}, \lambda_2 f_{M,t}, g_t\right) &= \psi \frac{\partial}{\partial \lambda_2} g_V\left(t, \theta, f, g\right) \\
\frac{1}{\lambda_2} - \frac{f_{M,t}}{\beta} &= \psi f_{M,t} \tau
\end{aligned}
$$

Eliminating ψ and using

$$\begin{aligned}
\theta_{M,t}\sigma_t &= \sigma_t^{-1}\left(\mu_t - \sigma_t g_t\right) \\
\left|\theta_{M,t}\sigma_t\right| &= \left|\sigma_t^{-1}\left(\mu_t - \sigma_t g_t\right)\right| \\
\theta_{M,t}\left(\mu_t - \sigma_t g_t\right) &= \sigma_t^{-2}\left(\mu_t - \sigma_t g_t\right)^2 = \left(\theta_{M,t}\sigma_t\right)^2
\end{aligned}$$

we get

$$\lambda_2 = u\left(t, \lambda_1\right) f_{M,t}$$

where the function u is defined by

$$u\left(t, z\right) \equiv 1 + \frac{\sqrt{\tau}\left|\theta_{M,t}\sigma_t\right|}{\mathcal{N}^{-1}\left(\alpha\right)}\left(1 - z\right) \tag{88}$$

and λ_1 makes the VaR constraint hold with equality. For the VaR constraint to hold with equality, $\lambda_1 > 0$ must satisfy

$$g_V\left(t, \lambda_1\theta_{M,t}, u\left(t, \lambda_1\right) f_{M,t}, g_t\right) = -\log\left(1 - a_V\right) \tag{89}$$

If there are no positive roots, it means we are in the case

$$\begin{aligned}
\theta_t &= \mathbf{0} \\
f_t &= R_t - \frac{1}{\tau}\log\left(1 - a_V\right)
\end{aligned}$$

considered before.

Equation (89) has the form

$$0 = c + b_1\lambda_1 + b_2\left|\lambda_1\right| + a\lambda_1^2$$

with

$$\begin{aligned}
a &= \frac{1}{2}\tau\left(\eta_t - g_t\right)^2 \\
b_1 &= -\left(\frac{\sqrt{\tau}\left|\eta_t - g_t\right|}{\mathcal{N}^{-1}\left(\alpha\right)}\tau\beta + \tau\left(\eta_t - g_t\right)^2\right) \\
b_2 &= -\sqrt{\tau}\left|\eta_t - g_t\right|\mathcal{N}^{-1}\left(\alpha\right) \\
c &= \log\left(1 - a_V\right) + \tau\left(\beta - R_t\right) + \tau\beta\frac{\sqrt{\tau}\left|\eta_t - g_t\right|}{\mathcal{N}^{-1}\left(\alpha\right)}
\end{aligned}$$

If $\eta_t \neq g_t$, solutions to (89) have the form

$$\lambda_1 = -\frac{b_1 \pm b_2}{2a} \pm \frac{\sqrt{\left(b_1 + b_2\right)^2 - 4ac}}{2a}$$

Since $a, b_2 > 0$, the only solutions that can be positive are

$$\lambda_1 = -\frac{b_1 + b_2}{2a} \pm \frac{\sqrt{\left(b_1 + b_2\right)^2 - 4ac}}{2a}$$

However, when $\lambda_1 > 0$, the smallest of these two solutions is not optimal since $\nabla_{\theta_t} h\left(t, \lambda_1 \theta_{M,t}, u\left(t, \lambda_1\right) f_{M,t}, g_t\right.$
0. Therefore, the only solution that can be optimal and positive is

$$\varphi_t \equiv \lambda_1 = -\frac{b_1 + b_2}{2a} + \frac{\sqrt{\left(b_1 + b_2\right)^2 - 4ac}}{2a}$$

which we label φ_t to distinguish it from the other potential solutions. To summarize, we have the following cases:

- If $\eta_t = g_t$ or $\varphi_t \leq 0$, then $\theta_t = \mathbf{0}$ and $f_t = R_t - \frac{1}{\tau} \log\left(1 - a_V\right)$.
- If $\eta \neq g$, we have two cases
  - If $\varphi_t \in (0, 1]$, the VaR constraint holds with equality and we have $\theta_t = \varphi_t \theta_{M,t}$ and $f_t = u\left(t, \varphi_t\right) f_{M,t}$
  - If $\varphi_t > 1$, the VaR constraint does not bind and $\theta_t = \theta_{M,t}$ and $f_t = f_{M,t}$.

Putting everything together, the optimal portfolio when is then characterized by

$$\begin{aligned} \theta_t &= \min\left\{1, \max\left\{0, \varphi_t\right\}\right\} \theta_{M,t} && (90) \\ f_t &= u\left(t, \min\left\{1, \varphi_t\right\}\right) f_{M,t} 1_{\{\varphi_t > 0\}} + \left(R_t - \frac{1}{\tau} \log\left(1 - a_V\right)\right) 1_{\{\varphi_t \leq 0\}} \\ \varphi_t \text{ is the largest root of:} \quad & g_V\left(t, \varphi_t \theta_{M,t}, u\left(t, \varphi_t\right) f_{M,t}, g_t\right) = -\log\left(1 - a_V\right) && (91) \end{aligned}$$

We now study when $\varphi_t > 0$ and $\varphi_t > 1$. If any one of the four conditions

1. $0 < c$ and $b_1 + b_2 < 0$ and $\left(b_1 + b_2\right)^2 - 4ac > 0$
2. $b_1 > 0$ and $c < 0$
3. $b_1 + b_2 < 0$ and $c < 0$
4. $b_1 < 0$ and $-b_2 < b_1$ and $c < 0$

is satisfied, then $\varphi_t > 0$. If any of the two conditions

1. $a + b_1 + b_2 + c < 0$ and $((b_1 + b_2 > 0$ and $b_1 < 0)$ or $(b_1 + 2b_2 > 0$ and $b_1 + b_2 < 0)$ or $(c > 0$ and $b_1 + 2b_2 < 0)$ or $b_1 > 0$ or $(a + b_1 + b_2 > 0$ and $b_1 + 2b_2 \leq 0))$

2. $b_1 + 2b_2 < 0$ and $c < 0$ and $a + b_1 + b_2 \leq 0$

is satisfied, then $\varphi_t > 1$.

Let $\hat{\gamma} = 1/\varphi_t$. We define the bank's effective risk aversion by

$$\begin{aligned}
\gamma_t &\equiv \frac{1}{\min\left\{1, \max\left\{0, \frac{1}{\hat{\gamma}}\right\}\right\}} \in [1, \infty) \\
&= \begin{cases} \infty & , \text{ if } \quad \hat{\gamma} < 0 \\ 1 & , \text{ if } \; 0 \leq \hat{\gamma} < 1 \\ \hat{\gamma} & , \text{ if } \quad \hat{\gamma} \geq 1 \end{cases}
\end{aligned}$$

Finally, we note that λ is the VaR Lagrange multiplier for the deterministic path-by-path problem, so the Lagrange multiplier for the original problem under the bank's probability measure is

$$\begin{aligned}
\lambda_{VaR,t}^{bank} &= \lambda e^{-\beta t} \\
&= \frac{1}{\tau}\left(\frac{1}{f_t} - \frac{1}{\beta}\right) e^{-\beta t}
\end{aligned}$$

and under the physical measure is

$$\begin{aligned}
\lambda_{VaR,t} &= \lambda e^{-\beta t} e^{\zeta_t} \\
&= \frac{1}{\tau}\left(\frac{1}{f_t} - \frac{1}{\beta}\right) e^{-\beta t} e^{\zeta_t}
\end{aligned} \tag{92}$$

Note that since

$$f_t \leq \beta$$

we have

$$\lambda_{VaR,t} \geq 0$$

C.1 State price density of the bank under complete markets

The solution to the bank problem derived above did not require complete markets. We now derive the bank's state price density (SPD) assuming that markets are complete. The bank problem under the physical measure is

$$
\begin{aligned}
V\left(X_0\right) &= \max_{\{\theta_t, f_t\}_{t\geq 0}} E_0\left[\int_0^{\infty} e^{-\beta t} e^{\zeta_t} \log\left(f_t X_t\right) dt\right] \\
& s.t. \\
\frac{dX_t}{X_t} &= \left(R_t - f_t + \theta_t \mu_t\right) dt + \theta_t \sigma_t dB_t \qquad (93) \\
g_V\left(t, \theta_t, f_t\right) &= -\left(R_t - f_t + \theta_t \mu_t - \frac{1}{2}\left(\theta_t \sigma_t\right)^2\right)\tau - \mathcal{N}^{-1}(\alpha)\left|\theta_t \sigma_t\right|\sqrt{\tau} \leq \log\frac{1}{1-a_V}
\end{aligned}
$$

Complete markets implies that the dynamic budget constraint (93) is equivalent to the static one[14]

$$
X_0 = E_0\left[\int_0^{\infty} Q_t f_t X_t dt\right] \qquad (94)
$$

where the banks take Q_t as given. The Lagrangian is :

$$
\mathcal{L} = E_0\left[\int_0^{\infty} e^{-\beta t} e^{\zeta_t} \log\left(f_t X_t\right) dt\right] + \lambda_{bc}\left(X_0 - E_0\left[\int_0^{\infty} Q_t f_t X_t dt\right]\right) - \int_0^{\infty} \lambda_{VaR,t}\left(g_V\left(t, \theta_t, f_t\right) - \log\frac{1}{1-a_V}\right.
$$

where $\lambda_{bc} > 0$ is a number but $\lambda_{VaR,t} > 0$ is a function of time since we have one VaR constraint for each t. The FOC for an interior solution are

$$
\begin{aligned}
\left[f_t\right] &: \; 0 = \frac{e^{-\beta t} e^{\zeta_t}}{f_t X_t} - \lambda_{bc} Q_t + \lambda_{VaR,t}\frac{\tau}{X_t} \qquad (95) \\
\left[\theta_t\right] &: \; 0 = \frac{e^{-\beta t} e^{\zeta_t}}{\beta}\nabla_\theta U\left(t, \theta_t, f_t\right) + \lambda_{VaR,t}\nabla_\theta g_V\left(t, \theta_t, f_t\right)
\end{aligned}
$$

Re-arranging (95) gives

$$
\begin{aligned}
Q_t &= \frac{e^{-\beta t} e^{\zeta_t}}{\lambda_{bc}}\frac{1}{f_t X_t} + \frac{\lambda_{VaR,t}\tau}{\lambda_{bc}}\frac{1}{X_t} \\
&= \frac{1}{\lambda_{bc} X_t}\left(e^{-\beta t} e^{\zeta_t}\frac{1}{f_t} + \lambda_{VaR,t}\tau\right)
\end{aligned}
$$

Using $\lambda_{VaR,t}$ from equation (92)

$$
\lambda_{VaR,t} = \frac{e^{-\beta t} e^{\zeta_{b,t}}}{\tau}\left(\frac{1}{f_t} - \frac{1}{\beta}\right)
$$

gives

$$
Q_t = \frac{e^{-\beta t} e^{\zeta_t}}{\lambda_{bc} X_t}\left(\frac{2}{f_t} - \frac{1}{\beta}\right)
$$

[14]See Huang & Pages (1992).

The multiplier λ_{bc} can be found from noting that we must have $Q_0 = 1$, which gives

$$\lambda_{bc} = \frac{e^{\zeta_0}}{X_0}\left(\frac{2}{f_0} - \frac{1}{\beta}\right)$$