AMBIGUITY AND INFORMATION PROCESSING IN A MODEL OF INTERMEDIARY ASSET PRICING

LEYLA JIANYU HAN, KENNETH KASA, AND YULEI LUO

Abstract. This paper incorporates ambiguity and information processing constraints into a model of intermediary asset pricing. Financial intermediaries (specialists) are assumed to possess greater information processing capacity. Households purchase this capacity, and then delegate their investment decisions to specialists. The delegation contract is constrained by two frictions: (1) As in He and Krishnamurthy (2012), an incentive constraint arises from a moral hazard problem, which takes the form of a minimum capital requirement, and (2) Because households can invest for themselves at any time, continued delegation is subject to a participation constraint that depends on the underlying heterogeneity in channel capacity. At the same time, both households and specialists have a preference for robustness, reflecting ambiguity about risky asset returns. Ambiguity takes the form of endogenously determined pessimistic drift distortions (Hansen and Sargent (2008)). When volatility increases, so does ambiguity, since it becomes more difficult to discriminate among models. Importantly, these endogenous drift distortions produce heterogeneous beliefs. In our model, ambiguity is scaled by the inverse of time preference, and we assume specialists are more patient. Hence, given their longer investment horizons, specialists have a stronger preference for robustness. As a result, when volatility is high specialists become relatively pessimistic, and this tightens the capital constraint and accelerates the onset of a financial crisis.

Keywords: Ambiguity, Information Processing, Asset Pricing, Financial Crisis.

JEL Classification Numbers: D81, G01, G12

1. Introduction

In a pair of influential papers, He and Krishnamurthy (2012, 2013) [henceforth HK12, HK13] argue that for many assets it is misleading to characterize prices using household Euler equations. This is because many assets are not held by households. They are held by leveraged financial intermediaries. Although these intermediaries may be investing on behalf of households, the contractual relationships between them are plagued by a variety of frictions. In HK12, asymmetric information produces a moral hazard problem that leads to a capital constraint, requiring the intermediary to maintain a minimum degree of 'skin in the game'. HK12 and HK13 show that the effective stochastic discount factor becomes

Date: November, 2018.

Leyla Jianyu Han: Faculty of Business and Economics, The University of Hong Kong, hanjyu@hku.hk
Kenneth Kasa: Department of Economics, Simon Fraser University, kkasa@sfu.ca
Yulei Luo: Faculty of Business and Economics, The University of Hong Kong, yulei.luo@gmail.com

We would like to thank seminar participants in University of Hong Kong for helpful comments. Luo thanks the General Research Fund (GRF No. HKU791913) in Hong Kong for financial support. All errors are the responsibility of the authors.

much more volatile, and that the nonlinearity induced by the constraint can account for observed state dependence in risk premia.[1]

Although the work of He and Krishnamurthy has been influential, it has not gone unquestioned. The key premise of HK12,13 is that some securities are too 'complex' for households to understand, so they delegate investment in these securities to specialists, whose actions cannot be precisely monitored. Cochrane (2017) questions how widespread and insurmountable this complexity problem really is,

> *Furthermore, if there is such an extreme agency problem, that delegated managers were selling during the buying opportunity of a generation, why do fundamental investors put up with it? Why not invest directly, or find a better contract?...So, in my view, institutional finance and small arbitrages are surely important frosting on the macro-finance cake, needed to get a complete description of financial markets in times of crisis...But are they also the cake?...Or can we understand the big picture of macro-finance without widespread frictions, and leave the frictions to understand the smaller puzzles, much as we conventionally leave the last 10 basis points to market microstructure.* (Cochrane (2017, p. 963-64))

Perhaps in anticipation of this critique, HK13 confine their analysis to the market for mortgage-backed securities.

In this paper, we argue that intermediary asset pricing is indeed 'the cake'. We operationalize complexity by assuming that agents face limits on their ability to process information, giving rise to so-called Rational Inattention (RI) (Sims (2003)). Although there have been many applications of Rational Inattention to financial markets, these applications either abstract from heterogeneity in information-processing capacity, or assume that any differences are fixed and immutable.[2] In contrast, we argue that trade in information-processing capacity is the raison d' etre of financial markets, and that when this trade is combined with the monitoring frictions of HK12, the scope of intermediary asset pricing models is greatly expanded. Although most households could manage their portfolios themselves, most *choose* not to do so.[3]

Another key ingredient of our analysis is the assumption that investment is subject to Knightian Uncertainty, or equivalently, ambiguity. Of course, this is not a new idea. Besides Knight (1921), Keynes (1936) argued that financial markets are by their very nature mechanisms for intermediating differences of opinion about ambiguous investment opportunities. However, it took many decades before this idea became operationalized in formal mathematical models. Our particular approach is based on the work of Hansen and Sargent (2008). Agents are assumed to have a (correctly specified) benchmark model of

---

[1]Brunnermeier and Sannikov (2014) develop a closely related model. He and Krishnamurthy (2018) survey the burgeoning literature on 'intermediary asset pricing', which followed HK12 and HK13.

[2][Provide citations]. Sims (2006) criticizes applications of RI in finance, arguing that in most financial applications information is scarce and costly, so the relevant constraint is on the supply-side, not the demand-side. Kacperczyk et. al. (2018) argue that differences in information-processing capacity contribute to wealth inequality, but do not allow agents to buy and sell this information-processing capacity.

[3]Pagel (2018) also bases portfolio delegation on inattention. However, in her model inattention is not based on information processing limits, but rather on 'information avoidance' (Golman et. al. (2017)), which arises from from loss aversion.

asset returns, which they distrust in a way that cannot be captured by a conventional finite-dimensional Bayesian prior. Rather than commit to a single model/prior, agents entertain a *set* of unstructured alternative models, and then optimize against the worst-case model. Since the worst-case model depends on an agent's own actions, agents view themselves as being immersed in a dynamic zero-sum game. Solutions of this game produce 'robust' portfolio policies. To prevent agents from being unduly pessimistic, in the sense that they attempt to hedge against empirically implausible alternatives, the hypothetical 'evil agent' who selects the worst-case model is required to pay a penalty that is proportional to the relative entropy between the benchmark model and the worst-case model.

Incorporating robustness into intermediary asset pricing models is important for a couple of reasons. First, it delivers a natural source of heterogeneous beliefs. In contrast to Maenhout (2004), we do *not* scale the entropy penalty parameter by the value function. Even with log preferences, a constant entropy penalty produces horizon effects in portfolio choice. In particular, the effective degree of ambiguity aversion depends on an agent's rate of time preference. Agents with a low rate of time preference are endogenously more ambiguity averse, since they care more about the future. Following HK12, we assume specialists are more patient than households, which in our model makes them more ambiguity averse. As a result, their pessimistic drift distortions are greater. This is important because it allows households to survive in the long-run, despite their greater impatience. In contrast, the model in HK12 does not possess a nondegenerate stationary equilibrium, which makes it difficult to evaluate empirically.[4]

The second reason robustness is important is that it tightens the specialist's capital constraint, making crisis episodes more likely. The constraint binds when households want to invest in the risky asset, but specialists do not. We assume throughout that differences in channel capacity are sufficiently great that households choose to remain in the contract. This imposes an upper bound on the fee the specialist can charge.[5] Because specialists are relatively ambiguity averse, they want to invest less in the risky asset. As a result, the constraint binds at higher levels of specialist wealth than without ambiguity. We inject cyclicality into this mechanism by assuming that dividend volatility is stochastic, and follows a 2-state jump process. In robust control models, pessimistic drift distortions 'hide behind' objective risk. When volatility increases, it becomes more difficult to discriminate among models, and this endogenously makes ambiguity increase as well. Since specialists have a higher degree of ambiguity aversion, their *relative* pessimism increases during volatile periods, thus making it more likely that the economy will hit the capital constraint.

The remainder of the paper is organized as follows. The next section explains how we incorporate ambiguity and information processing into the model of He and Krishnamurthy. Section 3 solves the model, and derives equilibrium diffusion processes for the risk premium and market price of risk. These processes depend on two state variables: (1) the

[4]HK13 remedies this defect by introducing nontradeable labor income. However, to keep the analysis tractable, they assume households live for a single-period and have a rather implausible bequest motive. HK12 note that when households are relatively impatient, their model can capture 'liquidation effects', in which asset values fall in response to financial disintermediation.

[5]In contrast, in HK12, where households have no ability to opt out, intermediation fees actually *increase* during crises.

endogenous wealth distribution between specialists and households, and (2) the exogenous level of dividend volatility. Section 4 calibrates the model, and studies its quantitative implications.

2. Ambiguity and Financial Intermediation

2.1. **Model Specification and the Full-Information Rational Expectations Solution.** We consider an infinite horizon continuous-time Lucas (1978)-type model. The economy is populated by two types of agents, specialists and households. There are two assets in the economy: one risky asset and one risk-free asset. The risky asset represents complex assets that require some expertise and information processing capacity. We assume the market is incomplete due to limited market participation as Basak and Cuoco (1998), where only experts who own the intermediaries can invest into the risky asset. Households can purchase channel capacity from specialists and make investments through intermediaries. Households thus face the decision to allocate portfolio between purchasing equity from intermediaries and the riskless short term bond. Figure 1 shows the market structure of the economy where the intermediary sector is indicated in the middle block.

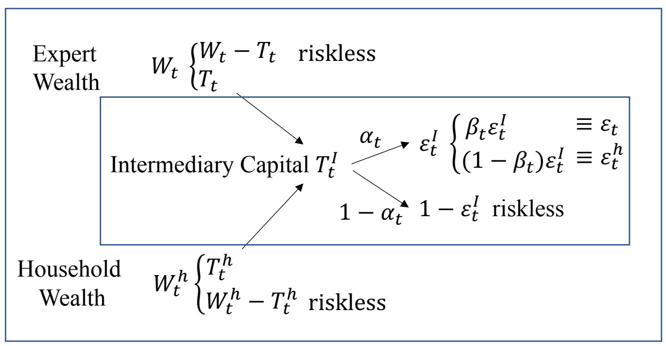

Figure 1: Market Structure and Intermediation Relationship

The total wealth of experts is W_t and households wealth is W_t^h. A superscript h denotes the households throughout the paper. Households allocate T_t^h to purchase intermediary equities and the remaining fraction is used to buy riskless bonds. Intermediaries absorb in sum T_t^I funds from households T_t^h and experts T_t, allocate a fraction α_t to the risky asset and $1-\alpha_t$ to the riskless bond. Assuming there is no short-selling constraint for the intermediary, we expect α_t to be larger than 1, i.e., specialists use leverage. In this case, specialists invest more than total intermediary capital into risky equity and borrow $(\alpha_t - 1)\, T_t^I$ from the bond market. The total risky asset position or intermediary's dollar exposure in risky asset is ε_t^I. Through an affine contract developed by HK(2012), $\beta_t \in [0, 1]$ is the share of returns going to specialists and $1-\beta_t$ to households. Thus, at time t, the specialist bears a total risk exposure of $\varepsilon_t = \beta_t \varepsilon_t^I$ and the household is offered an exposure of $(1-\beta t)\varepsilon_t^I$ to excess return.

The dividend of the risky asset is governed by a geometric brownian motion with return g and stochastic volatility σ_t,

$$\frac{dD_t}{D_t} = g_t dt + \sigma_t dZ_t, \tag{2.1}$$

where Z_t is a standard Brownian motion. Assume the volatility σ_t is a two-state Markov chain with state space $\Sigma = \{\sigma_H, \sigma_L\}$, where $\sigma_H > \sigma_L$. The intensity matrix of the continuous-time Markov chain is

$$\begin{bmatrix} -\lambda_H & \lambda_H \\ \lambda_L & -\lambda_L \end{bmatrix}.$$

λ_H is the rate of transition from the high volatility state to the low volatility state, and λ_L is the rate of transition from low to high. The return of the risky asset is defined as:

$$dR_t = \frac{D_t dt + dP_t}{P_t} = \mu_{R,t} dt + \sigma_{R,t} dZ_t. \tag{2.2}$$

where P_t is the risky asset price, $\mu_{R,t}$ is the expected return and $\sigma_{R,t}$ is the volatility of the risky asset. The riskless asset is in zero-net supply, and has an interest rate r_t. We define the risk premium in our model as $\pi_{R,t} \equiv \mu_{R,t} - r_t$. We assume that the measure of households are specialists are normalized to one. Both households and specialists are infinitely lived and have log preferences over consumption. Denote households (specialists) consumption rate as C_t^h (C_t). The household's objective is to:

$$\max_{\{C_t, \varepsilon_t^h\}} \mathbb{E}\left[\int_0^\infty e^{-\rho^h t} \ln C_t^h dt\right] \tag{2.3}$$

while the specialist's objective is to:

$$\max_{\{C_t, \varepsilon_t, \beta_t\}} \mathbb{E}\left[\int_0^\infty e^{-\rho t} \ln C_t dt\right] \tag{2.4}$$

where ρ^h and ρ denote the time discount rates for households and specialists, respectively. The dynamic budget constraints are

$$dW_t^h = \varepsilon_t^h (dR_t - r_t dt) - k_t \varepsilon_t^h dt + W_t^h r_t dt - C_t^h dt, \tag{2.5}$$

and

$$dW_t = \varepsilon_t (dR_t - r_t dt) + \max_{\beta_t \in [\frac{1}{1+m}, 1]} \left(\frac{1-\beta_t}{\beta_t}\right) k_t \varepsilon_t^* + W_t r_t dt - C_t dt. \tag{2.6}$$

where k_t is the exposure price that clears the intermediation market. Households obtain an exposure ε_t^h from the intermediary with an excess return indicated as the first term in the budget constraint, i.e., $\varepsilon_t^h(dR_t - r_t dt)$. Specialists bear a risky exposure ε_t by putting their own wealth into the intermediary. In order to use the intermediation service, households pay an intermediation fee $k_t \varepsilon_t^h \equiv K_t$. The second term denotes the transfer from households to intermediary. The specialist chooses the optimal contract share β_t to maximize the intermediation fee. The third term is the risk-free interest earns by the household (specialist) on his own wealth. The last term is the consumption expense. The optimal exposure supply schedule is $\beta_t^* = \frac{1}{1+m}$if $k_t > 0$ and $\beta_t^* \in \left[\frac{1}{1+m}, 1\right]$ if $k_t =$

0. Further, define the per-unit specialist fee q_t as $q_t \equiv \frac{K_t}{W_t} = \left(\frac{1-\beta_t^*}{\beta_t^*}\right) k_t \frac{\pi_{R,t}}{\sigma_{R,t}}$. The full-information rational expectations solutions for the above two maximization problems are:

$$C_t^{h*} = \rho^h W_t^h \text{and} \varepsilon_t^{h*} = \frac{\pi_{R,t} - k_t}{\sigma_{R,t}^2} W_t^h,$$

and

$$C_t^* = \rho W_t \text{and} \varepsilon_t^* = \frac{\pi_{R,t}}{\sigma_{R,t}^2} W_t.$$

Now define the scaled specialist wealth as an aggregate state variable in the economy:

$$x_t \equiv \frac{W_t}{D_t} \tag{2.7}$$

It is governed by the following stochastic process

$$\frac{dx_t}{x_t} = \mu_{x,t} dt + \sigma_{x,t} dZ_t, \tag{2.8}$$

where $\mu_{x,t}$ and $\sigma_{x,t}$ are the endogenously determined growth rate and volatility.

2.2. **Ambiguity and Robustness.** We are interested in studying an intermediation relationship between households and specialists when they have heterogeneous information processing capacities and a preference for robustness. In this section, we introduce ambiguity and robustness into the above otherwise standard He-Krishnamurthy model. Specifically, we assume that agents in our economy do not know the true model governing the evolution of the economy, and incorporate model uncertainty due to robustness into their decision problems. Following Hansen and Sargent (2001, 2006), we assume that when agents face model misspecifications, they take Equation (2.5) as the approximating model which is generated by the probability measure P. Assume the probability distribution in the distorted problem Q is absolutely continuous with respect to P. All the random variables and expectation operators for the robust problem below are defined on Q and $\mathcal{F}_t$ measurable. As argued in Anderson, Hansen and Sargent (2003) – henceforth AHS – the agents believe the approximating model is only a useful benchmark. However, they are concerned about the possibility that the approximating model is misspecified. In order to incorporate doubts about model specification, the agents conceive a class of models surrounding the approximating model, and make optimal decisions based on the range of possible models. An endogenous perturbation $\nu^h(W_t^h) = \nu_t^h$ is introduced to parameterize the change of measure from P to Q. By the Girsanov Theorem, the distortion only changes the drift, and the corresponding distorting model is

$$dW_t^h = \left(\varepsilon_t^h(\pi_{R,t} - k_t) + r_t W_t^h - C_t^h\right) dt + \sigma_{W,t}^h \left(\nu_t^h dt + dZ_t\right), \tag{2.9}$$

where $\sigma_{W,t}^h \equiv \sigma_{R,t}\varepsilon_t^h$. Note that the endogenous feedback effect of ν_t^h on W_t^h enables a wide range of model misspecifications. The alternative distorting models are vaguely specified and statistically difficult to distinguish within a given sample period T (AHS, 2003). This setting allows the approximating model to be perturbed by alternative functional forms, and enables the complexity of possible models to be nonlinear, and high-dimensional.[6] Thus ν_t^h reflects the pessimistic view of shock process from households, i.e., a pessimism parameterization. The evil agent chooses the drift adjustment ν_t^h to minimize the sum of the expected continuation payoff, but adjusted to reflect the additional drift component in (2.9), and of entropy penalty:

$$\inf_{\nu_t^h}\left[DV(W_t^h;Y_t^h)+\nu_t^h\sigma_{W,t}^hV_w+\frac{1}{2\theta_t^h}\mathcal{L}_t\right]$$

where

$$\mathcal{D}V(W_t^h;Y_t^h)=V_w\left[\varepsilon_t^h(\pi_{R,t}-k_t)+r_tW_t^h-C_t^h\right]+\frac{1}{2}V_{ww}(\varepsilon_t^h)^2\sigma_{R,t}^2+\mu_{Y,t}^h.$$

and $\mathcal{L}_t=\left(\nu_t^h\right)^2$ is the relative entropy (i.e., the expected log Radon-Nikodym derivative), which measures the distance between the two models. $\frac{1}{\theta_t^h}\geq 0$ is the weight on the entropy penalty term, i.e., the Lagrange multiplier imposed on the time t relative entropy constraint $\mathcal{L}_t$.[7] The household then solves the following HJB equation:

$$\sup_{\{C_t^h,\varepsilon_t^h\}}\inf_{\nu_t^h}\left[\ln C_t^h-\rho^hV+DV+\nu_t^h\sigma_{R,t}\varepsilon_t^hV_w+\frac{1}{2\theta_t^h}\left(\nu_t^h\right)^2\right].$$

subject to (2.9). Solving first the infimization part yields

$$\nu_t^{h*}=-\theta_t^hV_w\sigma_{W,t}^h.$$

Here we assume that the degree of ambiguity and robustness is constant, i.e., $\theta_t^h=\theta^h$ for two reasons. First, following the literature on robust control, Hansen, Sargent and Tallarini (1999) and AHS (2003) assume θ_t^h is fixed over time and state independent. Maenhout (2004) argues that θ_t^h should be state dependent in order to prevent robustness from diminishing and to ensure homotheticity in the CRRA utility setting. The log utility in this paper does not face this problem. Second, Hansen and Sargent (1998) showed that the model with time-varying θ_t^h but a constant relative entropy is observationally equivalent to the model with a constant robust parameter θ^h but has a time-varying entropy $\mathcal{L}_t$. Substituting for ν_t^h in the HJB equation gives:

$$0=\sup_{\{C_t^h,\varepsilon_t^h\}}\left[\ln C_t^h-\rho^hV+DV-\frac{\theta^h}{2}(\sigma_{R,t}\varepsilon_t^hV_w)^2\right]. \tag{2.10}$$

The following proposition summarizes the main results from the above model with:

Proposition 2.1. *Under robustness, the household's optimal consumption rule is*

$$C_t^{h*}=\rho^hW_t^h \tag{2.11}$$

[6] Learning is excluded here by assuming the impossibility to learn due to the model complexity. It is also possible to decompose ambiguity into time-varying non-learnable ambiguity (Markov hidden state) and time-invariant learnable ambiguity, e.g. Hansen and Sargent (2005, 2007, 2010) and Epstein and Schneider (2007).

[7] Hansen and Sargent (2001) show the observational equivalence between the robust control problem and multiple priors setup in the ambiguity literature (Gilboa and Schmeider, 1989; Epstein and Wang, 1994). Comparatively, θ^h indexes the set of priors used in multiple priors modeling.

and the optimal risk exposure is

$$\varepsilon_t^{h*} = \frac{\pi_{R,t} - k_t}{\gamma^h \sigma_{R,t}^2} W_t^h \tag{2.12}$$

where $\gamma^h = 1 + \theta^h/\rho^h$ and θ^h reflects the degree of robustness. Household's value function takes the form $V(W_t^h; Y_t^h) = \frac{1}{\rho^h} \ln W_t^h + Y_t^h$, where Y_t^h is a function of aggregate state x_t which satisfies a second-oder ODE.

Proof. See the Appendix 6.1 for the derivations. □

Now we turn to the specialist's problem. Take Equation (2.6) as the specialist's approximating model. The corresponding distorting model can thus be obtained by adding an endogenous distortion $\nu(W_t) = \nu_t$,

$$dW_t = (\varepsilon_t \pi_{R,t} + (q_t + r_t)W_t - C_t)\, dt + \sigma_{W,t}\,(\nu_t dt + dZ_t) \tag{2.13}$$

where $\sigma_{W,t} = \sigma_{R,t}\varepsilon_t$. Choosing a drift adjustment ν_t to:

$$\inf_{\nu_t} \left[DJ(W_t; Y_t) + \nu_t \sigma_{W,t} J_w + \frac{1}{2\theta_t} \mathcal{H}_t \right]$$

where

$$\mathcal{D}J(W_t; Y_t) = J_w \left[\varepsilon_t \pi_{R,t} + (q_t + r_t)W_t - C_t \right] + \frac{1}{2} J_{ww} {\varepsilon_t}^2 \sigma_{R,t}^2 + \mu_{Y,t}$$

and $\mathcal{H}_t = \nu_t^2$ denotes the relative entropy of the specialist's approximating model. $\frac{1}{\theta_t}$ is the weight that specialist puts on the entropy penalty term. We assume θ_t is time invariant which equals constant θ. The specialist solves the following HJB equation:

$$\sup_{\{C_t, \varepsilon_t\}} \inf_{\nu_t} \left[\ln C_t - \rho J + DJ + \nu_t \varepsilon_t \sigma_{R,t} J_w + \frac{1}{2\theta_t} \nu_t^2 \right]$$

subject to (2.13). Assuming θ_t is a constant and solving first the infimization part,

$$\nu_t^* = -\theta J_w \sigma_{W,t}.$$

Substituting for ν_t in the HJB equation gives

$$0 = \sup_{\{C_t, \varepsilon_t\}} \left[\ln C_t - \rho J + DJ - \frac{\theta}{2} \sigma_{R,t}^2 \varepsilon_t^2 J_w^2 \right] \tag{2.14}$$

The following proposition summarizes the main results from the above model with:

Proposition 2.2. *Under robustness, the specialist's optimal consumption rule is*

$$C_t^* = \rho W_t \tag{2.15}$$

and the optimal risk exposure is

$$\varepsilon_t^* = \frac{\pi_{R,t}}{\gamma \sigma_{R,t}^2} W_t. \tag{2.16}$$

where $\gamma = 1 + \theta/\rho$ and θ reflects the degree of robustness. The specialist's value function takes the form $J(W_t; Y_t) = \frac{1}{\rho} \ln W_t + Y_t$, where Y_t is a function of aggregate state x_t which satisfies a second-oder ODE.

Proof. See the Appendix 6.1 for the derivations. □

It is clear from the above optimal rules that the robustness parameter changes the desired exposures of households and specialists. When $\theta = \theta^h = 0$, the optimal choices of the household and specialist reduce to the original FI-RE model. It is worth noting that the effective coefficient of risk aversion $\begin{bmatrix} \gamma^h \\ \gamma \end{bmatrix}$ is determined by the interaction of the risk aversion ($\mathbf{1}$) in the logarithm utility specification, the degree of robustness $\begin{bmatrix} \theta^h \\ \theta \end{bmatrix}$, and the subjective discount factor $\begin{bmatrix} \rho^h \\ \rho \end{bmatrix}$ via the relationship $\begin{bmatrix} \gamma^h \\ \gamma \end{bmatrix} = \mathbf{1} + \begin{bmatrix} \theta^h/\rho^h \\ \theta/\rho \end{bmatrix}$. This representation coincides with Tallarini (2000) where $\begin{bmatrix} \gamma^h \\ \gamma \end{bmatrix}$ is more plausibly interpreted as a measure of the agents' aversion to model uncertainty instead of atemporal risk aversion. From the optimal risk exposure choices, (2.12) and (2.16), it is clear that the effective coefficient of risk aversion negatively influences the desired risk exposure for both agents. In particular, the effective coefficient of risk aversion increases with the degree of robustness, and decreases with the rate of time preference. The higher the degree of robustness, the lower the demand for risk exposure. When agents are more concerned about model uncertainty, they become more risk averse, and reduce their desired exposure to the risky assets. Moreover, when the discount rate is higher, agents value present well-being more than than the future, and thus behave less risk-averse to the present risks since they are so unlikely to forgo the present. Hansen, Sargent, Tallarini (1999) show the observational equivalence of the locus of robustness parameter and subjective discount factor $\begin{bmatrix} (\theta^h, \rho^h) \\ (\theta, \rho) \end{bmatrix}$ where movements along which preserve all equilibrium quantities. The pairs of $\begin{bmatrix} \theta^h \\ \theta \end{bmatrix}$ and $\begin{bmatrix} \rho^h \\ \rho \end{bmatrix}$ have explicitly negative relationship and variations in one can be completely offset by appropriate changes in the other one.[8] Later it will also be shown that γ^h and γ are the keys to determine the effective financial constraint. When $\rho^h/\rho = \theta^h/\theta$, these two problems are observationally equivalent to the FI-RE model in the constrained region. Following HK12 and HK13, we assume households are less patient than the specialists ($\rho^h > \rho$). Hence, observation equivalence would require households to be more ambiguity averse than the specialists ($\theta^h > \theta$).

The specialist's exposure supply is a step function:

$$\left\{ \begin{array}{ll} \frac{1-\beta_t^*}{\beta_t^*}\varepsilon_t^* \in [0, m\varepsilon_t^*], \text{for any} \beta_t^* \in [\frac{1}{1+m}, 1] & \text{if} k_t = 0, \\ m\varepsilon_t^* \text{with} \beta_t^* = \frac{1}{1+m} & \text{if} k_t > 0. \end{array} \right.$$

with $\varepsilon_t^* = \frac{\pi_{R,t}}{\gamma\sigma_{R,t}^2} W_t$, and k_t denotes the per-unit exposure price, which is the only determinant for optimal contract β_t^*. In contrast, the household's exposure demand is $\varepsilon_t^h = \frac{\pi_{R,t} - k_t}{\gamma^h \sigma_{R,t}^2} W_t^h$, which follows the risk sharing constraint $\varepsilon_t^h \leq m\varepsilon_t^*$. From Figure 2,

[8] Hansen, Sargent, Tallarini (1999) and Luo (2016) discuss the robustness effect from the precautionary saving perspective.

it is clear that both the exposure supply and demand functions are influenced by the robust parameters θ and θ^h.

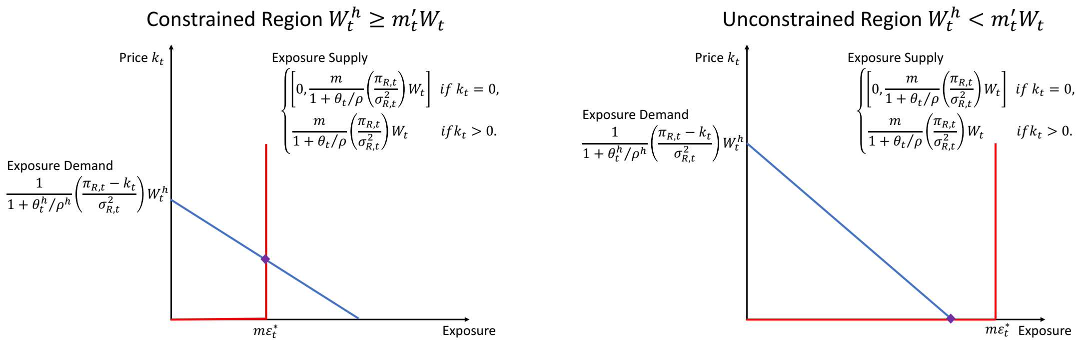

Figure 2: Unconstrained and Constrained Region in Equilibrium

3. Capacity-constrained Filtering and Robust Filtering

Now we assume that the evolution of the expected dividend growth rate g_t is unobservable to the agents, but it is assumed to follow a mean reverting process:

$$dg_t = \rho_g \left(\overline{g} - g_t\right) dt + \rho_{Dg}\sigma_g dZ_t + \sqrt{1-\rho_{Dg}^2}\sigma_g dZ_{g,t}, \tag{3.17}$$

where Z_t and $Z_{g,t}$ are independent standard Brownian motions with $Z_{g,t}$ capturing innovations to the growth rate that are not correlated with the dividend process. The assumption of the mean-reverting expected growth rate of dividend is consistent with the facts on real business cycles.

Here we assume that the typical investor learn the state (g_t) via finite information-processing capacity (rational inattention, or RI). The main idea of Sims' RI theory is that agents with finite capacity react to the innovations to the state gradually and incompletely because the channel along which information flows cannot carry an infinite amount of information. Following Peng (2004) and Kasa (2006), we adopt the noisy-information specification and assume that the investor observes only a noisy signal containing imperfect information about g_t:

$$dg_t^* = g_t dt + d\xi_t, \tag{3.18}$$

where ξ_t is the noise shock, and is a Brownian motion with mean 0 and variance Λ (in the RI setting, the variance, Λ, is a choice variable for the agent). Following the RI literature, we assume that ξ_t is independent of the Brownian motions, Z_t and $Z_{g,t}$.

To model RI due to finite capacity, we follow Sims (2003) and impose the following constraint on the investor's information-processing ability:

$$\mathcal{H}\left(g_{t+\Delta t}|\mathcal{I}_t\right) - \mathcal{H}\left(g_{t+\Delta t}|\mathcal{I}_{t+\Delta t}\right) \leq \kappa\Delta t, \tag{3.19}$$

where κ is the investor's information channel capacity; $\mathcal{H}\left(g_{t+\Delta t}|\mathcal{I}_t\right)$ denotes the entropy of the state prior to observing the new signal at $t+\Delta t$; and $\mathcal{H}\left(g_{t+\Delta t}|\mathcal{I}_{t+\Delta t}\right)$ is the entropy after observing the new signal. κ imposes an upper bound on the amount of information – that is, the change in the entropy – that can be transmitted in any given period. To apply this information constraint to the state transition equation, we first rewrite (3.17) in the time interval of $[t, t+\Delta t]$:[9]

$$g_{t+\Delta t} = \rho_0 + \rho_1 g_t + \rho_2 \sqrt{\Delta t} \varepsilon_{t+\Delta t}, \tag{3.20}$$

where $\rho_0 = \overline{g}\left(1 - \exp\left(-\rho_g \Delta t\right)\right)$, $\rho_1 = \exp\left(-\rho_g \Delta t\right)$, $\rho_2 = \sigma_g \sqrt{\left(1 - \exp\left(-2\rho_g \Delta t\right)\right) / \left(2\rho_g \Delta t\right)}$, and $\varepsilon_{t+\Delta t}$ is the time-$(t+\Delta t)$ standard normal distributed innovation to permanent income. Taking conditional variances on both sides of (**??**) and substituting it into (**??**), we have

$$\ln\left(\rho_1^2 \Sigma_t + \rho_2^2\right) - \ln\left(\Sigma_{t+\Delta t}\right) = 2\kappa\Delta t,$$

which reduces to

$$\dot{\Sigma}_t = 2\left(-\rho_g - \kappa\right)\Sigma_t + \sigma_g^2,$$

as $\Delta t \to 0$, where $\Sigma_t = E_t\left[\left(g_t - \widehat{g}_t\right)^2\right]$ the conditional variance at t. In the steady state in which $\dot{\Sigma}_t = 0$, the steady state conditional variance can be written as:

$$\Sigma = \frac{\sigma_g^2}{2\left(\kappa + \rho_g\right)}. \tag{3.21}$$

Proof. The IPC,

$$\ln\left(\rho_1^2 \Sigma_t + \rho_2^2\right) - \ln \Sigma_{t+\Delta t} = 2\kappa\Delta t,$$

can be rewritten as

$$\begin{aligned}
\ln\left(\exp\left(-2\rho_g \Delta t\right)\Sigma_t + \frac{1 - \exp\left(-2\rho_g \Delta t\right)}{2\rho_g \Delta t}\Delta t \sigma_g^2\right) - \ln \Sigma_{t+\Delta t} &= 2\kappa\Delta t, \\
\ln\left(\exp\left(-2\rho_g \Delta t\right)\Sigma_t + \frac{1 - \exp\left(-2\rho_g \Delta t\right)}{2\rho_g}\sigma_g^2\right) - \ln \Sigma_{t+\Delta t} &= 2\kappa\Delta t,
\end{aligned}$$

which can be reduced to

$$\Sigma_{t+\Delta t} - \Sigma_t = \left(\exp\left(2\left(-\rho_g - \kappa\right)\Delta t\right) - 1\right)\Sigma_t + \frac{\exp\left(2\left(-\rho_g - \kappa\right)\Delta t\right) - \exp\left(-2\kappa\Delta t\right)}{-2\rho_g}\sigma_g^2.$$

Dividing Δt on both sides of this equation and letting $\Delta t \to 0$, we have the following continuous-time updating equation for Σ_t:

$$\dot{\Sigma}_t = \lim_{\Delta t \to 0} \frac{\Sigma_{t+\Delta t} - \Sigma_t}{\Delta t} = 2\left(-\rho_g - \kappa\right)\Sigma_t + \sigma_g^2.$$

□

[9] Note that here we use the fact that $\Delta B_t = \varepsilon_t \sqrt{\Delta t}$, where ΔB_t represents the increment of a Wiener process.

In the filtering problem, we assume that the agent's objective is to minimize the following minimum square errors (MSE):

$$J_t = \inf_{\{\widehat{g}_t\}} E_t \left[(g_t - \widehat{g}_t)^2 \right], \tag{3.22}$$

subject to the information-processing constraint. Note that in the steady state the value of J_t is just $\Sigma = \frac{\sigma_g^2}{2(\kappa+\rho_g)}$.

Specifically, we may think that the model with imperfect state observations can be decomposed into a two-stage optimization procedure:

(1) The optimal filtering problem determines the optimal evolution of the perceived state;
(2) The optimal control problem in which the decision makers treat the perceived state as the underlying state when making optimal decisions.

Here we assume ex post Gaussian distributions and Gaussian noise but adopt log preferences. In stage 1, consumers need to estimate the unobserved state (g_t) using its prior distribution and all processed and available information (i.e., their noisy observations, $\mathcal{F}_t = \left\{ g_j^* \right\}_{j=0}^{t}$). Specifically, consumers rationally compute the conditional distribution of the unobserved state and represent the original optimization problem as a Markovian one. Given the Gaussian prior $g_0 \sim N(\widehat{g}_0, \Sigma_0)$, finding the posterior distribution of g_t becomes a standard filtering problem that can be solved using the Kalman-Bucy filtering method. Specifically, the optimal estimate for g_t given $\mathcal{F}_t = \left\{ g_j^* \right\}_{j=0}^{t}$ in the mean square sense coincides with the conditional expectation: $\widehat{g}_t = E_t[g_t]$, where $E_t[\cdot]$ is based on $\mathcal{F}_t$. Applying Theorem 12.1 in Liptser and Shiryaev (2001), we can obtain the filtering differential equations for $\widehat{g}_t$ and Σ_t as follows:

$$d\widehat{g}_t = \rho_g (\overline{g} - \widehat{g}_t) dt + K_t d\eta_t, \tag{3.23}$$

$$\dot{\Sigma}_t = -\Lambda K_t^2 - 2\rho_g \Sigma_t + \sigma_g^2, \tag{3.24}$$

given $g_0 \sim N(\widehat{g}_0, \Sigma_0)$, where

$$K_t = \frac{\Sigma_t}{\Lambda} \tag{3.25}$$

is the Kalman gain and

$$d\eta_t = \sqrt{\Lambda} dB_t^*, \tag{3.26}$$

with mean $E[d\eta_t] = 0$ and $var(d\eta_t) = \Lambda dt$, where B_t^* is a standard Brownian motion and Λ is to be determined. Note that η_t is a Brownian motion with mean 0. Although the Brownian variable, ξ_t, is not observable, the innovation process, η_t, is observable because it is derived from observable processes (i.e., dg_t^* and $\rho_g(\overline{g} - \widehat{g}_t) dt$). In this case, the path of the conditional expectation, $\widehat{s}_t$, is generated by the path of the innovation process, η_t. In the steady state, we have the following proposition:

Proposition 3.1. *Given finite capacity κ, in the steady state, the evolution of the perceived state can be written as:*

$$d\widehat{g}_t = \rho_g (\overline{g} - \widehat{g}_t) dt + \widehat{\sigma} dB_t^*, \tag{3.27}$$

where

$$\widehat{\sigma} \equiv \Sigma/\sqrt{\Lambda} = f(\kappa)\,\sigma_g, \tag{3.28}$$

$f(\kappa) = \sqrt{\kappa/(\kappa+\rho_g)} < 1$ *(i.e., the standard deviation of the estimated state is greater than that of the true state),*

$$\Lambda = \frac{\sigma_g^2}{4\kappa(\kappa+\rho_g)} \tag{3.29}$$

is the steady state conditional variance, and

$$K = 2\kappa \tag{3.30}$$

is the corresponding Kalman gain.

Proof. In the steady state in which $\dot{\Sigma}_t = 0$, substituting the definition of the Kalman gain, (3.25), into $-\Lambda K_t^2 + 2r\Sigma_t + \sigma_g^2 = 0$ and using $\Sigma = \frac{\sigma_g^2}{2(\kappa+\rho_g)}$, we can easily obtain that:

$$\Lambda = \frac{\sigma_g^2}{4\kappa(\kappa+\rho_g)} \text{ and } K = 2\kappa.$$

□

3.1. **Robust Filtering.** In the robust filtering problem, as shown in Basar and Bernhard (1995), Pan and Basar (1996), Ugrinovskii and Petersen (2002), and Kasa (2006), a robust filter can be characterized by the following dynamic zero-sum game:

$$L_t = \inf_{\{\widehat{g}_s\}} \sup_{\{Q\}} \left\{ E^Q\left[(g_s - \widehat{g}_s)^2\right] - \theta^{-1} H(Q|P) \right\}, \tag{3.31}$$

where P and Q are the approximating and distorted models, respectively, and H_∞ is the relative entroy and is bounded from above. As shown in Dai Pra, Meneghini, and Runggaldier (1996), the entropy constrained robust filtering problem is equivalent with the following risk-sensitive filtering problem:

$$\frac{1}{\theta}\log\left(\int \exp\left(\theta F(g_s, \widehat{g}_s)\right) dP\right) = \sup_Q \left\{ \int F(g_s, \widehat{g}_s)\, dQ - \theta^{-1} H_\infty(Q|P) \right\}, \tag{3.32}$$

where $F \equiv (g_s - \widehat{g}_s)^2$ is the loss function. The steady state conditional variance is determined by the following Riccati equation:[10]

$$2\rho_g \Sigma = \sigma_g^2 - \left(\frac{1}{\sigma^2} - \theta\right)\Sigma^2,$$

which means that

$$\Sigma^* = \frac{\sigma^2\left[-\rho_g + \sqrt{\rho_g^2 + (1-\theta\sigma^2)\, SNR}\right]}{1-\theta\sigma^2}, \tag{3.33}$$

where $SNR = \sigma_g^2/\sigma^2$ is the signal-to-noise ratio. The corresponding Kalman gain is

$$K = \frac{\Sigma^*}{\sigma^2} = \frac{-\rho_g + \sqrt{\rho_g^2 + (1-\theta\sigma^2)\, SNR}}{1-\theta\sigma^2}.$$

[10] For simplicity here we assume that $\rho_{Dg} = 0$.

The value function can be written as

$$\begin{aligned} J_t &= \frac{1}{\theta}\log\left(\int \exp\left(\theta F\left(g_s,\widehat{g}_s\right)\right)dP\right) = \frac{1}{\theta}\log\left(E_t \exp\left(\theta\left(g_t-\widehat{g}_t\right)^2\right)\right) \\ &= E_t\left[\left(g_t-\widehat{g}_t\right)^2\right] + \frac{1}{2}\theta var_t\left[\left(g_t-\widehat{g}_t\right)^2\right] \\ &= \Sigma_t + \frac{1}{2}\theta\left(E_t\left[\left(g_t-\widehat{g}_t\right)^4\right] - \left(E_t\left[\left(g_t-\widehat{g}_t\right)^2\right]\right)^2\right) \\ &= \Sigma_t + \theta\Sigma_t^2 \end{aligned}$$

where we use the fact that $E_t\left[\left(g_t-\widehat{g}_t\right)^4\right] = (3!!)\Sigma_t^2$. In the steady state,

$$J = \Sigma + \theta\Sigma^2, \tag{3.34}$$

i.e., the welfare loss under robustness is larger than that obtained in the model without robustness.

3.2. **Observational Equivalence (OE) between Capacity Constrained Filtering and Robust Filtering.** We can establish the first observational equivalence between capacity constrained filtering and robust filtering when the filtering problems lead to the same Kalman gain:

$$\kappa = \frac{-\rho_g + \sqrt{\rho_g^2 + \left(1-\theta\sigma^2\right)SNR}}{2\left(1-\theta\sigma^2\right)}. \tag{3.35}$$

The following figure illustrates the relationship between κ and θ.

itbpF4.0145in3.5129in0inoe1.eps

Since both capacity and ambiguity increase the Kalman gain of the filtering problem, one way agents can implement a robust filter is by re-allocating some of their limited capacity to decisions that demand a relatively high degree of ambiguity. This result is the same as that obtained Kasa (2006).

4. Theoretical Implications

4.1. **Market Equilibrium.** Here we provide a detailed definition of market equilibrium in our model economy:

Definition 4.1. *An equilibrium for the economy is a set of progressively, measurable price processes* $\{P_t, r_t, R_t\}$ *and* $\{k_t\}$*, households' decisions* $\{C_t^{h*}, \varepsilon_t^{h*}\}$*, and experts' decisions* $\{C_t^*, \varepsilon_t^*, \beta_t^*\}$ *such that*

(1) *Given the processes, decisions optimally solve (2.3) and (2.4).*
(2) *The intermediation market reaches equilibrium with risk exposure clearing condition,*

$$\varepsilon_t^{h*} = \frac{1-\beta_t^*}{\beta_t^*}\varepsilon_t^*. \tag{4.36}$$

(3) *The stock market clears:*

$$\varepsilon_t^* + \varepsilon_t^{h*} = P_t. \tag{4.37}$$

(4) *The goods market clears:*

$$C_t^* + C_t^{h*} = D_t. \tag{4.38}$$

(5) *Transversality conditions satisfy:*

$$\lim_{t\to\infty} \mathbb{E}\left[\exp\left(-\rho^h t\right) V(W_t^h, t)\right] = 0 \tag{4.39}$$

$$\lim_{t\to\infty} \mathbb{E}\left[\exp\left(-\rho t\right) J(W_t, t)\right] = 0 \tag{4.40}$$

In the unconstrained region (see the right panel of Figure 2), the exposure supply exceeds the demand. There exists an abundance of intermediary supply so that specialists must set the intermediation fee to zero to attract all the exposure demand from the household. In this case, both the per-unit exposure price k_t and per-unit of specialist wealth fee q_t are zero. The incentive-compatibility constraint is slack ($\beta_t^* > \frac{1}{1+m}$), as well as the risk-sharing constraint is slack, such that

$$\varepsilon_t^h|_{k_t=0} < m\varepsilon_t \Longleftrightarrow \frac{\pi_{R,t}}{\gamma^h \sigma_{R,t}^2} W_t^h < m \frac{\pi_{R,t}}{\gamma \sigma_{R,t}^2} W_t,$$

where we assume risk premium $\pi_{R,t}$ is positive. Define $\tilde{m} \equiv \frac{\gamma^h}{\gamma} m$, the risk-sharing constraint is translated into the equity capital constraint

$$W_t^h < \tilde{m} W_t.$$

Intermediary earns higher exposure, so that households put all the wealth into the intermediation, $T_t^h = W_t^h$.

In the constrained region (see the left panel of Figure 2), the exposure supply is less than demand, $k_t \geq 0$. The incentive-compatibility constraint is binding ($\beta_t^* = \frac{1}{1+m}$) and the equity capital constraint is binding, such that

$$\begin{aligned} \varepsilon_t^h = m\varepsilon_t \Longleftrightarrow \frac{\pi_{R,t} - k_t}{\gamma^h \sigma_{R,t}^2} W_t^h &= m \frac{\pi_{R,t}}{\gamma \sigma_{R,t}^2} W_t \\ \Longleftrightarrow W_t^h &\geq \tilde{m} W_t. \end{aligned} \tag{4.41}$$

In equilibrium, the specialist earns a rent $q_t = \frac{k_t m \varepsilon_t^*}{W_t} = \frac{m k_t}{\gamma} \frac{\pi_{R,t}}{\sigma_{R,t}^2} \geq 0$ for scarce intermediary service. When k_t increases, ε_t^{h*} decreases, hence exposure demand drops. Households would not put all their wealth into the intermediary, $T_t^h = \tilde{m} W_t \leq W_t^h$, thus induces the financial constraint for the intermediation. Define effective financial constraint

Definition 4.2.

$$\tilde{m} \equiv \frac{\gamma^h}{\gamma} m. \tag{4.42}$$

Robustness concerns change the binding conditions for the economy through the effective financial constraint $\tilde{m}$. When $\gamma^h = \gamma \Leftrightarrow \theta^h/\theta = \rho^h/\rho$, ambiguity parameters don't change the financial constraint, i.e. $\tilde{m} = m$. However, as HK12 and HK13, assuming the specialist is more patient than the household, $\rho^h > \rho$, the existence of ambiguity causes the effective financial constraint to be scaled by relative ambiguity aversion. Thus, specialists become more constrained even if they face the same level of ambiguity as the households, i.e. when $\theta^h = \theta$, $\tilde{m} < m$. Later we will see in addition to the wealth distribution among households and specialists, robustness parameters influence the equity capital binding conditions as well as the conditions for whether the economy is in the constrained region or not. This is similar to the role of the financial constraint. We incorporate ambiguity into the financial constraint to make it "endogenous" by the agents' ambiguity. The effective financial constraint can also be treated as an "adjusted" financial constraint with the adjustment of $\frac{\gamma^h}{\gamma} = \frac{1+\theta^h/\rho^h}{1+\theta/\rho}$, which is the relative ratio of effective risk aversion between the household and specialist. During a financial crisis, people fear more about the adverse state and are more uncertain about the true state. This fear will influence the asset market through effective financial constraint directly. Using (4.42), we have

$$\frac{d\tilde{m}}{d\theta} = -\frac{\gamma^h}{\rho\gamma^2} m < 0 \text{ and } \frac{d\tilde{m}}{d\theta^h} = \frac{1}{\rho^h \gamma} m > 0.$$

From these results, we can see that θ and θ^h play opposite roles in determining $\tilde{m}$ which captures the inverse of agency friction. When the specialist's ambiguity aversion θ is larger, the effective financial constraint $\tilde{m}$ is smaller, which makes the equity capital constraint easier to bind for a given wealth distribution. On the other hand, when household's ambiguity θ^h increases, $\tilde{m}$ is larger, thus relieving the financial constraint faced by the intermediary for a given wealth. The intuition is that when experts are more uncertain they are less willing to put the required 'skin in the game', and so the agency friction worsens. However, when household become more ambiguous, they have a lower demand for financial intermediation, which relaxes the agency friction. HK12 document the constraint effect and the sensitivity effect of the intermediation multiplier which are also applicable here.

4.2. **Asset Pricing Implications.**

4.2.1. *The Equity Capital Constraint and the Price-Dividend Ratio.* Since bonds are in zero net supply, the asset market clears when aggregate wealth equals the market value of the risky asset,

$$W_t^h + W_t = P_t. \tag{4.43}$$

In equilibrium, from the goods market clearing condition (4.38) and the optimal consumption rules of households and specialists,

$$\rho W_t + \rho^h W_t^h = D_t.$$

Thus, the equilibrium price/dividend ratio is

$$\frac{P_t}{D_t} = \frac{1}{\rho^h} + (1 - \frac{\rho}{\rho^h}) x_t = \frac{1 + \varDelta\rho x_t}{\rho^h} \tag{4.44}$$

where $x_t \equiv W_t/D_t$ is the aggregate state variable and $\varDelta\rho \equiv \rho^h - \rho$. Notice that robustness concerns do not have a first order effect on the price/dividend ratio. Robustness only indirectly influences it through a wealth effect. When the risk sharing constraint just starts to bind, the threshold level of the state x^c could be derived as

$$\varepsilon_t^h = m\varepsilon_t,$$

which means that $\frac{P_t - W_t}{\gamma^h} = \frac{mW_t}{\gamma}$. Together with the equilibrium price/dividend ratio above yields:

$$x^c = \frac{1}{\tilde{m}\rho^h + \rho}. \tag{4.45}$$

When $x_t \leq x^c$, the economy is within the constrained region, otherwise, when $x_t > x^c$, the economy is unconstrained. Agents are ambiguous, thus both the robust concerns from households and expert s influence the critical level of x^c through the effective financial constraint $\tilde{m}$. As Figure 3 shows, when household become more ambiguous about the world, $\tilde{m}$ is larger and x^c is smaller thus the constrained region is smaller. On the other hand, when expert s becomes more ambiguous, $\tilde{m}$ is smaller so that they face tighter effective financial constraint which is easier to bind for a given wealth distribution. Hence, the constrained region is larger thus the probability for the economy to drop into the constrained region is higher.

4.2.2. *Specialist's Portfolio Share.* The specialist makes a portfolio choice to invest a share α_t of the total equity $T_t^I = W_t + T_t^h$ into the risky asset and the rest into the riskless bond. Thus, the total exposure is

$$\varepsilon_t^I = \alpha_t T_t^I$$

which yields the following implementation constraint:

$$\varepsilon_t^* + \varepsilon_t^{h*} = \alpha_t (W_t + T_t^h). \tag{4.46}$$

This implementation constraint requires the specialist to choose α_t to reach the optimal risk exposure ε_t^*. Household obtains the desired exposure ε_t^{h*} by choosing how much wealth T_t^h to contribute to the intermediation.

Proposition 4.3. *In unconstrained region, the share of the return is*

$$\beta_t^U = \frac{1}{1 + \frac{\gamma}{\gamma^h}\left(\frac{1}{\rho^h x_t} - \frac{\rho}{\rho^h}\right)}. \tag{4.47}$$

In constrained region,

$$\beta_t = \frac{1}{1+m}. \tag{4.48}$$

Proof: In the unconstrained case, per-unit exposure price is zero. Recall that the share of return contract $\beta_t \equiv \varepsilon_t^* / \varepsilon_t^I$. Since the robust concern distorts the specialist's desired risk exposure ε_t^*, the choice of share contract turns into

$$\beta_t^U = \frac{W_t}{W_t + \frac{\gamma}{\gamma^h} W_t^h} \text{and}\, k_t = 0.$$

Now the specialist and household no longer hold the equity claims according to their wealth contributions as in HK(2012)'s case, but with a distortion term $\frac{\gamma}{\gamma^h}$ which equals the inverse of distortion on the financial constraint. Note that although agency friction m doesn't enter β_t^U in unconstrained region, both robustness parameters distort the contract share alternatively. Replacing W_t^h with asset market clearing condition (4.43) yields:

$$\beta_t^U = \frac{1}{1 + \frac{\gamma}{\gamma^h}\left(\frac{1}{\rho^h x_t} - \frac{\rho}{\rho^h}\right)}.$$

By the imposed assumption that $0 < \beta_t^U \leq 1$, x_t should be limited within $(0, 1/\rho]$. Later we will show that in order for the risk-free rate to be valid whenever robustness exists, $x_t \neq 1/\rho$. From now on, we assume

$$x_t \in \{ \begin{array}{ll} (0,1/\rho] & \text{for}\, \theta = \theta^h = 0 \\ (0,1/\rho) & \text{others.} \end{array}$$

Furthermore, we have

$$\frac{d\beta_t^U}{d\theta} = \frac{-\left(\frac{1}{\rho x_t} - 1\right)\rho^h\gamma^h}{\left[\rho^h\gamma^h + \gamma\left(\frac{1}{x_t} - \rho\right)\right]^2} \leq 0 \text{and} \frac{d\beta_t^U}{d\theta^h} = \frac{\left(\frac{1}{\rho x_t} - 1\right)\rho\gamma}{\left[\rho^h\gamma^h + \gamma\left(\frac{1}{x_t} - \rho\right)\right]^2} \geq 0.$$

Note that when $\theta = \theta^h = \bar{\theta}$, we have

$$\beta_t^U = \frac{\rho^h + \bar{\theta}}{\rho^h - \rho + (1 + \bar{\theta}/\rho)\frac{1}{x_t}}, \frac{d\beta_t^U}{d\bar{\theta}} = \frac{-\Delta\rho\left(\frac{1}{\rho x_t} - 1\right)}{\left(\Delta\rho + (1 + \bar{\theta}/\rho)\frac{1}{x_t}\right)^2} \leq 0.$$

A higher ambiguity from the specialists decreases the share of returns that go to them. In other words, when specialists have more doubt about their approximating models, they prefer a lower return share which comes from a lower risk exposure. In contrast, a higher ambiguity from household increases the contract share in the unconstrained region, where they want to bear lower risk and transfer the risk to the intermediation. By calculating the scale effect of the two robustnesses $\left|\frac{d\beta_t^U}{d\theta}\right|$ and $\left|\frac{d\beta_t^U}{d\theta^h}\right|$, it is clear to see that if $\rho^h\gamma^h > \rho\gamma \iff \theta^h + \rho^h > \theta + \rho$, the expert 's preference for robustness has a stronger effect on the contract share than the household in the unconstrained case. In the constrained region, the share of return is determined by the incentive constraint of specialist. In order to prevent the specialist from shirking, households need to pay a positive intermediation fee and exposure price to the intermediary, thus

$$\beta_t = \frac{1}{1+m} \text{and}\, k_t > 0.$$

expert

Proposition 4.4. *In the unconstrained region, the desired risk exposure and optimal portfolio choice are*

$$\varepsilon_t^{U*} = \frac{1}{1 + \frac{\gamma}{\gamma^h}\left(\frac{1}{\rho^h x_t} - \frac{\rho}{\rho^h}\right)} P_t. \tag{4.49}$$

$$\alpha_t^{U*} = 1. \tag{4.50}$$

In the constrained region,

$$\varepsilon_t^* = \frac{1}{1+m} P_t. \tag{4.51}$$

$$\alpha_t^* = \frac{\frac{1}{x_t} + \rho^h - \rho}{(1+\tilde{m})\rho^h}. \tag{4.52}$$

Proof. In the unconstrained region, $T_t^h = W_t^h$, both households and specialists put all their wealth into the intermediation, such that the total risk exposure equals $\varepsilon_t^I = \alpha_t(W_t + W_t^h)$. The equilibrium conditions (4.37) and (4.43) yield $\alpha_t^{U*} = 1$. The risk exposure for the specialist can be derived as follows:

$$\varepsilon_t^* + \varepsilon_t^{h*} = W_t + W_t^h$$

$$\varepsilon_t^* + \frac{1-\beta_t^U}{\beta_t^U}\varepsilon_t^* = P_t$$

$$\varepsilon_t^* = \beta_t^U P_t \Longrightarrow \varepsilon_t^{U*} = \frac{1}{1+\frac{\gamma}{\gamma^h}\left(\frac{1}{\rho^h x_t} - \frac{\rho}{\rho^h}\right)} P_t.$$

In the constrained region, the expert holds $\beta_t = \frac{1}{1+m}$ share of risk. Hence, the expert 's risk exposure

$$\varepsilon_t^* = \beta_t \varepsilon_t^I = \frac{1}{1+m} P_t.$$

Further, the expert 's portfolio share is

$$\alpha_t = \frac{\varepsilon_t^I}{W_t + T_t^h} = \frac{P_t}{(1+\tilde{m})W_t} \Longrightarrow \alpha_t^* = \frac{\frac{1}{x_t} + \rho^h - \rho}{(1+\tilde{m})\rho^h}.$$

From (4.52), we have

$$\frac{d\alpha_t^*}{d\theta} > 0 \text{and} \frac{d\alpha_t^*}{d\theta^h} < 0,$$

which means that the two robustness parameters play opposite roles in determining the equilibrium portfolio share α_t in the risky asset. From Figure 3, in the unconstrained region, $\alpha_t = 1$ such that the expert invests all of the intermediary's equity capital into the risky asset. Once the constraint is binding. $\alpha_t > 1$ means the expert holds above 100% of the total equity and borrows $(\alpha_t - 1)(W_t + T_t^h)$ riskless bonds. There are three effects of heterogeneous ambiguity:The first is the "*constraint effect*" under a given wealth distribution.[11] When the expert 's (household's) ambiguity aversion θ (θ^h) is larger (smaller), the effective financial constraint $\tilde{m}$ is tighter (looser), which induces the expert to hold a larger (smaller) portfolio share α_t of risky assets. The constraint effect could be further decomposed into two channels*, (i) the "general equilibrium channel*" and (ii) the "*intermediary expertise effect*". Specifically, a higher ambiguity aversion θ induces the expert to have more doubts about the approximating model thus more unlikely to expose to the risky asset. Hence, the exposure supply decreases, risky asset price decreases accordingly in equilibrium using equation (4.51). Lower price triggers a higher demand for risky asset portfolio holding for a given return. This is the general equilibrium effect. When the households realize that the expert loses the expertise in determining the probability distributions, they put even less wealth into the intermediation, which worsens the agency problem. Since the households are more sensitive to the expertise of the expert implied by the ambiguity aversion, the expert loses more equity capital due to the household participation decline by lowering the risk exposure. As a consequence, the equilibrium portfolio

[11] HK(2012) illustrate the constraint effect of intermediation multiplier m as an accelerator of the tightness for capital constraint. In our model, ambiguity is endogenous as a scale on m and plays the similar role in terms of constraint effect through the effective financial constraint.

share has to adjust above 1 to make the risk exposure optimal, where the extra leverage is borrowed from the short-term bond market. This is the intermediary expertise effect implied by the expert ambiguity change. On the other hand, when the households become more ambiguous, i.e. θ^h is larger, the effective financial constraint $\tilde{m}$ becomes looser, such that expert will invest less in the risky asset. The general equilibrium effect raises when the households more concern about their model uncertainty, such that they want to bear less risk exposure and reduce risk exposure demand, hence the risky asset price decreases in equilibrium. The expert needs to reduce the optimal exposure from equation (4.51). Total risk exposure declines but equity capital rises. As a result, portfolio share has to decrease to meet the optimization in equilibrium. Furthermore, households trust the expertise of expert , reflected in less ambiguity, more than themselves when they are more ambiguous, so that they don't want to make portfolio decisions and put more wealth into the intermediation and let the expert do the portfolio choice, thus releasing the effective financial constraint for the expert . The intermediation expertise effect makes the agency problem less severe hence the expert bears less risk exposure and invests less in the risky asset. The relative ratio of both agents' relative risk aversions determines the equilibrium result. The agent who's effective risk aversion changes larger during financial crisis will dominate the equilibrium portfolio choice even though the expert is the marginal investor. This coincides with Bossaerts et al (2010) that prices reflect the average beliefs of heterogeneous agents. Moreover, the ambiguity dispersion of two agents will scale up the tightness of financial constraint. The effective financial constraint $\tilde{m}$ plays a crucial role of financial multiplier or accelerator, which incorporates the sensitivities of agents' heterogeneous ambiguities, in transmitting the belief dispersion into the asset market. In the special case when $\gamma = \gamma^h$, the heterogeneous ambiguity constraint effect neutralizes. The second effect is the "*wealth effect*" under certain effective financial constraint. Later, Figure show that the unconditional mean of scaled expert wealth x_t is a decreasing function of θ which will even amplify the constraint effect. However, there is no steady state solution $\mathbb{E}[x_t \mid \theta = 0]$ for any θ^h. From equation (4.52), constraint effect is a first order effect of ambiguity to portfolio share while wealth effect is not. This further emphasizes the most important channel of heterogeneous ambiguity through effective financial constraint transmitting into asset market dynamics. The third is the *"sensitivity effect"* to exogenous shocks. HK (2012) document the sensitivity effect such that one percent of expert s' wealth drop will induce m (> 1) percent equity participation of households' wealth into the intermediation. The economy is more sensitive to the changes in the aggregate state in the constrained region. It is easy to show,

$$\frac{d}{d\theta}\left(\frac{d\alpha_t^*}{dx_t}\right) = \frac{1}{(1+\tilde{m})^2\rho^h x_t^2}\frac{d\tilde{m}}{d\theta} < 0 \text{and} \frac{d}{d\theta^h}\left(\frac{d\alpha_t^*}{dx_t}\right) = \frac{1}{(1+\tilde{m})^2\rho^h x_t^2}\frac{d\tilde{m}}{d\theta^h} > 0.$$

When expert robustness is higher, the effective financial constraint $\tilde{m}$ is smaller, such that a change in the expert s' wealth leads to a smaller change in α_t, i.e. a weaker sensitivity effect. However, when household robustness θ^h is larger, the effective financial constraint becomes looser, a change in the expert wealth results in a larger change in α_t. On the other hand, the robust concern from the household plays an opposite role. When θ^h is larger, the effective financial constraint becomes looser, a change in the expert wealth results in a larger change in α_t. This would cause a severer binding during crisis and accelerate the sensitivity effect especially during financial crisis.

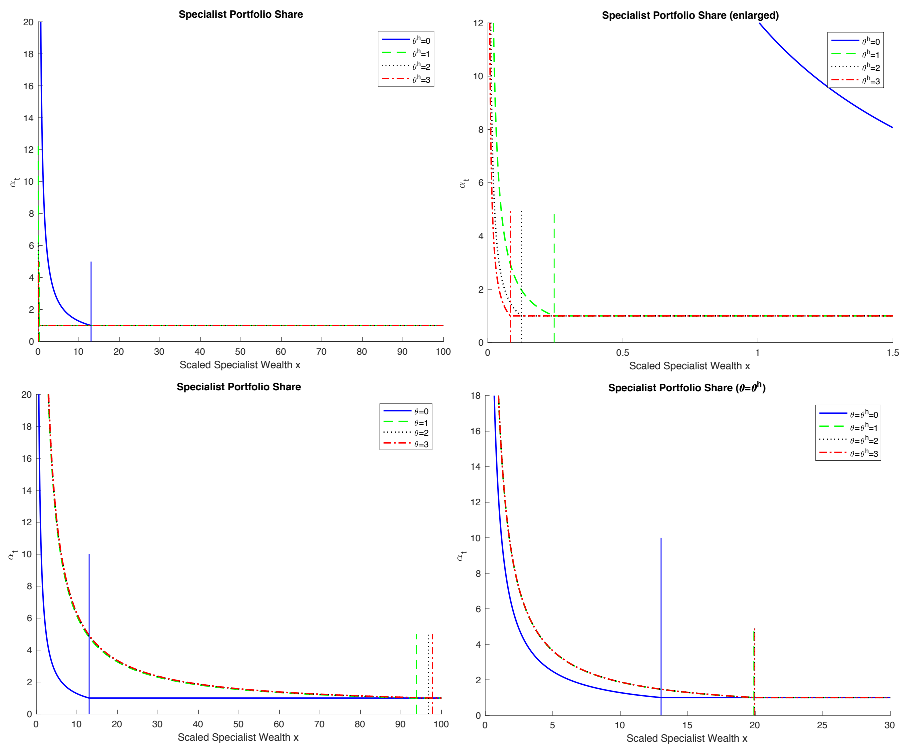

Figure 3:

The expert's portfolio choice for risky asset α_t is graphed against the expert wealth x_t for different robust parameters (θ and θ^h) varying from 0 to 3. The threshold value x^c (vertical line) separates the constrained (left) and unconstrained (right) region. The top panels shut down expert's robust concern ($\theta = 0$) and the top right panel is the enlarged version of top left panel. Bottom left panel shuts down household's robust concern ($\theta^h = 0$). Bottom right panel plots the homogeneous robust concern from two agents ($\theta = \theta^h = \bar{\theta}$).

4.2.3. *Risky Asset Volatility.*

Proposition 4.5. *In the unconstrained region,*

$$\sigma^U_{R,t} = \sigma\left(\frac{1}{1+\Delta\rho x_t}\right)\frac{\left(\rho^h\gamma^h - \rho\gamma\right)x_t + \gamma}{\rho\left(\gamma^h - \gamma\right)x_t + \gamma}. \tag{4.53}$$

In the constrained region,

$$\sigma_{R,t} = \sigma\left(\frac{\rho^h}{1+\Delta\rho x_t}\right)\left(\frac{1+m}{m\rho^h+\rho}\right). \tag{4.54}$$

Proof. The return volatility can be derived from matching the diffusion terms of equation (2.1), (2.2) and (4.44) that

$$\sigma_{R,t} = \frac{\sigma D_t}{\rho^h P_t - (\rho^h - \rho)\varepsilon_t^*} = \left(\frac{1}{P_t/D_t}\right)\frac{\sigma}{\rho^h - (\rho^h - \rho)\beta_t}. \tag{4.55}$$

Using Proposition 4.4 and 4.3, we have

$$\sigma_{R,t}^U = \frac{\sigma}{\rho^h}\left(\frac{1}{P_t/D_t}\right)\frac{\rho\left(\gamma^h - \gamma\right)x_t + \gamma}{\left(\rho^h\gamma^h - \rho\gamma\right)x_t + \gamma}. \tag{4.56}$$

$$\sigma_{R,t} = \left(\frac{1}{P_t/D_t}\right)\left(\frac{\sigma}{\rho^h - \frac{\rho^h - \rho}{1+m}}\right). \tag{4.57}$$

From equation (4.56) and (4.57), price/dividend ratio increases when x_t increases, thus the return volatilities decrease. In the constrained region, as x_t drops, the constraint tightens, thus return volatility rises only through price/dividend ratio. However, in the unconstrained region, decreasing in x_t not only increases $\sigma_{R,t}^U$ through price/dividend ratio from the first term in parentheses of equation (4.56), but also decreases $\sigma_{R,t}^U$ through the second term. Thus, the effect of x_t to $\sigma_{R,t}^U$ is ambiguous in the unconstrained case. When there is no ambiguity, i.e. $\theta = \theta^h = 0$, $\sigma_{R,t}^U = \sigma$ which is independent of x_t.

Lemma 4.6.

$$\frac{d\sigma_{R,t}^U}{d\theta} = \frac{-\sigma\left(\frac{1}{\rho} - x_t\right)\left[\rho\gamma\left(1-\rho^h\right)\left(\frac{1}{\rho} - x_t\right) + \rho^h\gamma^h\left(1-\rho\right)x_t\right]}{\rho^h\left(1 + \Delta\rho x_t\right)\left(\rho\Delta\gamma x_t + \gamma\right)^2} \leq 0.$$

$$\frac{d\sigma_{R,t}^U}{d\theta^h} = \frac{\sigma\left(\frac{1}{\rho} - x_t\right)\rho\Delta\rho\gamma x_t}{\rho^h\left(1 + \Delta\rho x_t\right)\left(\rho\Delta\gamma x_t + \gamma\right)^2} \geq 0.$$

$$\frac{d\sigma_{R,t}^U}{d\bar{\theta}} = -\frac{\sigma\left(\frac{1}{\rho} - x_t\right)\Delta\rho}{\rho^h\left(1 + \Delta\rho x_t\right)\left(\rho\Delta\gamma x_t + \gamma\right)^2} \leq 0$$

where $\Delta\gamma \equiv \gamma^h - \gamma$ denotes the dispersion in effective risk aversion. And,

$$\frac{d\sigma_{R,t}}{d\theta} = \frac{d\sigma_{R,t}}{d\theta^h} = \frac{d\sigma_{R,t}}{d\bar{\theta}} = 0.$$

Lemma 4.6 shows the opposite influence from two agents' ambiguities. From equation (4.55), the risky asset volatility come from two parts: price/dividend ratio and risk share contract. Price/dividend ratio is not a function of ambiguity under given wealth. In Proposition 4.3, heterogeneous agents' ambiguities have first order effect on β_t^U in the unconstrained region but no effect in constrained region, thus first order influence the risky asset volatility.

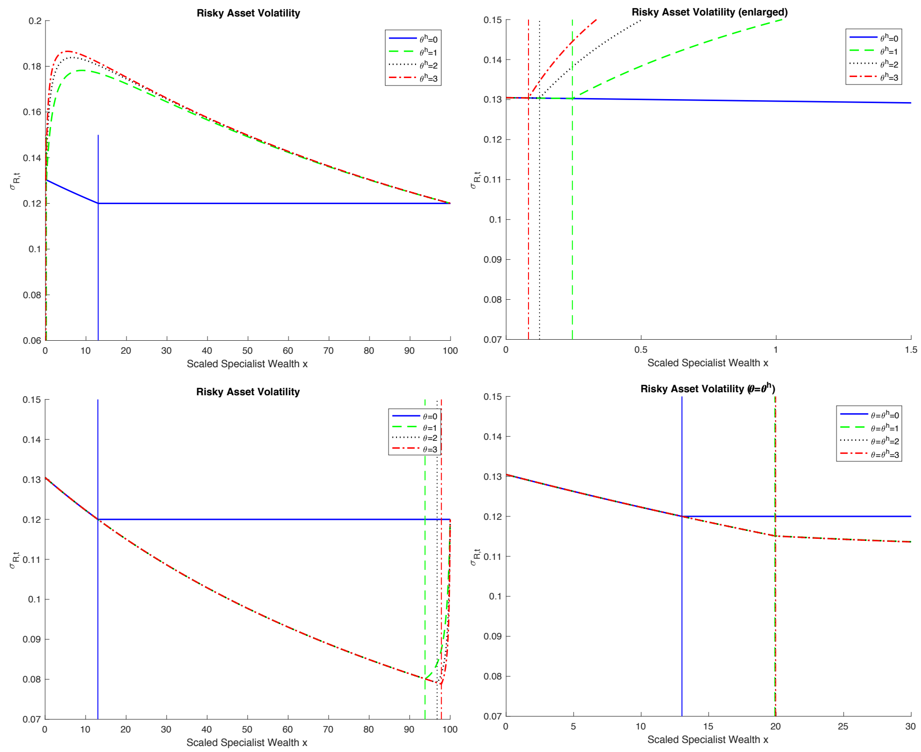

Figure 4:

The risky asset volatility $\sigma_{R,t}$ is graphed against the expert wealth x_t for different robust parameters (θ and θ^h) varying from 0 to 3. The threshold value x^c (vertical line) separates the constrained (left) and unconstrained (right) region. The top panels shut down expert's robust concern ($\theta = 0$) and the top right panel is the enlarged version of top left panel. Bottom left panel shuts down household's robust concern ($\theta^h = 0$). Bottom right panel plots the homogeneous robust concern from two agents ($\theta = \theta^h = \bar{\theta}$).

4.2.4. *Risk Premium and Financial Constraint.* The risk premium could be solved through optimal exposure supply by the expert (2.16),

$$\pi_{R,t} = \frac{\gamma \sigma_{R,t}^2 \varepsilon_t^*}{W_t}.$$

Thus, we have the following results.

Proposition 4.7. *In the unconstrained region,*

$$\pi_{R,t}^U = \frac{\sigma^2 \gamma \gamma^h}{(1 + \Delta \rho x_t)} \frac{\left[\left(\rho^h \gamma^h - \rho \gamma\right) x_t + \gamma\right]}{\left[\rho\left(\gamma^h - \gamma\right) x_t + \gamma\right]^2}. \tag{4.58}$$

In the constrained region,

$$\pi_{R,t} = \frac{\sigma^2 \rho^h \gamma}{x_t \left(1 + \Delta \rho x_t\right)} \frac{1 + m}{\left(m \rho^h + \rho\right)^2}. \tag{4.59}$$

Proof. See Appendix 6.2.1.

Lemma 4.8.

$$\frac{d\pi_{R,t}^{U}}{d\theta}=\frac{2\sigma^{2}\left(\gamma^{h}\right)^{2}\left[\rho^{h}\gamma^{h}+\gamma\left(\frac{1}{x_t}-\rho\right)\right]}{\rho^{3}x_t\left[1+\Delta\rho x_t\right]\left[\gamma^{h}-\gamma+\gamma\frac{1}{\rho x_t}\right]^{3}}>0.$$

$$\frac{d\pi_{R,t}^{U}}{d\theta^{h}}=\frac{\sigma^{2}\gamma^{2}\left(\frac{1}{x_t}-\rho\right)\left(\rho^{h}+\Delta\rho\beta_t^{U}\right)\left(\beta_t^{U}\right)^{2}}{\gamma^{h}x_t\left(1+\Delta\rho x_t\right)\left(\rho^{h}-\Delta\rho\beta_t^{U}\right)^{3}}\geq 0.$$

$$\frac{d\pi_{R,t}^{U}}{d\bar{\theta}}\geq 0.$$

$$\frac{d\pi_{R,t}}{d\theta}>0 \text{ and } \frac{d\pi_{R,t}}{d\theta^{h}}=0 \text{ and } \frac{d\pi_{R,t}}{d\bar{\theta}}>0.$$

Proof. See Appendix 6.3.1. It is interesting to notice that, θ positively changes the risk premium both in the unconstrained and constrained region, while θ^h also has a positive impact but *only* in the unconstrained region, as shown in Figure 5. The intuition is, whenever there is an increase in ambiguity aversion from two agents, both of them require higher risk premium in unconstrained region.[12] However, during financial crisis, the model predicts a first order effect of θ which reflects the major influence from marginal investor. The higher risk premium induces the expert who has high ambiguity aversion or low wealth to buy the exposure. Further, since $d\left[(1+m)\left(\frac{\sigma}{m\rho^h+\rho}\right)^2\right]/dm=-\frac{2\rho^h\sigma^2(1+m)}{\left(m\rho^h+\rho\right)^3}<0$, the risk premium is a decreasing function of $m only$ in the constrained region. When the capital constraint tightens, it will induce a higher risk premium.

4.2.5. *Market Price of Risk and Uncertainty.* The market price of risk is defined as the Sharpe ratio. Using Proposition 4.5 and 4.7 directly gets the following result.

Proposition 4.9. *In the unconstrained region, the market price of risk is*

$$\frac{\pi_{R,t}^{U}}{\sigma_{R,t}^{U}}=\frac{\sigma\gamma\gamma^{h}}{\rho\left(\gamma^{h}-\gamma\right)x_t+\gamma}. \tag{4.60}$$

In the constrained region,

$$\frac{\pi_{R,t}}{\sigma_{R,t}}=\sigma\left(\frac{\gamma}{m\rho^{h}+\rho}\right)\frac{1}{x_t}. \tag{4.61}$$

In the constrained region, only the expert robustness concern has first order effect on the sharpe ratio. This is consistent with the argument in intermediary asset pricing that marginal investors rather than households truly dominate the asset market. From equation (6.96), wealth growth and consumption growth are direct functions of sharpe ratio over

[12]This coincides with the literature where robustness can explain equity premium puzzle, e.g. Hansen, Sargent and Tallarini (1999), Chen and Epstein (2002), Maenhout (2004).

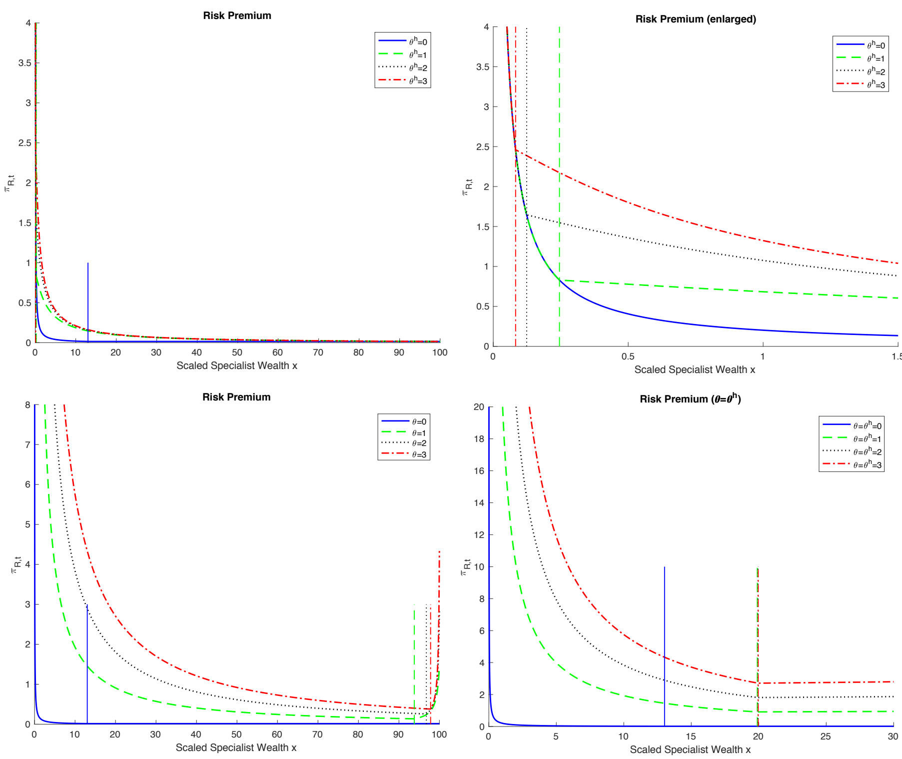

Figure 5:

The risk premium $\pi_{R,t}$ is graphed against the expert wealth x_t for different robust parameters (θ and θ^h) varying from 0 to 3. The threshold value x^c (vertical line) separates the constrained (left) and unconstrained (right) region. The top panels shut down expert's robust concern ($\theta = 0$) and the top right panel is the enlarged version of top left panel. Bottom left panel shuts down household's robust concern ($\theta^h = 0$). Bottom right panel plots the homogeneous robust concern from two agents ($\theta = \theta^h = \bar{\theta}$).

expert robustness parameter, which generate larger first and second order amplification effect from expert . Further,

$$\frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\theta} = \frac{\sigma\gamma^{h2}x_t}{\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^2} > 0$$

$$\frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\theta^h} = \frac{\sigma\gamma^2\left(1-\rho x_t\right)}{\rho^h\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^2} \geq 0.$$

$$\frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\bar{\theta}} = \frac{\sigma\left(1+\bar{\theta}/\rho\right)\left[\Delta\rho\left(\rho^h-\frac{\bar{\theta}^2}{\rho}\right)x_t+\rho^h\left(1+\bar{\theta}/\rho\right)^2\right]}{\left[-\Delta\rho\bar{\theta}x_t+\rho^h\left(1+\bar{\theta}/\rho\right)\right]^2} > 0.$$

Proof. See Appendix 6.98. Also, it is easy to see that,

$$\frac{d\left(\pi_{R,t}/\sigma_{R,t}\right)}{d\theta} > 0 \text{ and } \frac{d\left(\pi_{R,t}/\sigma_{R,t}\right)}{d\theta^h} = 0 \text{ and } \frac{d\left(\pi_{R,t}/\sigma_{R,t}\right)}{d\bar{\theta}} > 0.$$

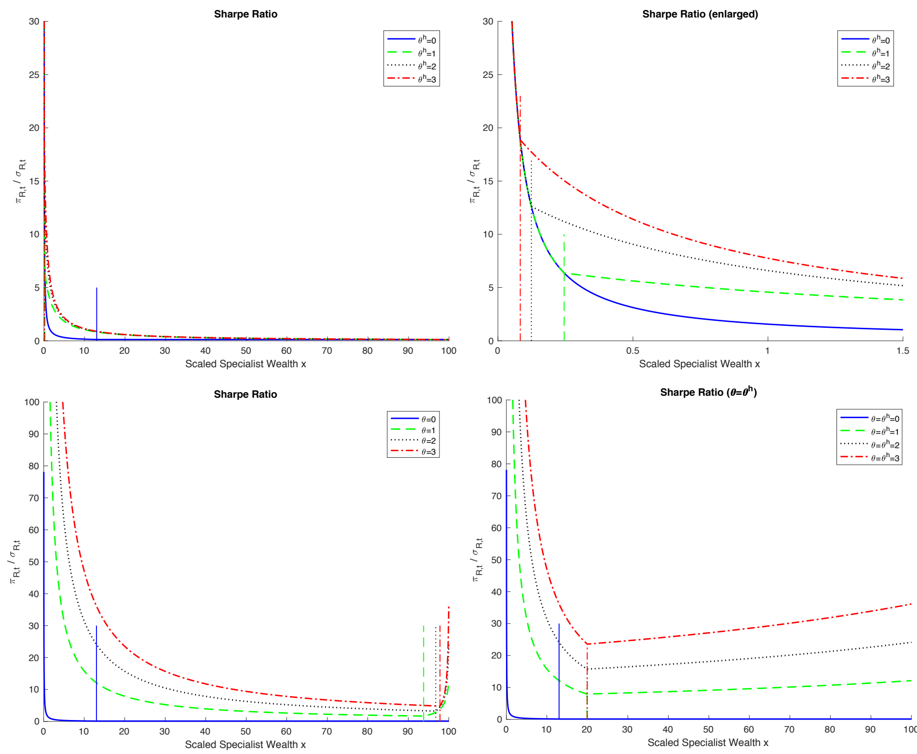

Figure 6:

The sharpe ratio $\pi_{R,t}/\sigma_{R,t}$ is graphed against the expert wealth x_t for different robust parameters (θ and θ^h) varying from 0 to 3. The threshold value x^c (vertical line) separates the constrained (left) and unconstrained (right) region. The top panels shut down expert's robust concern ($\theta = 0$) and the top right panel is the enlarged version of top left panel. Bottom left panel shuts down household's robust concern ($\theta^h = 0$). Bottom right panel plots the equal robust concern from two agents ($\theta = \theta^h = \bar{\theta}$).

4.2.6. *Exposure Price and Intermediation Fee.*

Proposition 4.10. *In the unconstrained region, the per-unit exposure price* $k_t^U = 0$ *and intermediation fee* $q_t^U = 0$. *In the constrained region,*

$$k_t = \frac{\sigma^2(1+m)}{(m\rho^h + \rho)^2}\left(\gamma - \frac{\rho^h\gamma^h m x_t}{1-\rho x_t}\right)\frac{\rho^h}{(1+\Delta\rho x_t)\,x_t} \tag{4.62}$$

$$q_t = \frac{\sigma^2 m}{(m\rho^h + \rho)^2}\left(\gamma - \frac{\rho^h\gamma^h m x_t}{1-\rho x_t}\right)\left(\frac{1}{x_t}\right)^2. \tag{4.63}$$

Proof. See Appendix 6.2.2.

$$\frac{dk_t}{d\theta} = \frac{\sigma^2(1+m)}{\rho\,(m\rho^h + \rho)^2}\frac{\rho^h}{(1+\Delta\rho x_t)\,x_t} > 0$$

$$\frac{dk_t}{d\theta^h} = -\frac{\sigma^2(1+m)}{(m\rho^h+\rho)^2}\frac{\gamma^h m x_t}{1-\rho x_t}\frac{\rho^h}{(1+\Delta\rho x_t)\, x_t} < 0$$
$$\frac{dk_t}{d\bar{\theta}} = \frac{\sigma^2(1+m)}{(m\rho^h+\rho)^2}\frac{\rho^h\left[1-\rho x_t\,(1+m)\right]}{(1+\Delta\rho x_t)\, x_t} > 0.$$

Proof. See Appendix 6.3.3. And,

$$\frac{dq_t}{d\theta} > 0 \text{ and } \frac{dq_t}{d\theta^h} < 0 \text{ and } \frac{dq_t}{d\bar{\theta}} > 0.$$

From Figure 7 and 8, the exposure price and intermediation fee change similar in terms of θ and θ^h. The exposure price clears the intermediation market where the exposure supply and demand are directly reduced by an increase in θ and θ^h, respectively. There are two reasons for it. First, During the crisis, an increased ambiguity induces the household a lower demand for risky asset and higher demand for the riskless bond. This directly reduces the exposure demand hence the per-unit exposure price drops in equilibrium.[13] Secondly, equation (6.95) implies $dk_t/d\tilde{m} < 0$, which indicates another channel that θ and θ^h change k_t through $\tilde{m}$. The intuition is, both higher ambiguity of households and lower ambiguity of expert s will make households trust the expertise of the expert more so that agency friction is relaxed, making it easier for expert to manage the intermediary thus charging a lower intermediation fee.

4.2.7. *Interest Rate.*

Proposition 4.11. *In the unconstrained region, the interest rate is*

$$r_t^U = \rho^h + g - \rho\Delta\rho x_t + \sigma^2\frac{\gamma\gamma^h - \left(\gamma^h+\gamma\right)\left[\rho x_t\left(\gamma^h-\gamma\right)+\gamma\right]}{\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^2}. \tag{4.64}$$

In the constrained region,

$$r_t = \rho^h + g - \rho\Delta\rho x_t - \sigma^2\frac{(1-\rho x_t)\left[\rho\,(1+\gamma m)+\rho^h m^2\gamma^h\right]+\rho^h m^2\left(\rho^h x_t-\gamma^h\right)}{(1-\rho x_t)\left(\rho+m\rho^h\right)^2 x_t}. \tag{4.65}$$

Proof. See Appendix 6.2.3. Moreover,

$$\frac{dr_t^U}{d\theta} = \frac{2\sigma^2 x_t\gamma^h\,(1-\rho x_t)}{\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^3}\left(\gamma^h-\gamma\right).$$

$$\Rightarrow \begin{cases} \frac{dr_t^U}{d\theta} < 0 & \text{if}\, \gamma^h < \gamma \\ \frac{dr_t^U}{d\theta} \geq 0 & \text{if}\, \gamma^h \geq \gamma. \end{cases}$$

$$\frac{dr_t^U}{d\theta^h} = -\frac{2\sigma^2\rho\gamma x_t\,(1-\rho x_t)\left(\gamma^h-\gamma\right)}{\rho^h\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^3}.$$

[13] The exposure supply also drops due to higher θ^h, but indirectly through risky asset price decline, which is smaller compared to the first order effect of θ^h on demand reduction.

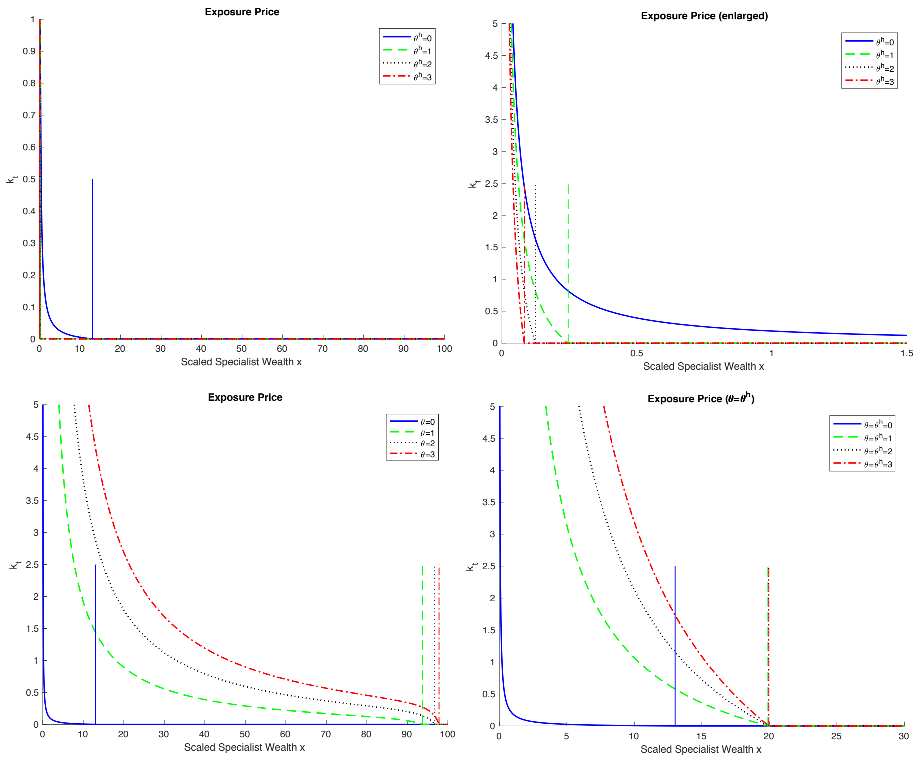

Figure 7:

The exposure price k_t is graphed against the expert wealth x_t for different robust parameters (θ and θ^h) varying from 0 to 3. The threshold value x^c (vertical line) separates the constrained (left) and unconstrained (right) region. The top panels shut down expert's robust concern ($\theta = 0$) and the top right panel is the enlarged version of top left panel. Bottom left panel shuts down household's robust concern ($\theta^h = 0$). Bottom right panel plots the equal robust concern from two agents ($\theta = \theta^h = \bar{\theta}$).

$$\Rightarrow \begin{cases} \frac{dr_t^U}{d\theta^h} > 0 & \text{if } \gamma^h < \gamma \\ \frac{dr_t^U}{d\theta^h} \leq 0 & \text{if } \gamma^h \geq \gamma. \end{cases}$$

$$\frac{dr_t^U}{d\bar{\theta}} = -\frac{2\sigma^2\bar{\theta}\rho^h \left(\varDelta\rho\right)^2 x_t \left(1-\rho x_t\right)}{\rho\left[-\varDelta\rho\bar{\theta}x_t + \rho^h\left(1+\bar{\theta}/\rho\right)\right]^3} \leq 0.$$

In the constrained case,

$$\frac{dr_t}{d\theta} = -\frac{\sigma^2 m}{\left(m\rho^h + \rho\right)^2 x_t} < 0.$$

$$\frac{dr_t}{d\theta^h} = \frac{\rho\sigma^2 m^2}{\left(m\rho^h + \rho\right)^2 \left(1-\rho x_t\right)} > 0.$$

$$\frac{dr_t}{d\bar{\theta}} = -\frac{\sigma^2 m \left[1-\rho x_t \left(1+m\right)\right]}{\left(m\rho^h + \rho\right)^2 \left(1-\rho x_t\right) x_t} < 0.$$

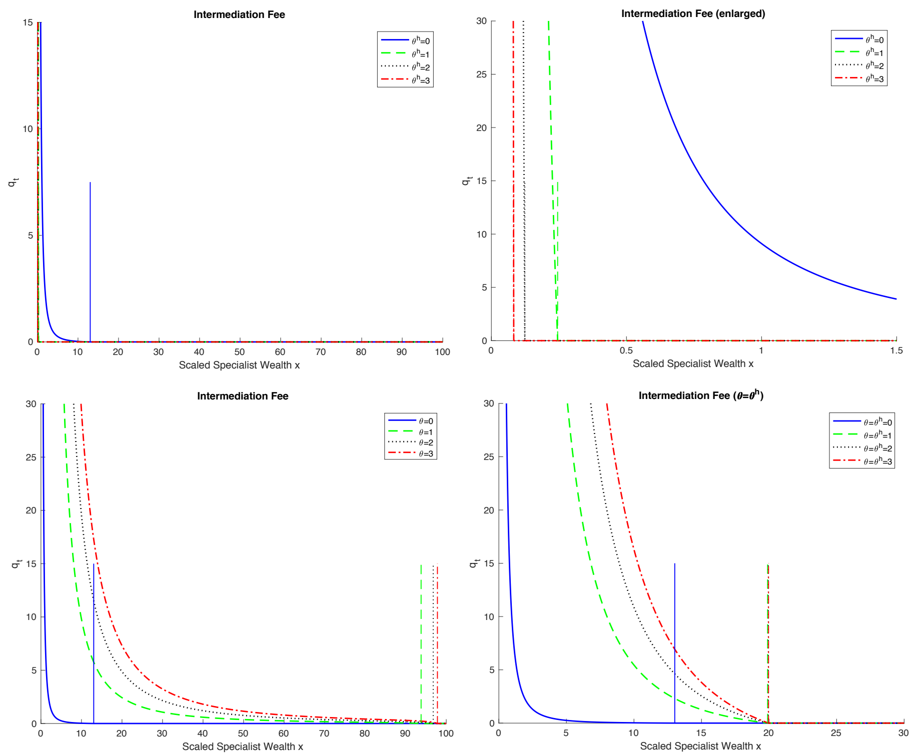

Figure 8:

The per-unit intermediation fee q_t is graphed against the expert wealth x_t for different robust parameters (θ and θ^h) varying from 0 to 3. The threshold value x^c (vertical line) separates the constrained (left) and unconstrained (right) region. The top panels shut down expert's robust concern ($\theta = 0$) and the top right panel is the enlarged version of top left panel. Bottom left panel shuts down household's robust concern ($\theta^h = 0$). Bottom right panel plots the equal robust concern from two agents ($\theta = \theta^h = \bar{\theta}$).

Proof: See Appendix 6.3.4.

4.3. **Observational Equivalence for Financial Friction.** There exists an observational equivalence (OE) for $\gamma^h = \gamma \Leftrightarrow \theta^h/\rho^h = \theta/\rho \Leftrightarrow \tilde{m} = m$. Following the assumption by HK(2012) that $\rho^h > \rho$, we must get $\theta^h > \theta$, i.e. households are more patient and more ambiguous than the expert . This is intuitive in the way that marginal investor has the expertise in the sense that he concerns less about the robustness. Moreover, insider information also makes them less ambiguous about the true state. Under OE, the effective financial constraint drops to the non-adjusted one, where the effect of heterogeneous ambiguity cancel out with each other in aggregate equilibrium. Thus, the constraint effect, which describes the channel of how heterogeneous ambiguities influence asset market through effective financial constraint, is neutralized. We could see from Figure 10, the portfolio share and risky asset volatility won't change with different ambiguity levels. However, the power of ambiguity to explain risk premium puzzle still remains both in

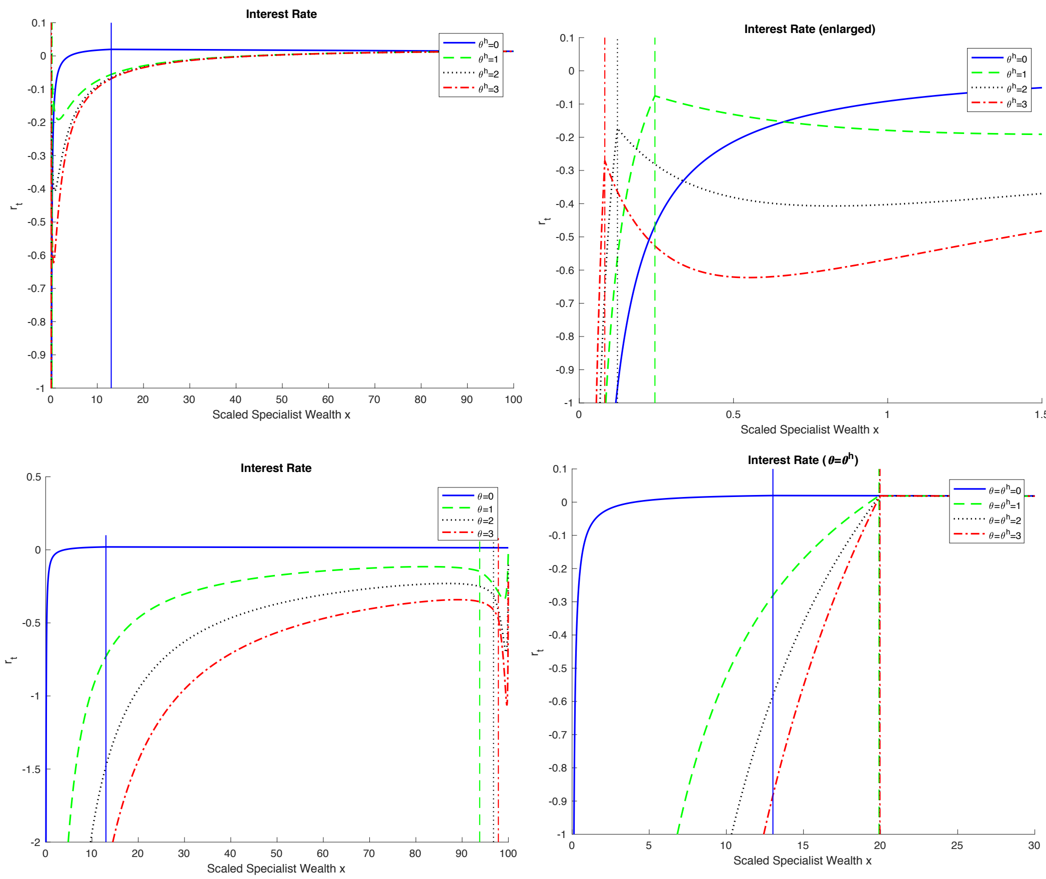

Figure 9:

Interest rate r_t is graphed against the expert wealth x_t for different robust parameters (θ and θ^h) varying from 0 to 3. The threshold value x^c (vertical line) separates the constrained (left) and unconstrained (right) region. The top panels shut down expert's robust concern ($\theta = 0$) and the top right panel is the enlarged version of top left panel. Bottom left panel shuts down household's robust concern ($\theta^h = 0$). Bottom right panel plots the equal robust concern from two agents ($\theta = \theta^h = \bar{\theta}$).

constrained and unconstrained case (middle panels), where risk premium and sharpe ratio are higher with higher ambiguity.

4.4. **The Wealth Distribution.**

Proposition 4.12. *The expert wealth process follows,*

$$\frac{dx_t}{x_t} = \mu_{x,t}dt + \sigma_{x,t}dZ_t.$$

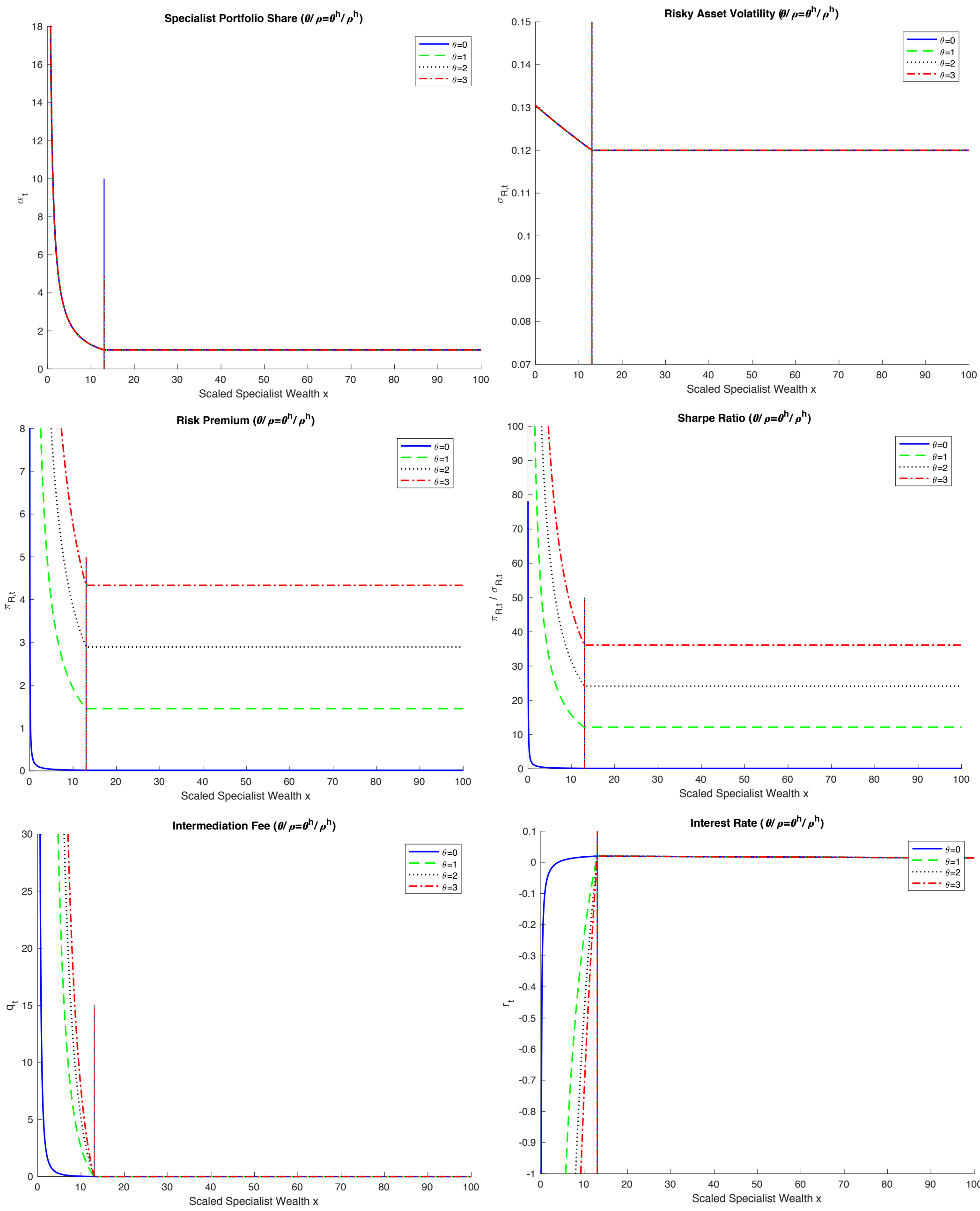

Figure 10:

Major asset pricing variables against the expert wealth x_t when equalizing effective financial constraint ($m = \tilde{m}$) under different robust parameters (θ and θ^h) varying from 0 to 3. The threshold value x^c (vertical line) separates the constrained (left) and unconstrained (right) region.

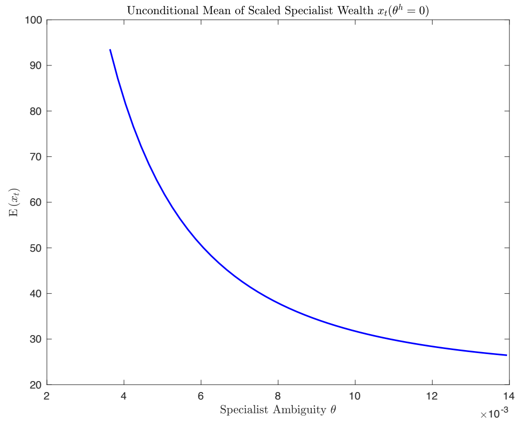

Figure 11:

Unconditional mean of expert ambiguity θ conditional on household ambiguity $\theta^h = 0$.

In unconstrained region,

$$\mu^U_{x,t} = \Delta\rho\,(1-\rho x_t) + \sigma^2 + \sigma^2 \frac{A^U_0 x_t^2 + A^U_1 x_t + A^U_2}{(1-\rho x_t)\left[\rho\left(\gamma^h - \gamma\right)x_t + \gamma\right]^2} \tag{4.66}$$

$$\begin{aligned} A^U_0 &= \rho^2\left(\gamma^h - \gamma\right)\left(2\gamma^h + \gamma\right) \\ A^U_1 &= \rho\left(2\gamma^2 - 3\gamma^{h2} + 2\gamma\gamma^h\right) \\ A^U_2 &= \gamma^{h2} - \gamma^2 - \gamma\gamma^h \end{aligned}$$

and

$$\sigma^U_{x,t} = \sigma\left[\frac{\gamma^h}{\rho\left(\gamma^h - \gamma\right)x_t + \gamma} - 1\right]. \tag{4.67}$$

In constrained region,

$$\mu_{x,t} = \Delta\rho\,(1-\rho x_t) + \sigma^2 + \sigma^2 \frac{A_0 x_t^2 + A_1 x_t + A_2}{\left(m\rho^h + \rho\right)^2 (1-\rho x_t)\, x_t^2} \tag{4.68}$$

$$\begin{aligned} A_0 &= \rho^h\left(\rho\gamma^h - \rho^h\right)m^2 + \rho\left(\rho\gamma + \rho^h\right)m + 2\rho^2 \\ A_1 &= -\rho^h\gamma^h m^2 - \left(\rho^h + 2\rho\gamma\right)m - 3\rho \\ A_2 &= \gamma m + 1 \end{aligned}$$

and

$$\sigma_{x,t} = \sigma\left[\frac{1}{\left(m\rho^h + \rho\right)x_t} - 1\right]. \tag{4.69}$$

Proof. See Appendix 6.4.

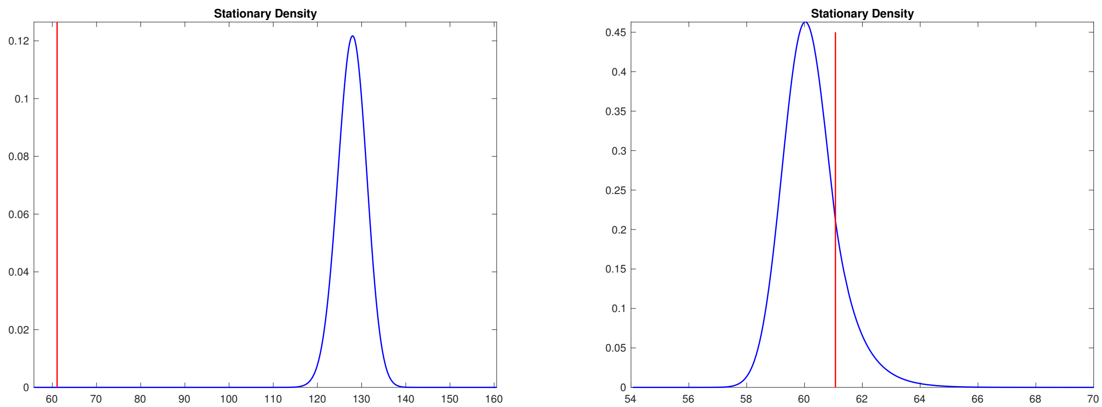

Figure 12: Theoretical Stationary Distribution

The left (right) figure is the stationary distribution of expert's scaled wealth when $\sigma = \sigma_L$ ($\sigma = \sigma_H$).

Figure 11 shows the unconditional mean of x_t with different expert ambiguity degree θ assuming no household ambiguity, i.e. $\mu_{x,t}(\theta) = 0$ conditional on $\mathbb{E}\left[x_t \mid \theta^h = 0\right]$. However, there is no solution of $\mu_{x,t}(\theta^h) = 0$ for $\mathbb{E}\left[x_t \mid \theta = 0\right]$.

Define scaled household wealth $x_t^h \equiv \frac{W_t^h}{D_t}$,

$$x_t^h = \frac{P_t - W_t}{D_t} = \frac{1 - \rho x_t}{\rho^h}. \tag{4.70}$$

We solve the stochastic process for the wealth process however highly non-linear, which prevents us from obtaining a close form solution of the density function. However, it enables us to do the simulation easily and solve other variables in terms of the state variable in close form. Figure 14 shows the stationary scaled household wealth distribution is fat-tailed or Pareto, which coincides with the fundamental results in empirical evidence (Gabaix, 2009) and theoretical results from dynamic heterogeneous agent models in studying wealth distributions (Benhabib, Bisin and Zhu, 2011, 2015). Furthermore, we consider the heterogeneity of ambiguity or robust concern as reflecting the expertise level. Expert as marginal investor has higher expertise thus lower ambiguity. The heterogeneity in expertise would lead to the stationary fat-tailed wealth distribution is analyzed by Eisfeldt, Lustig and Zhang (2017).

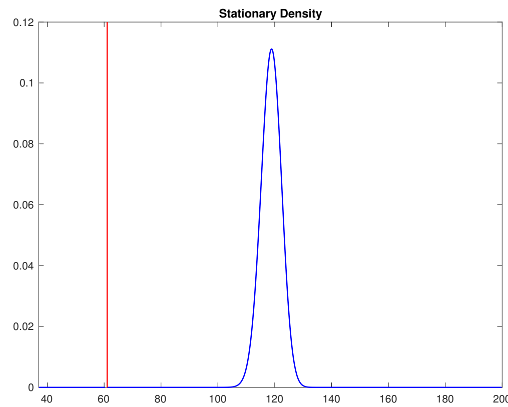

Figure 13: Theoretical Stationary Distribution

Stationary distribution of expert's scaled wealth when $\sigma = \bar{\sigma}$, where $\bar{\sigma} = 0.094$ is the unconditional mean of σ in the steady state.

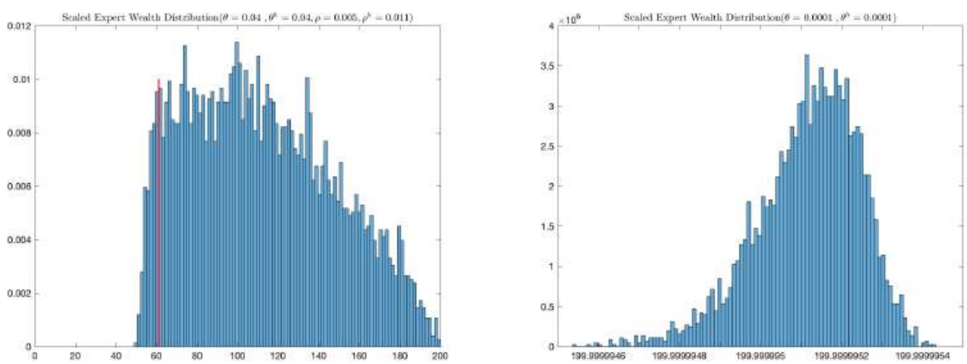

Figure 14: Simulated Expert Scaled Wealth Distribution

Histogram of expert scaled wealth distribution at the end period of simulation. The red line is the just binding constraint x^c. The left figure is the case when $\theta = \theta^h = 0.04$ and the right one is the benchmark model when $\theta = \theta^h = 0.0001$.

In HK(2012), a big problem is that the state variable process degenerates in steady state. Under observational equivalence $\gamma^h = \gamma$ in subsection 4.3, which includes HK(2012) where $\gamma = \gamma^h = 1$.

In the unconstrained region,

$$\mu^U_{x,t}\mid_{\gamma=\gamma^h=1} = \Delta\rho\left(1-\rho x_t\right) \text{ and } \sigma^U_{x,t}\mid_{\gamma=\gamma^h=1} = 0.$$

Thus, x_t satisfies a second-order linear differential equation $dx_t = \Delta\rho(1-\rho x_t)x_t dt$,

$$\begin{aligned}\int \frac{1}{\Delta\rho\left(1-\rho x_t\right)x_t}dx_t &= \int dt = t + const. \Leftrightarrow \frac{1}{\Delta\rho}\int\left(\frac{1}{x_t} - \frac{1}{x_t - \frac{1}{\rho}}\right)dx_t = t + const.\\ \Leftrightarrow Ce^{\Delta\rho t} &= \frac{x_t}{x_t - \frac{1}{\rho}}\end{aligned}$$

where $x^c \leq x_t < \frac{1}{\rho}$ and C is a constant. When $t = 0$, x_t drops into unconstrained region, $x_0 = x^c = \frac{1}{m\rho^h + \rho}$,

$$C = \frac{x^c}{x^x - \frac{1}{\rho}} = -\frac{\rho}{m\rho^h}.$$

Finally, we get

$$-\frac{\rho}{m\rho^h}e^{\Delta\rho t} = \frac{x_t}{x_t - \frac{1}{\rho}} \Leftrightarrow x_t = \frac{1}{m\rho^h e^{-\Delta\rho t} + \rho}.$$

It can be seen that under $\rho^h > \rho$, when $t \to \infty$,

$$\lim_{t\to\infty} x_t = \frac{1}{\rho}.$$

The top left panel in Figure 15 shows that, the convergence speed is exponentially fast until 1000 periods. HK(2012) corresponds to case ($\theta = 0$) on the top right panel. When x_t is close to zero, $\mu_{x,t}$ is sufficiently large and $\sigma_{x,t} \to \infty$ such that x_t jumps into the unconstrained region easily. Once inside the unconstrained region, x_t will converge to the steady state value $1/\rho$ and, with zero diffusion, never jump back to the constrained region again.

Decompose equation (4.66),

$$\mu^U_{x,t} = \underbrace{\Delta\rho\left(1-\rho x_t\right)}_{\mu^U_{x,t}|_{\gamma=\gamma^h=1}} + \underbrace{\sigma^2}_{\pi^U_{R,t}|_{\gamma=\gamma^h=1}} + \underbrace{\sigma^2\frac{A^U_0 x^2_t + A^U_1 x_t + A^U_2}{\left(1-\rho x_t\right)\left[\rho\rho^h\left(\gamma^h - \gamma\right)x_t + \rho^h\gamma\right]^2}}_{\text{subjective expected wealth (under } \mathbb{Q})}.$$

The third term is generated by ambiguity heterogeneity which is the mean driving force in scaled expert wealth evolution process. Under OE, In the constrained region, decompose equation (4.68),

$$\mu_{x,t} = \Delta\rho\left(1-\rho x_t\right) + \sigma^2 - \left(\sigma x^c\right)^2\left[\frac{\left(\rho^h m\right)^2}{1-\rho x_t} + \frac{m^2\rho^h + \left(2\rho+\rho^h\right)m + 3\rho}{x_t} - \frac{m+1}{x^2_t}\right].$$

Solve for inverse function $x_t(\theta) = x^c$ conditional on $\mu_{x,t} = 0$ under OE, we get $\theta_{x^c} = \frac{\rho^h\Delta\rho}{\sigma^2}$. When $\theta \leq \theta_{x^c}$, since $\Delta\rho$ is close to zero, θ_{x^c} is close to 0 (around 0.0078), it is more possible for x_t to drop into constrained region. Small θ corresponds to big $\sigma_{x,t}$ thus it would be highly possible to jump to infinity (see bottom panel of Figure 15). On the contrary, when ambiguity parameter $\theta > \theta_{x^c}$, it is less possible to drop into the constrained

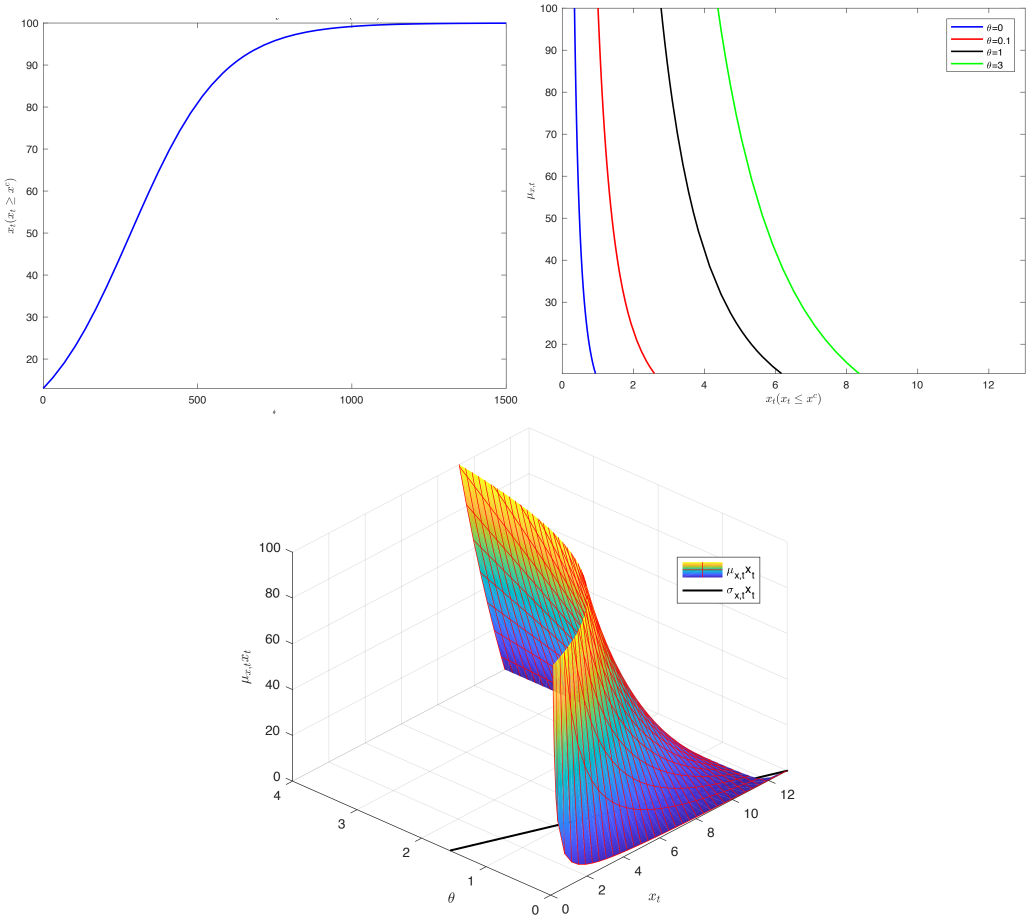

Figure 15:

Top left: The expert wealth x_t evolves over time in unconstrained region under HK(2012) and observational equivalence (OB, $\tilde{m} = m$). Top right: Unconditional mean of $\mu_{x,t}$ in constrained region over different θs under OB. Bottom: Unconditional mean and standard variance of scaled expert wealth distribution in constrained region under CE.

region. Conditional on $x_t \leq x^c$, the larger θ, the smaller $\sigma_{x,t}$, making the distribution non-degenerate and stationary. However, under OE, we cannot get the fat-tailed household wealth distribution. Thus, we show that the heterogeneity in effective risk aversion matters in order to fit the empirical evidence.

The wealth evolution for households is

$$\frac{dW_t^h}{W_t^h} - \frac{dW_t}{W_t} = -\frac{1}{1-\rho x_t}\left(\frac{dx_t}{x_t} + \frac{\sigma}{\gamma}\frac{\pi_{R,t}}{\sigma_{R,t}}dt - \sigma^2 dt\right).$$

In HK(2012) case, $\frac{\pi_{R,t}}{\sigma_{R,t}} = \sigma^2$, $\gamma = 1$, then $\frac{dW_t^h}{W_t^h} - \frac{dW_t}{W_t} = -\Delta\rho \leq 0$. Households will eventually become extinct and only experts will survive in the economy.

5. Quantitative Results

5.1. **Calibrating Robustness Parameter Using Detection-error Probabilities.**

Obtain the relative entropies from household and expert optimal robust problem,

$$g^h(x_t) = \sqrt{\mathcal{L}_t} = \theta^h \sigma_{W,t}^h V_w \tag{5.71}$$

$$g(x_t) = \sqrt{\mathcal{H}_t} = \theta \sigma_{W,t} J_w. \tag{5.72}$$

Finally we get

$$g^h(x_t) = \sigma\gamma\left(\gamma^h - 1\right)\left[\frac{m\rho^h}{\gamma(m\rho^h + \rho)(1-\rho x_t)}\mathbf{1}_{x_t \in (0,x^c]} + \frac{1}{\rho(\gamma^h - \gamma)x_t + \gamma}\mathbf{1}_{x_t \in (x^c, \frac{1}{\rho}]}\right]$$

$$g(x_t) = \sigma\gamma^h(\gamma - 1)\left[\frac{1}{\gamma^h(m\rho^h + \rho)x_t}\mathbf{1}_{x_t \in (0,x^c]} + \frac{1}{\rho(\gamma^h - \gamma)x_t + \gamma}\mathbf{1}_{x_t \in (x^c, \frac{1}{\rho}]}\right] \tag{5.73}$$

where $\mathbf{1}$ denotes the indicator function. See Appendix 6.6 for the proof. Follow the method of Maenhout (2006), define Radon-Nikodym derivative $\Xi_{1,t}^h \equiv \frac{dQ^h}{dP}\left(\Xi_{1,t} \equiv \frac{dQ}{dP}\right)$ as households (expert s) distorted model $Q^h(Q)$ with respect to approximating model P. Then log likelihoods for two agents are

$$\xi_{1,t}^h \equiv \log \Xi_{1,t}^h = -\int_0^t g^h(x_s)\, dZ_s - \frac{1}{2}\int_0^t \left\|g^h(x_s)\right\|^2 ds \tag{5.74}$$

$$\xi_{1,t} \equiv \log \Xi_{1,t} = -\int_0^t g(x_s)\, dZ_s - \frac{1}{2}\int_0^t \|g(x_s)\|^2 ds. \tag{5.75}$$

The log of Radon-Nikodym derivative $\Xi_{2,t}^h \equiv \frac{dP}{dQ^h}\left(\Xi_{2,t} \equiv \frac{dP}{dQ}\right)$ of the approximating model $\mathbb{P}$ with respect to households (expert s) distorted model $Q^h(Q)$ is

$$\xi_{2,t}^h \equiv \log \Xi_{2,t}^h = \int_0^t g^h(x_s)\, dZ_s + \frac{1}{2}\int_0^t \left\|g^h(x_s)\right\|^2 ds \tag{5.76}$$

$$\xi_{2,t} \equiv \log \Xi_{2,t} = \int_0^t g(x_s)\, dZ_s + \frac{1}{2}\int_0^t \|g(x_s)\|^2 ds. \tag{5.77}$$

When approximating model P generates the data, q_P measures the probability of the likelihood ratio of making detection errors in selecting model Q. Define

$$q_P^h = \Pr\left(\xi_{1,t}^h > 0 | P, \mathcal{F}_0\right) \tag{5.78}$$

$$q_P = \Pr\left(\xi_{1,t} > 0 | P, \mathcal{F}_0\right). \tag{5.79}$$

Similarly, when model Q generates the data,

$$q_Q^h = \Pr\left(\xi_{2,t}^h > 0 | Q^h, \mathcal{F}_0\right) \tag{5.80}$$

$$q_Q = \Pr\left(\xi_{2,t} > 0 | Q^h, \mathcal{F}_0\right). \tag{5.81}$$

Given the equal weight of prior probabilities over model P and Q, the conditional probability of the detection error for two agents over sample length N are

$$p^h\left(\theta^h; N\right) = \frac{1}{2}q_P^h + \frac{1}{2}q_Q^h \tag{5.82}$$

$$p(\theta; N) = \frac{1}{2}q_P + \frac{1}{2}q_Q. \tag{5.83}$$

5.2. **Measurements from Simulation.** In this section we evaluate the model's ability to match the empirical moments. The calibration proceeds in two steps. First, a subset of parameter values in Table 1 are set using standard values from the literature. Second, given these parameter values, the two ambiguity parameters are simultaneously determined by solving the model to target jointly five moments from the data.

In the first step, we adopt the parameter values from HK (2012). In the second step, we compare the theoretical moments of our model with heterogeneous ambiguity with the equivalent model under rational expectation. The baseline moments in the first column of Table 2 estimated from the data are taken from HK(2013). The second column is the approximating model without considering two agents' ambiguity degrees. Left columns are simulation results with calibrated ambiguity degrees.

[14]Indeed, all of the measurements display more severe dynamics after considering ambiguity aversion. The main points to notice in Table 2 are the following: (i) The calibration suggests that the expert has a very low level of ambiguity aversion. This is because the ambiguity indicates the heterogeneity in expertise. Expert has high expertise thus low ambiguity degree than the households. (ii) Expert ambiguity dominates the key moments of risk premium and Sharpe ratio. A small shift of the expert ambiguity degree from 0.01 to 0.02 would cause the risk premium and Sharpe ratio to jump twice values than before. (iii) Heterogeneity should be within a certain range, neither too low or too high. For $\theta = 0.01$, $\theta^h \in (0.1, 1]$ will result in reasonable moments. (iv) Heterogeneity is important while non-linear in determining the risk free rate level and volatility. That is the joint result of "flight to quality" wealth effect and expertise effect. The equilibrium interest rate decline depends on which effect dominates. (iv) household wealth distribution is fat-tailed and Pareto due to the heterogeneity in ambiguity.

After the expert's ambiguity shift from 0.1 to 0.2 conditional on invariant of household's ambiguity, column 5 of Table 2 shows the risk premium soar immediately from 36.30% to 59.33% (63% change). This is consistent with the fact documented in Muir (2017) that an increase in risk premia occurs right during or after the banking or financial panic. We interpret the panic as an increase in expert's robust preference or ambiguity.

[14]The simulation uses quarterly data over 10,000 sample paths and remain the last 33 years as non-burn-in periods. Numerically we prove that it is long enough to ensure the stationary distribution of state variable x_t. Also, x_t is independent of the initial values.

Table 1: Parameters and Targets

Panel A. Preferences		
ρ	Time discount rate of expert	0.5%
ρ^h	Time discount rate of household	1.136%
θ	Ambiguity attitude of expert	0.0413
θ^h	Ambiguity attitude of household	0.0413
Panel B. Intermediation		
m	intermediation multiplier	2
g	Dividend growth rate	2%
σ_H	Dividend volatility in the high state	9%
σ_H	Dividend volatility in the low state	15%
λ_H	Transition rate from high to low	0.7
λ_L	Transition rate from low to high	0.05

Table 2: Measurements

	Data	Model
θ	0.0001	0.04
θ^h	0.0001	0.04
γ	1.02	9.26
γ^h	1.01	4.63
Risk Premium (%)	0.92	5.29
Sharpe Ratio (%)	9.59	61.62
Interest Rate (%)	1.59	1.77
Interest Rate Volatility (%)	0.31	0.35
Return Volatility (%)	9.40	8.35
P/D Mean	200.00	150.40
P/D Volatility	0.00	3.99
Portfolio Share	1	1.0031
Expert Scaled Wealth Mean	200.00	111.44
Expert Scaled Wealth Volatility	0.00	7.12
Expert Detection Error Probability	0.25	0.29
Household Detection Error Probability	0.25	0.28
Probability of Sharpe Ratio Exceed Twice of the Mean (%)	0	0.32

This table reports the unconditional simulated results. We simulate 5000 years and 5000 sample paths with quarterly frequency. To match the data from 1970-2017, we report 47 years simulated results in stationary distribution.

6. Conclusion

TBA.

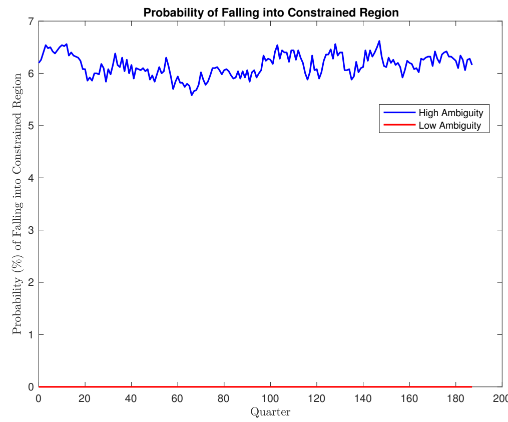

Figure 16: Probability of Constraint Binds

Time path for the probability of falling into constrained region.

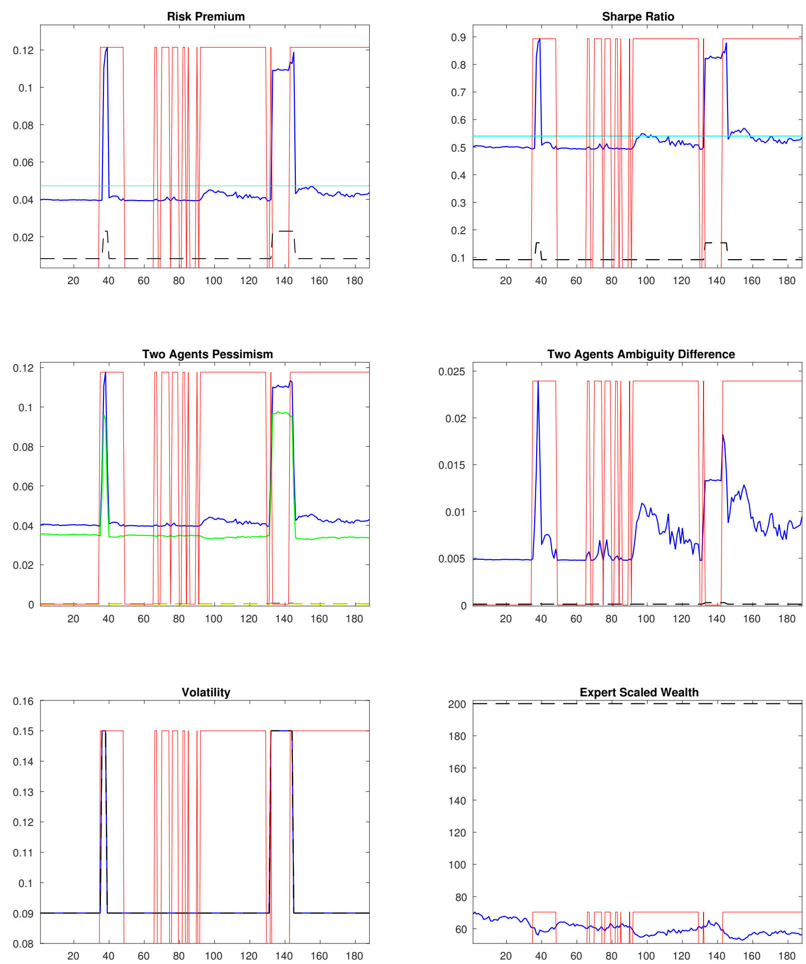

Figure 17:

All blue lines indicate the calibrated case with $\theta = \theta^h = 0.04$. The dotted black line are $\theta = \theta^h = 0.0001$, which represent almost no ambiguity. Red rectangular regions indicate the constrained region of the $\theta = \theta^h = 0.04$ case. The light blue horizontal lines in the first two figures are unconditional means for risk premium and Sharpe Ratio, respectively.

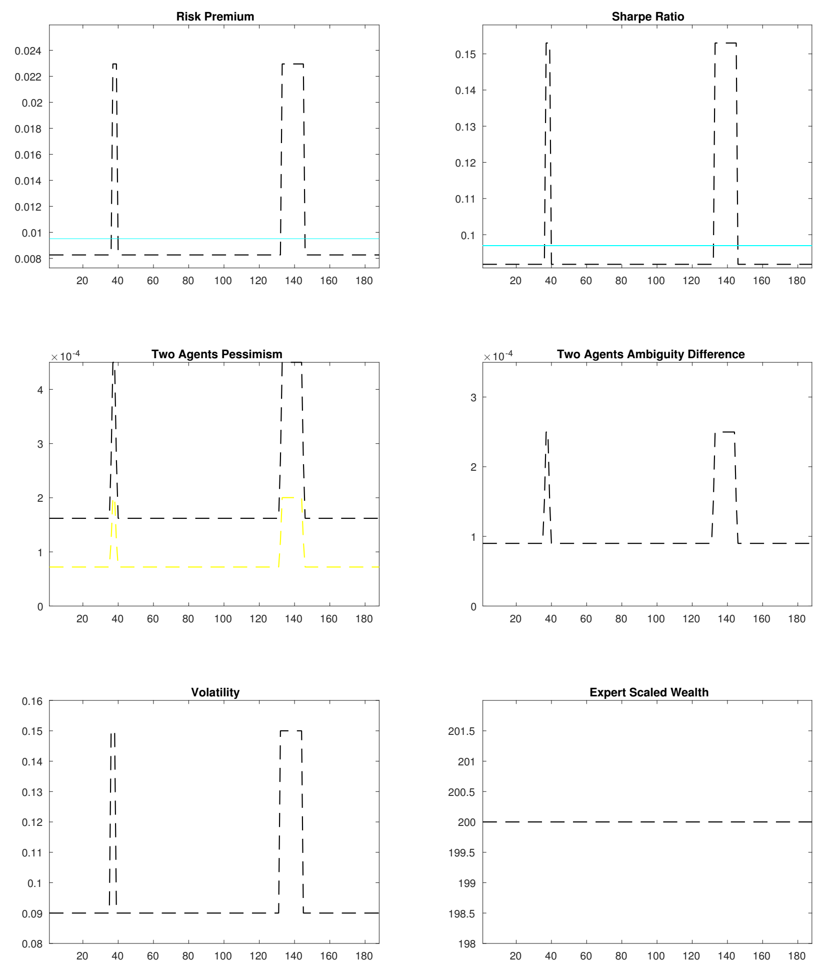

Figure 18:

This figure plots the enlarged graph for $\theta = \theta^h = 0.0001$, which represent almost no ambiguity. The light blue horizontal lines in the first two figures are unconditional means for risk premium and Sharpe Ratio, respectively.

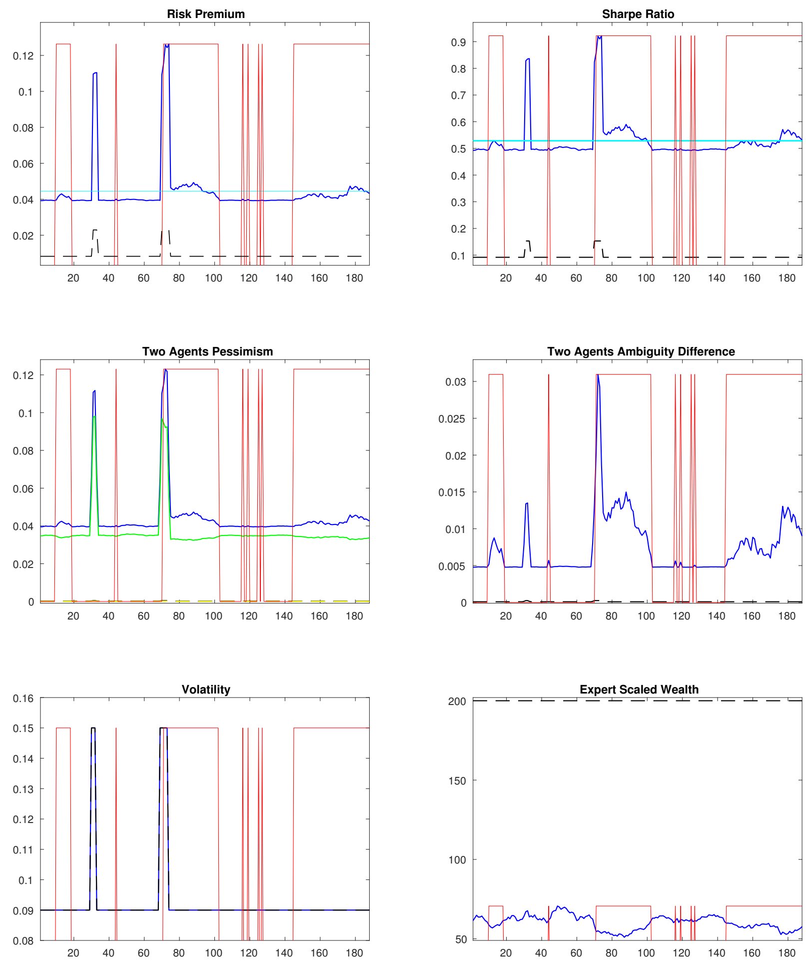

Figure 19:

All blue lines indicate the calibrated case with $\theta = \theta^h = 0.04$. The dotted black line are $\theta = \theta^h = 0.0001$, which represent almost no ambiguity. Red rectangular regions indicate the constrained region of the $\theta = \theta^h = 0.04$ case. The light blue horizontal lines in the first two figures are unconditional means for risk premium and Sharpe Ratio, respectively.

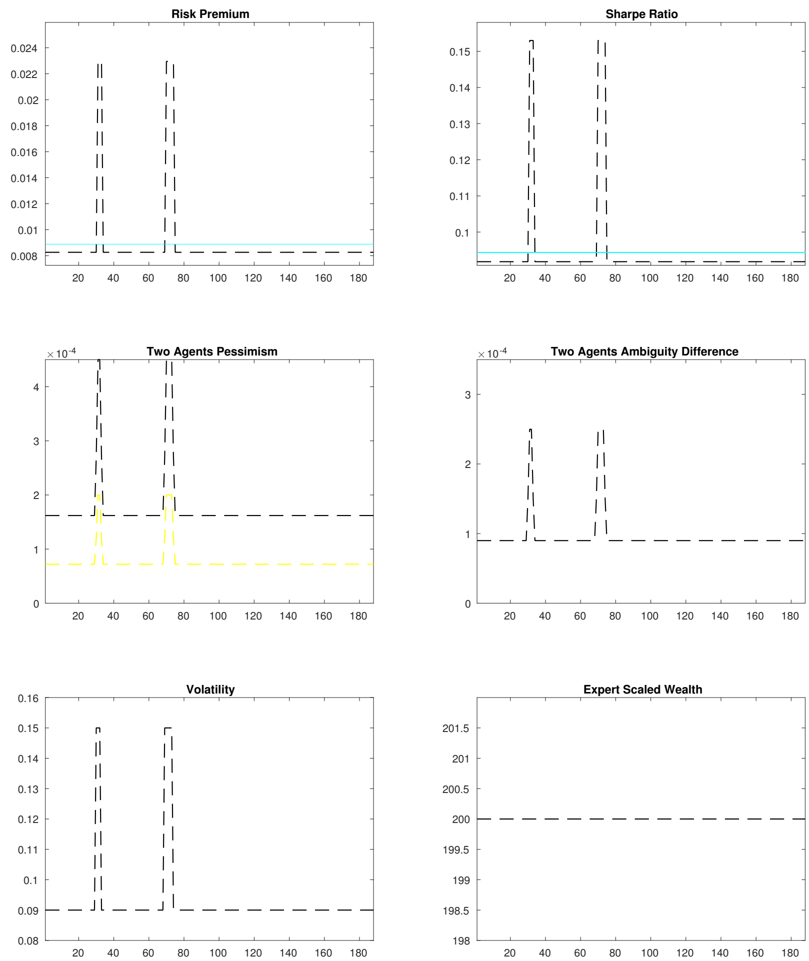

Figure 20:

This figure plots the enlarged graph for $\theta = \theta^h = 0.0001$, which represent almost no ambiguity. The light blue horizontal lines in the first two figures are unconditional means for risk premium and Sharpe Ratio, respectively.

References

[1] Adrian, T., Shin, H.S., 2014. Procyclical Leverage and Value-at-Risk. *The Review of Financial Studies* 27, 373-403

[2] Ahn, D., Choi, S., Gale, D., Kariv, S., 2014. Estimating ambiguity aversion in a portfolio choice experiment. *Quantitative Economics* 5, 195-223

[3] Anderson, E.W., Hansen, L.P., Sargent, T.J., 2003. A Quartet of Semigroups for Model Specification, Robustness, Prices of risk, and Model Detection. *Journal of the European Economic Association* 1, 68-123

[4] Antoniou, C., Harris, R.D.F., Zhang, R., 2015. Ambiguity aversion and stock market participation: An empirical analysis. *Journal of Banking and Finance* 58, 57-70

[5] Backus, D., Ferriere, A., Zin, S., 2015. Risk and ambiguity in models of business cycles. *Journal of Monetary Economics* 69, 42-63

[6] Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics* 115, 1-24

[7] Benhabib, J., Bisin, A., Zhu, S., 2011. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica* 79, 123-157

[8] Benhabib, J., Bisin, A., Zhu, S., 2015. The wealth distribution in Bewley economies with capital income risk. *Journal of Economic Theory* 159, 489-515

[9] Bhandari, A., Borovička, J., & Ho, P. (2016). Identifying Ambiguity Shocks in Business Cycle Models Using Survey Data (No. w22225). *National Bureau of Economic Research.*

[10] Bloom, N., 2014. Fluctuations in Uncertainty. *Journal of Economic Perspectives* 28, 153-76

[11] Borovička, J., 2016. Survival and long-run dynamics with heterogeneous beliefs under recursive preferences. *Job Market Paper*

[12] Bossaerts, P., Ghirardato, P., Guarnaschelli, S., Zame, W.R., 2010. Ambiguity in asset markets : theory and experiment. *Review of Financial Studies* 23, 1325-1359

[13] Boyarchenko, N., 2012. Ambiguity shifts and the 2007 –2008 financial crisis. *Journal of Monetary Economics* 59, 493-507

[14] Brenner, M., Izhakian, Y.Y., 2018. Asset pricing and ambiguity: Empirical evidence. *Journal of Financial Economics, forthcoming*

[15] Brunnermeier, M.K., Pedersen, L.H., 2009. Market Liquidity and Funding Liquidity. *The Review of Financial Studies* 22, 2201-2238

[16] Brunnermeier, M.K., Sannikov, Y., 2016. Macro, Money, and Finance: A Continuous-Time Approach. *Handbook of Macroeconomics*, pp. 1497-1545.

[17] Caballero, R.J., Krishnamurthy, A., 2008. Collective Risk Management in a Flight to Quality Episode. *Journal of Finance* 63, 2195-2230

[18] Caballero, R.J., Simsek, A., 2013. Fire Sales in a Model of Complexity. *Journal of Finance* 68, 2549-2587

[19] Cao, H.H., Wang, T., Zhang, H.H., 2005. Model Uncertainty, Limited Market Participation, and Asset Prices. *The Review of Financial Studies* 18, 1219-1251

[20] Chao, T., 2017. Ambiguity and Investment Decisions: An Empirical Analysis on Mutual Fund Investor Behaviour. *Academic Journal of Economic Studies* 3, 38-46

[21] Chen, Z., Epstein, L., 2002. Ambiguity, Risk, and Asset Returns in Continuous Time. *Econometrica* 70, 1403-1443

[22] Dimmock, S.G., Kouwenberg, R., Mitchell, O.S., Peijnenburg, K., 2016. Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics* 119, 559-577

[23] Driouchi, T., Trigeorgis, L., So, R.H.Y., 2015. Option implied ambiguity and its information content: Evidence from the subprime crisis. *Annals of Operations Research*

[24] Easley, D., O 'Hara, M., 2009. Ambiguity and Nonparticipation: The Role of Regulation. *The Review of Financial Studies* 22, 1817-1843

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

[26] Epstein, L.G., Wang, T., 1994. Intertemporal Asset Pricing under Knightian Uncertainty. *Econometrica* 62, 283-322
[27] Epstein, L.G., Schneider, M., 2007. Learning under ambiguity. *Review of Economic Studies* 74, 1275-1303
[28] Eisfeldt, A.L., Lustig, H., Zhang, L., 2017. Complex Asset Markets. *NBER working paper*
[29] Gabaix, X., 2009. Power Laws in Economics and Finance. *Annual Review of Economics* 1, 255-294
[30] Gilboa, I., Schmeidler, D., 1989. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18, 141-153
[31] Hansen, L.P., Sargent, T.J., 2001. Robust Control and Model Uncertainty. *American Economic Review* 91, 60-66
[32] Hansen, L.P., Sargent, T.J., 2005. Robust Estimation and Control under Commitment. *Journal of Economic Theory* 124, 258-301
[33] Hansen, L.P., Sargent, T.J., 2007. Recursive Robust Estimation and Control without Commitment. *Journal of Economic Theory* 136, 1-27
[34] Hansen, L.P., Sargent, T.J., 2010. Fragile Beliefs and the Price of Uncertainty. *Quantitative Economics* 1, 129-162
[35] Hansen, L.P., Sargent, T.J., 2011. Robustness and Ambiguity in Continuous Time. *Journal of Economic Theory* 146, 1195-1223
[36] Hansen, L.P., Sargent, T.J., Tallarini, T.D., 1999. Robust Permanent Income and Pricing. *Review of Economic Studies* 66, 873-907
[37] He, Z., Krishnamurthy, A., 2012. A Model of Capital and Crises. *Review of Economic Studies* 79, 735-777
[38] He, Z., Krishnamurthy, A., 2013. Intermediary Asset Pricing. *American Economic Review* 103, 732-770
[39] He, Z., Kelly, B., Manela, A., 2017. Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics* 126, 1-35
[40] Ju, N., Miao, J., 2012. Ambiguity, Learning, and Asset Returns. *Econometrica* 80, 559-591
[41] Kasa, K., Lei, X., 2018. Risk, uncertainty, and the dynamics of inequality. *Journal of Monetary Economics* 94, 60-78
[42] Kocher, M.G., Lahno, A.M., Trautmann, S.T., 2018. Ambiguity aversion is not universal. *European Economic Review* 101, 268-283
[43] Li, K., 2016. Asset Pricing with a Financial Sector. *Working paper, HKUST.*
[44] Luo, Y., 2016. Robustly Strategic Consumption –Portfolio Rules with Informational Frictions. *Management Science* 63, 4158-4174
[45] Luo, Y., Young, E., 2016. Induced uncertainty, market price of risk, and the dynamics of consumption and wealth. *Journal of Economic Theory* 163, 1
[46] Maenhout, P.J., 2004. Robust Portfolio Rules and Asset Pricing. *The Review of Financial Studies* 17, 951-983
[47] Maenhout, P.J., 2006. Robust Portfolio Rules and Detection-error Probabilities for a Mean-reverting Risk Premium. *Journal of Economic Theory* 128, 136-163
[48] Malmendier, U., Nagel, S., 2011. Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *Quarterly Journal of Economics* 126, 373
[49] Miao, J., Rivera, A., 2016. Robust Contracts in Continuous Time. *Econometrica* 84, 1405-1440
[50] Tallarini, T.D., 2000. Risk-sensitive Real Business Cycles. *Journal of Monetary Economics* 45, 507-532

Appendix

0.1. **Solving the Optimal Choices of Households and Experts under Robust Concern.** *Proof of Proposition* **??**. Optimal household consumption and portfolio rule under robustness are

$$C_t^h = \frac{1}{V_w} \tag{0.1}$$

$$\varepsilon_t^h = \frac{-V_w}{V_{ww} - \theta^h V_w^2} \frac{(\pi_{R,t} - k_t)}{\sigma_{R,t}^2}. \tag{0.2}$$

Put (0.1) and (0.2) back into (**??**) gives

$$\begin{aligned} 0 &= -\ln V_w - \rho^h V + V_w \left[\varepsilon_t^h(\pi_{R,t} - k_t) + r_t W_t^h - \frac{1}{V_w}\right] + \frac{1}{2} V_{ww} (\varepsilon_t^h)^2 \sigma_{R,t}^2 + \mu_{Y,t}^h - \frac{\theta^h}{2} (\sigma_{R,t} \varepsilon_t^h V_w)^2 \\ &= -\ln V_w - \rho^h V - 1 + V_w \left[\varepsilon_t^h(\pi_{R,t} - k_t) + r_t W_t^h\right] + \frac{1}{2}\left(V_{ww} - \theta^h V_w^2\right)(\varepsilon_t^h)^2 \sigma_{R,t}^2 + \mu_{Y,t}^h. \end{aligned} \tag{0.3}$$

Guess value function takes the form

$$V(W_t^h; Y_t^h) = \frac{1}{\rho^h} \ln W_t^h + Y_t^h \tag{0.4}$$

where Y_t^h is a function of aggregate state variable x_t. Now define Y_t^h andY_t as a function of x_t,

$$dY^h(x_t) = \mu_{Y,t}^h dt + \sigma_{Y,t}^h dZ_t$$
$$dY(x_t) = \mu_{Y,t} dt + \sigma_{Y,t} dZ_t.$$

Using Ito's formula,

$$\mu_{Y,t}^h = Y^{h\prime}(x_t)\mu_{x,t}x_t + \frac{1}{2} Y^{h\prime\prime}(x_t)\sigma_{x,t}^2 x_t^2 \tag{0.5}$$

$$\sigma_{Y,t}^h = Y^{h\prime}(x_t)\sigma_{x,t}x_t \tag{0.6}$$

$$\mu_{Y,t} = Y\prime(x_t)\mu_{x,t}x_t + \frac{1}{2} Y\prime\prime(x_t)\sigma_{x,t}^2 x_t^2 \tag{0.7}$$

$$\sigma_{Y,t} = Y\prime(x_t)\sigma_{x,t}x_t. \tag{0.8}$$

Under this conjecture, $V_w = \frac{1}{\rho^h W_t^h}$ and $V_{ww} = -\frac{1}{\rho^h (W_t^h)^2}$. Substituting these conjectures into FOCs (0.1) and (0.2),

$$C_t^h = \rho^h W_t^h$$
$$\varepsilon_t^h = \frac{\pi_{R,t} - k_t}{\gamma^h \sigma_{R,t}^2} W_t^h.$$

Put all those expressions back to (0.3) yields (for simplicity, I dropped the time script),

$$\begin{aligned} 0 &= \ln \rho^h W^h - \ln W^h - \rho^h Y^h - 1 + \frac{1}{\rho^h W^h}\left[\frac{(\pi_R - k)^2}{\gamma^h \sigma_R^2} W^h + r W^h\right] \\ &\quad -\frac{1}{2}\left(\frac{1}{\rho^h (W^h)^2} + \frac{\theta^h}{(\rho^h W^h)^2}\right) \frac{W^{h2}}{\gamma^{h2}} \frac{(\pi_R - k)^2}{\sigma_R^4} \sigma_R^2 + \mu_Y^h \\ 0 &= \ln \rho^h - \rho^h Y^h - 1 + \frac{(\pi_R - k)^2}{\rho^h \gamma^h \sigma_R^2} + \frac{r}{\rho^h} - \frac{1}{2\rho^h \gamma^h} \frac{(\pi_R - k)^2}{\sigma_R^2} + \mu_Y^h. \end{aligned}$$

Substituting equation (0.5) yields

$$Y_t^{h\prime}\mu_{x,t}x_t + \frac{1}{2}Y_t^{h\prime\prime}\sigma_{x,t}^2 x_t^2 = \rho^h Y_t^h - \ln\rho^h + 1 - \frac{(\pi_{R,t}-k_t)^2}{2\rho^h\gamma^h\sigma_{R,t}^2} - \frac{r_t}{\rho^h}. \quad (0.9)$$

From propositions **??????????**, Y_t^h could be derived in terms of x_t and satisfies the above second-order ODE. Thus, optimal robust consumption rule for the household is

$$C_t^{h*} = \rho^h W_t^h$$

and the robust optimal risk exposure is

$$\varepsilon_t^{h*} = \frac{\pi_{R,t}-k_t}{\gamma^h\sigma_{R,t}^2}W_t^h.$$

Proof of Proposition **??**. Optimal expert consumption and portfolio rules under RB are

$$C_t = \frac{1}{J_w}$$

$$\varepsilon_t = \frac{-J_w}{J_{ww}-\theta J_w^2}\frac{\pi_{R,t}}{\sigma_{R,t}^2}$$

where $\frac{-J_w}{J_{ww}-\theta J_w^2}$ is the adjustment for risk aversion of mean-variance coefficient. Conjecture the value function for expert takes the form

$$J(W_t;Y_t) = \frac{1}{\rho}\ln W_t + Y_t. \quad (0.10)$$

So that $J_w = \frac{1}{\rho W_t}$ and $J_{ww} = -\frac{1}{\rho W_t^2}$. Put those expressions into FOCs and back into (**??**),

$$0 = \ln\rho W - \ln W - \rho Y - 1 + \frac{1}{\rho W}\left[\frac{W\pi_R^2}{\gamma\sigma_R^2} + (q+r)W\right] - \frac{1}{2}\left(\frac{1}{\rho W^2} + \frac{\theta}{\rho^2 W^2}\right)\frac{W^2}{\gamma^2}\frac{\pi_R^2}{\sigma_R^2} + \mu_Y.$$

Substituting equation (0.7) yields

$$Y_t\prime\mu_{x,t}x_t + \frac{1}{2}Y_t\prime\prime\sigma_{x,t}^2x_t^2 = \rho Y_t - \ln\rho + 1 - \frac{q_t+r_t}{\rho} - \frac{\pi_{R,t}^2}{2\rho\gamma\sigma_{R,t}^2}. \quad (0.11)$$

From propositions **??????????**, Y_t could be derived in terms of x_t and satisfies the above second-order ODE. Thus, optimal robust consumption rule for the expert is

$$C_t^* = \rho W_t$$

and the robust optimal risk exposure is

$$\varepsilon_t^* = \frac{\pi_{R,t}}{\gamma\sigma_{R,t}^2}W_t.$$

0.2. **Solving for Asset Prices.**

0.2.1. *Solving the risk premium.*

$$\pi_{R,t} = \frac{\gamma\sigma_{R,t}^2\varepsilon_t^*}{W_t} = \frac{\gamma\sigma_{R,t}^2\beta_t P_t}{W_t} = \frac{\gamma\sigma_{R,t}^2\beta_t(P_t/D_t)}{x_t}.$$

In the unconstrained region,

$$\pi_{R,t}^U = \frac{\gamma\sigma_{R,t}^{U2}\beta_t^U(P_t/D_t)}{x_t}$$

$$\Rightarrow \pi_{R,t}^U = \frac{\sigma^2\gamma\gamma^h}{(1+\Delta\rho x_t)}\frac{\left[\left(\rho^h\gamma^h-\rho\gamma\right)x_t+\gamma\right]}{\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^2}.$$

In the constrained region,

$$\pi_{R,t} = \frac{\gamma\sigma_{R,t}^2\beta_t(P_t/D_t)}{x_t}$$

$$\Rightarrow \pi_{R,t} = \frac{\sigma^2\gamma\rho^h}{x_t\left(1+\Delta\rho x_t\right)}\frac{1+m}{\left(m\rho^h+\rho\right)^2}.$$

0.2.2. *Solving the exposure price and intermediation fee.* In the constrained region, $k_t \geq 0$. When household desired exposure demand (**??**) equals expert exposure supply (**??**), we have

$$\varepsilon_t^{h*}(k_t) = m\varepsilon_t^* \Leftrightarrow \frac{\pi_{R,t}-k_t}{\gamma^h\sigma_{R,t}^2}W_t^h = m\frac{\pi_{R,t}}{\gamma\sigma_{R,t}^2}W_t$$

$$\Rightarrow k_t = \left(1-\frac{\tilde{m}\rho^h x_t}{1-\rho x_t}\right)\pi_{R,t} = \left(1-\frac{\gamma^h}{\gamma}\frac{m\rho^h x_t}{1-\rho x_t}\right)\pi_{R,t} \tag{0.12}$$

$$\Rightarrow k_t = \frac{\sigma^2(1+m)}{\left(m\rho^h+\rho\right)^2}\left(\gamma-\frac{\rho^h\gamma^h m x_t}{1-\rho x_t}\right)\frac{\rho^h}{\left(1+\Delta\rho x_t\right)x_t}.$$

And then

$$q_t \equiv \frac{K_t}{W_t} = \frac{mk_t\varepsilon_t^*}{W_t} = \frac{mk_t}{\gamma}\frac{P_t}{\beta_t} = \frac{m}{(1+m)}\frac{P_t/D_t}{x_t}k_t$$

$$\Rightarrow q_t = \frac{\sigma^2 m}{\left(m\rho^h+\rho\right)^2}\left(\gamma-\frac{\rho^h\gamma^h m x_t}{1-\rho x_t}\right)\left(\frac{1}{x_t}\right)^2.$$

In order to make sure $k_t \geq 0$, x_t should satisfy the condition

$$\left(1-\frac{\tilde{m}\rho^h x_t}{1-\rho x_t}\right)\pi_{R,t} \geq 0 \Rightarrow \frac{(\tilde{m}\rho^h+\rho)x_t-1}{1-\rho x_t} \leq 0$$

$$\Rightarrow x_t \leq \frac{1}{\tilde{m}\rho^h+\rho} = x_c.$$

Since k_t is positive only when $x_t \leq x_c$, thus the condition satisfies automatically.

0.2.3. *Solving the risk free rate.* From household's Euler equation under distorted model,

$$r_t dt = \rho^h dt + \mathbb{E}_t\left[\frac{dC_t^{h*}}{C_t^{h*}}\right] - \text{var}_t\left[\frac{dC_t^{h*}}{C_t^{h*}}\right].$$

$$\frac{dC_t^{h*}}{C_t^{h*}} = \frac{d\left(\rho^h W_t^h\right)}{\rho^h W_t^h} = \frac{d\left(P_t - W_t\right)}{P_t - W_t}.$$

$$dW_t = \left(\varepsilon_t \pi_{R,t} + (q_t + r_t)W_t - C_t\right)dt + \sigma_{R,t}\varepsilon_t\left(\sigma_{R,t}\varepsilon_t\nu_t dt + dZ_t\right)$$

$$\begin{aligned}\Rightarrow \frac{dW_t}{W_t} &= \left[\frac{1}{\gamma}\left(1 - \frac{\gamma - 1}{\gamma}\right)\frac{\pi_{R,t}^2}{\sigma_{R,t}^2} + q_t + r_t - \rho\right]dt + \frac{\pi_{R,t}}{\gamma\sigma_{R,t}}dZ_t \\ &= \phi_{W,t}dt + r_t dt + \frac{\pi_{R,t}}{\gamma\sigma_{R,t}}dZ_t \qquad (0.13)\end{aligned}$$

where $\phi_{W,t} = \left(\frac{\pi_{R,t}}{\gamma\sigma_{R,t}}\right)^2 + q_t - \rho$.

$$\frac{d\left(P_t - W_t\right)}{P_t - W_t} = \frac{(gdt + \sigma dZ_t)D_t - \rho dW_t}{D_t - \rho W_t} = \frac{(gdt + \sigma dZ_t) - \rho\frac{dW_t}{W_t}x_t}{1 - \rho x_t} \qquad (0.14)$$

$$\begin{aligned}\Rightarrow \mathbb{E}_t\left[\frac{d(P_t - W_t)}{P_t - W_t}\right] &= \frac{g - \rho x_t\left(\phi_{W,t} + r_t\right)}{1 - \rho x_t}dt \\ \text{var}_t\left[\frac{d(P_t - W_t)}{P_t - W_t}\right] &= \left(\frac{\sigma - \frac{\rho x_t}{\gamma}\frac{\pi_{R,t}}{\sigma_{R,t}}}{1 - \rho x_t}\right)^2 dt\end{aligned}$$

$$r_t = \rho^h + \frac{g - \rho x_t\left(\phi_{W,t} + r_t\right)}{1 - \rho x_t} - \left(\frac{\sigma - \frac{\rho x_t}{\gamma}\frac{\pi_{R,t}}{\sigma_{R,t}}}{1 - \rho x_t}\right)^2$$

$$\Rightarrow r_t = \rho^h + g - \rho\left(\rho^h - \rho\right)x_t - \rho q_t x_t - \frac{\rho x_t\left(\left[\frac{\pi_{R,t}}{\gamma\sigma_{R,t}}\right]^2 - 2\sigma\frac{\pi_{R,t}}{\gamma\sigma_{R,t}}\right) + \sigma^2}{1 - \rho x_t}.$$

Using the expressions for $\pi_{R,t}/\sigma_{R,t}$ and q_t in constrained and unconstrained regions by propositions **??** and **??**,

$$r_t^U = \rho^h + g - \rho\Delta\rho x_t + \sigma^2\frac{\gamma\gamma^h - \left(\gamma^h + \gamma\right)\left(\rho x_t\left(\gamma^h - \gamma\right) + \gamma\right)}{\left[\rho\left(\gamma^h - \gamma\right)x_t + \gamma\right]^2}.$$

$$r_t = \rho^h + g - \rho\Delta\rho x_t - \sigma^2\frac{(1 - \rho x_t)\left[\rho\left(1 + \gamma m\right) + \rho^h m^2\gamma^h\right] + \rho^h m^2\left(\rho^h x_t - \gamma^h\right)}{\left(1 - \rho x_t\right)\left(\rho + m\rho^h\right)^2 x_t}.$$

From equation (0.14),

$$\frac{dW_t^h}{W_t^h} = \frac{d(P_t - W_t)}{P_t - W_t} = \frac{(gdt + \sigma dZ_t) - \rho\frac{dW_t}{W_t}x_t}{1 - \rho x_t}$$

$$\begin{aligned}
\Leftrightarrow \frac{dW_t^h}{W_t^h} &= \left(1-\frac{1}{1-\rho x_t}\right)\frac{dW_t}{W_t}+\frac{1}{1-\rho x_t}\frac{dD_t}{D_t}\\
\frac{dW_t^h}{W_t^h}-\frac{dW_t}{W_t} &= -\frac{1}{1-\rho x_t}\left(\frac{dx_t}{x_t}+\frac{\sigma}{\gamma}\frac{\pi_{R,t}}{\sigma_{R,t}}dt-\sigma^2 dt\right).
\end{aligned}$$

0.3. **Comparative Analysis.**

0.3.1. *Risk premium.*

$$\pi_{R,t}^U=\frac{\sigma^2(1+\theta/\rho)\left(1+\theta^h/\rho^h\right)}{\rho^2 x_t\left[1+(\rho^h-\rho)x_t\right]}\frac{\left[\left(\rho^h-\rho\right)+\left(\theta^h-\theta\right)+(1+\theta/\rho)\frac{1}{x_t}\right]}{\left[(1+\theta^h/\rho^h)-(1+\theta/\rho)+(1+\theta/\rho)\frac{1}{\rho x_t}\right]^2}$$

$$\Rightarrow \frac{d\pi_{R,t}^U}{d\theta} = \frac{2\sigma^2\left(\gamma^h\right)^2\left[\rho^h+\theta^h+\gamma\left(\frac{1}{x_t}-\rho\right)\right]}{\rho^3 x_t\left[1+(\rho^h-\rho)x_t\right]\left[\gamma^h-\gamma+\gamma\frac{1}{\rho x_t}\right]^3}.$$

If $\gamma^h>\gamma\geq 1$,

$$x_t>0>-\frac{\gamma}{\rho\left(\gamma^h-\gamma\right)}\Leftrightarrow \gamma^h-\gamma+\gamma\frac{1}{\rho x_t}>0.$$

If $\gamma>\gamma^h\geq 1$,

$$\frac{\gamma}{\gamma-\gamma^h}>1\Leftrightarrow\frac{\gamma}{\rho\left(\gamma-\gamma^h\right)}>\frac{1}{\rho}\geq x_t\Leftrightarrow \gamma^h-\gamma+\gamma\frac{1}{\rho x_t}>0.$$

If $\gamma^h=\gamma\geq 1$,

$$\gamma^h-\gamma+\gamma\frac{1}{\rho x_t} = \gamma\frac{1}{\rho x_t}>0.$$

$$\Rightarrow \frac{d\pi_{R,t}^U}{d\theta}=\frac{2\sigma^2\left(\gamma^h\right)^2\left[\rho^h+\theta^h+\gamma\left(\frac{1}{x_t}-\rho\right)\right]}{\rho^3 x_t\left[1+\Delta\rho x_t\right]\left[\gamma^h-\gamma+\gamma\frac{1}{\rho x_t}\right]^3}>0.$$

$$\pi_{R,t}^U=\frac{1}{P/D}\frac{\sigma^2\left(1+\theta/\rho\right)\beta_t^U}{x_t\left[\rho^h-\left(\rho^h-\rho\right)\beta_t^U\right]^2}$$

$$\begin{aligned}
\Rightarrow \frac{d\pi_{R,t}^U}{d\theta^h} &= \frac{1}{P/D}\frac{\sigma^2\left(1+\theta/\rho\right)}{x_t\left[\rho^h-\left(\rho^h-\rho\right)\beta_t^U\right]^3}\left[\frac{d\beta_t^U}{d\theta^h}\left(\rho^h-\left(\rho^h-\rho\right)\beta_t^U\right)+2\left(\rho^h-\rho\right)\frac{d\beta_t^U}{d\theta^h}\beta_t^U\right]\\
&= \frac{\rho^h}{\left[1+\left(\rho^h-\rho\right)x_t\right]x_t}\frac{\sigma^2\left(1+\theta/\rho\right)^2\left(\frac{1}{x_t}-\rho\right)\left(\rho^h+\left(\rho^h-\rho\right)\beta_t^U\right)\left(\beta_t^U\right)^2}{\left(\rho^h-\left(\rho^h-\rho\right)\beta_t^U\right)^3\rho^h\left(1+\theta^h/\rho^h\right)}.
\end{aligned}$$

$$0\leq\beta_t^U\leq 1\Rightarrow\rho\leq\rho^h-\left(\rho^h-\rho\right)\beta_t^U\leq\rho^h$$

$$\Rightarrow \frac{d\pi_{R,t}^U}{d\theta^h}=\frac{\sigma^2\gamma^2\left(\frac{1}{x_t}-\rho\right)\left(\rho^h+\Delta\rho\beta_t^U\right)\left(\beta_t^U\right)^2}{\gamma^h x_t\left(1+\Delta\rho x_t\right)\left(\rho^h-\Delta\rho\beta_t^U\right)^3}\geq 0.$$

0.3.2. *Sharpe ratio.*

$$\frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\theta} = \frac{\sigma\left(\rho^h+\theta^h\right)^2 x_t}{\left[\left(\rho\theta^h-\rho^h\theta\right)x_t+\rho^h\left(1+\theta/\rho\right)\right]^2}.$$

$$\Rightarrow \frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\theta} = \frac{\sigma\gamma^{h2}x_t}{\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^2} > 0.$$

$$\frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\theta^h} = \frac{\sigma\left(1+\theta/\rho\right)\rho^h\left(\theta+\rho\right)\left(\frac{1}{\rho}-x_t\right)}{\left[\left(\rho\theta^h-\rho^h\theta\right)x_t+\rho^h\left(1+\theta/\rho\right)\right]^2}$$

$$\Rightarrow \frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\theta^h} = \frac{\sigma\gamma^2\left(1-\rho x_t\right)}{\rho^h\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^2} \geq 0.$$

$$\frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\bar{\theta}} = \frac{\sigma\left(1+\bar{\theta}/\rho\right)\left[\Delta\rho\left(\rho^h-\frac{\bar{\theta}^2}{\rho}\right)x_t+\rho^h\left(1+\bar{\theta}/\rho\right)^2\right]}{\left[-\Delta\rho\bar{\theta}x_t+\rho^h\left(1+\bar{\theta}/\rho\right)\right]^2}.$$

When $\Delta\rho\left(\rho^h-\frac{\bar{\theta}^2}{\rho}\right)x_t+\rho^h\left(1+\bar{\theta}/\rho\right)^2>0$, $\frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\bar{\theta}}>0$. Hence,

$$\Delta\rho\left(\rho^h-\frac{\bar{\theta}^2}{\rho}\right)x_t+\rho^h\left(1+\bar{\theta}/\rho\right)^2>0$$

$$\begin{array}{lll} \rho^h-\frac{\bar{\theta}^2}{\rho}=0 & & \Leftrightarrow \bar{\theta}=\sqrt{\rho\rho^h} \\ \Leftrightarrow \{ \quad \begin{array}{l} \rho^h-\frac{\bar{\theta}^2}{\rho}<0 \\ \{ \quad x_t<-\frac{\rho^h\left(1+\bar{\theta}/\rho\right)^2}{\Delta\rho\left(\rho^h-\frac{\bar{\theta}^2}{\rho}\right)} \end{array} & \begin{array}{l} \Leftrightarrow \bar{\theta}>\sqrt{\rho\rho^h} \\ =\phi \end{array} & \end{array}$$

Take first derivative of ϕ with respect to $\bar{\theta}$,

$$\frac{d\phi}{d\bar{\theta}} = \frac{-2\rho^h\left(1+\bar{\theta}/\rho\right)}{\Delta\rho\left(\rho\rho^h-\bar{\theta}^2\right)^2}\left(\rho^h-\bar{\theta}-2\frac{\bar{\theta}^2}{\rho}\right).$$

$$\rho^h-\bar{\theta}-2\frac{\bar{\theta}^2}{\rho}=0 \Leftrightarrow 2\bar{\theta}^2+\rho\bar{\theta}-\rho\rho^h=0$$

$$\Rightarrow \bar{\theta}=\frac{-\rho\pm\sqrt{\rho^2+8\rho\rho^h}}{4}.$$

Since $\bar{\theta}\geq 0$, the negative root is not valid. Now show $\bar{\theta}=\frac{-\rho+\sqrt{\rho^2+8\rho\rho^h}}{4}<\sqrt{\rho\rho^h}$,

$$\frac{\bar{\theta}}{\sqrt{\rho\rho^h}} = \frac{-\sqrt{\frac{\rho}{\rho^h}}+\sqrt{\frac{\rho}{\rho^h}+8}}{4} < \frac{3}{4} < 1.$$

where we have used $0<\frac{\rho}{\rho^h}\leq 1$.

$$\Rightarrow \rho^h-\bar{\theta}-2\frac{\bar{\theta}^2}{\rho}<0\,\mathrm{for}\,\bar{\theta}\sqrt{\rho\rho^h}.$$

Thus,

$$\frac{d\phi}{d\theta} > 0.$$

Now show that $\phi > 1/\rho$,

$$-\frac{\rho^h \left(1+\bar{\theta}/\rho\right)^2}{\Delta\rho\left(\rho^h - \frac{\bar{\theta}^2}{\rho}\right)} > \frac{1}{\rho} \quad \Leftrightarrow \quad -\rho\rho^h\left(1+\frac{2\bar{\theta}}{\rho}+\frac{\bar{\theta}^2}{\rho}\right) > \rho^{h2} - \frac{\rho^h}{\rho}\bar{\theta}^2 - \rho\rho^h + \bar{\theta}^2$$

$$\Leftrightarrow \quad \left(\bar{\theta}+\rho^h\right)^2 > 0$$

$$\Rightarrow x_t < \frac{1}{\rho} < \phi.$$

Finally,

$$\Rightarrow \frac{d\left(\pi^U_{R,t}/\sigma^U_{R,t}\right)}{d\bar{\theta}} > 0.$$

0.3.3. *Exposure price.*

$$\frac{dk_t}{d\bar{\theta}} = \frac{\sigma^2(1+m)}{(m\rho^h+\rho)^2}\frac{\rho^h\left(1-\rho x_t - \rho m x_t\right)}{\left[1+(\rho^h-\rho)x_t\right]x_t}.$$

Since

$$\bar{x}^c = \frac{1}{\tilde{m}\rho^h+\rho} = \frac{1}{\rho}\left(\frac{1}{1+\frac{\rho^h+\bar{\theta}}{\rho+\bar{\theta}}m}\right) \leq \frac{1}{\rho}\left(\frac{1}{1+m}\right)$$

$$\Rightarrow x_t \leq \bar{x}^c \leq \frac{1}{\rho}\left(\frac{1}{1+m}\right)$$

$$\Leftrightarrow 1-\rho x_t\left(1+m\right) > 0. \tag{0.15}$$

$$\frac{dk_t}{d\bar{\theta}} = \frac{\sigma^2(1+m)}{(m\rho^h+\rho)^2}\frac{\rho^h\left[1-\rho x_t\left(1+m\right)\right]}{\left[1+(\rho^h-\rho)x_t\right]x_t} > 0.$$

0.3.4. *Interest rate.* Denote sharpe ratio $\frac{\pi_{R,t}}{\sigma_{R,t}} = sp$,

$$r_t = \rho^h + g - \rho\left(\rho^h-\rho\right)x_t - \rho q_t x_t - \frac{\rho x_t\left[\left(\frac{1}{\gamma}sp\right)^2 - 2\sigma\frac{1}{\gamma}sp\right]+\sigma^2}{1-\rho x_t}$$

$$\Rightarrow \frac{dr_t}{d\theta} = -\rho x_t\frac{dq_t}{d\theta} - \frac{2x_t}{\left(1-\rho x_t\right)\gamma^2}\left[\left(\frac{sp}{\gamma}-\sigma\right)\left(-sp+\frac{dsp}{d\theta}\rho\gamma\right)\right] \tag{0.16}$$

$$\Leftrightarrow \frac{dr^U_t}{d\theta} = \frac{2\sigma^2 x_t\gamma^h\left(1-\rho x_t\right)}{\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^3}\left(\gamma^h-\gamma\right).$$

$$\Rightarrow \begin{cases} \frac{dr^U_t}{d\theta} < 0 & \text{if } \gamma^h < \gamma \\ \frac{dr^U_t}{d\theta} \geq 0 & \text{if } \gamma^h \geq \gamma. \end{cases}$$

$$\frac{dr_t}{d\theta^h} = -\rho x_t\frac{dq_t}{d\theta^h} - \frac{2\rho x_t}{\left(1-\rho x_t\right)\gamma}\left(\frac{sp}{\gamma}\frac{dsp}{d\theta^h} - \sigma\frac{dsp}{d\theta^h}\right) \tag{0.17}$$

$$\frac{dr_t^U}{d\theta^h} = -\frac{2\sigma^2\rho\gamma x_t\left(1-\rho x_t\right)\left(\gamma^h-\gamma\right)}{\rho^h\left[\rho\left(\gamma^h-\gamma\right)x_t+\gamma\right]^3}.$$

$$\Rightarrow \left\{ \begin{array}{ll} \frac{dr_t^U}{d\theta^h} > 0 & \text{if } \gamma^h < \gamma \\ \frac{dr_t^U}{d\theta^h} \leq 0 & \text{if } \gamma^h \geq \gamma. \end{array} \right.$$

$$\frac{dr_t^U}{d\bar{\theta}} = \frac{2\sigma^2\bar{\theta}\Delta\rho x_t\left[\Delta\rho\left(\rho^h\bar{\theta}+\rho\rho^h\right)x_t-\Delta\rho\rho^h\left(1+\bar{\theta}/\rho\right)\right]}{\rho\left(1+\bar{\theta}/\rho\right)\left[-\Delta\rho\bar{\theta}x_t+\rho^h\left(1+\bar{\theta}/\rho\right)\right]^3}$$

$$\Rightarrow \frac{dr_t^U}{d\bar{\theta}} = -\frac{2\sigma^2\bar{\theta}\rho^h\left(\Delta\rho\right)^2 x_t\left(1-\rho x_t\right)}{\rho\left[-\Delta\rho\bar{\theta}x_t+\rho^h\left(1+\bar{\theta}/\rho\right)\right]^3} \leq 0.$$

In constrained case, from equations (0.16) and (0.17),

$$\frac{dr_t}{d\theta} = -\frac{\sigma^2 m}{\left(m\rho^h+\rho\right)^2 x_t} - \frac{2x_t}{\left(1-\rho x_t\right)\gamma^2}\left[\left(\frac{\sigma}{\left(m\rho^h+\rho\right)x_t}-\sigma\right)\left(-\frac{\sigma\gamma}{\left(m\rho^h+\rho\right)x_t}+\frac{\sigma\left(\rho+\theta\right)}{\rho\left(m\rho^h+\rho\right)x_t}\right)\right]$$

$$\frac{dr_t}{d\theta} = -\frac{\sigma^2 m}{\left(m\rho^h+\rho\right)^2 x_t} < 0.$$

$$\frac{dr_t}{d\theta^h} = \frac{\rho\sigma^2 m^2}{\left(m\rho^h+\rho\right)^2\left(1-\rho x_t\right)} > 0.$$

$$\frac{dr_t}{d\bar{\theta}} = -\frac{\sigma^2 m\left[1-\rho x_t\left(1+m\right)\right]}{\left(m\rho^h+\rho\right)^2\left(1-\rho x_t\right)x_t} < 0$$

where we used the condition (0.15).

0.4. **Solving the Stochastic Process of Aggregate State.** In order to derive the unconditional mean and variance of risk premium and interest rate, we need to know the distribution of the state variable x_t. Using Ito's formula,

$$dx_t = d\left(\frac{W_t}{D_t}\right) = \frac{dW_t}{D_t} - \frac{W_t dD_t}{D_t^2} - \frac{dW_t dD_t}{D_t^2} + \frac{W_t}{D_t^3}\left[dD_t, dD_t\right]$$

$$\Rightarrow \frac{dx_t}{x_t} = \frac{dW_t}{W_t} - \frac{dD_t}{D_t} - \frac{dD_t}{D_t}\frac{dW_t}{W_t} + \left(\frac{dD_t}{D_t}\right)^2.$$

Plug in the assumed dividend process (**??**) and derived expert wealth process (0.13),

$$\frac{dx_t}{x_t} = \frac{dW_t}{W_t} - \frac{dD_t}{D_t} - \frac{\sigma}{\gamma}\frac{\pi_{R,t}}{\sigma_{R,t}}dt + \sigma^2 dt$$

$$\frac{dx_t}{x_t} = \left(\sigma^2 - g - \rho + q_t + r_t + \frac{1}{\gamma^2}\frac{\pi_{R,t}^2}{\sigma_{R,t}^2} - \frac{\sigma}{\gamma}\frac{\pi_{R,t}}{\sigma_{R,t}}\right)dt + \left(\frac{\pi_{R,t}}{\gamma\sigma_{R,t}} - \sigma\right)dZ_t.$$

Denote $E=\frac{\pi_{R,t}}{\gamma\sigma_{R,t}}$, drift and diffusion of aggregate state process as $\mu_{x,t}$ and $\sigma_{x,t}$, respectively.

$$\begin{aligned}\mu_{x,t} &= \sigma^2-g-\rho+q_t+r_t+E^2-\sigma E\\ &= \sigma^2+\rho^h-\rho-\rho\left(\rho^h-\rho\right)x_t+\left(1-\rho x_t\right)q_t+\frac{E^2\left(1-2\rho x_t\right)+\left(3\rho x_t-1\right)\sigma E-\sigma^2}{1-\rho x_t}\\ \sigma_{x,t} &= E-\sigma\end{aligned}$$

From equation (**??**)

$$\frac{dx_t}{x_t}=\mu_{x,t}dt+\sigma_{x,t}dZ_t.$$

The mean of dx_t/x_t is a quartic equation with polynomial of degree four. Under OE,

$$\mu_{x,t}=x^c\Rightarrow x_t=\frac{\gamma m+1}{3\rho+m\rho^h+\gamma m^2\rho^h+2\gamma m\rho}.$$

-

0.5. **Leverage Exposure and Liquidity.** Define the state variable in HK(2013) as $y_t\equiv\frac{W_t}{P_t}$. From goods market clearing condition,

$$\rho(y_tP_t)+\rho^h(P_t-P_ty_t)=D_t\quad\Leftrightarrow\quad\frac{P_t}{D_t}=\frac{1}{\rho^h-\Delta\rho y_t}.$$

$$\Rightarrow x_t=\frac{y_t}{\rho^h-\Delta\rho y_t}.$$

$$\Rightarrow\frac{\lambda_t}{\lambda_t^U}=\frac{1}{1+m}\left[1-\frac{\rho\gamma}{\rho^h\gamma^h}\left(\frac{1}{\rho x_t}-1\right)\right]=\frac{1}{1+m}\left[1+\left(1-\frac{\gamma^h-\gamma}{\gamma^h}\right)\left(\frac{1-y_t}{y_t}\right)\right].$$

Compare with equation (20) in HK(2013), it is obvious that the liquidity $\lambda=\frac{\gamma^h-\gamma}{\gamma^h}$. And the effective financial constraint $\tilde{m}=\frac{\gamma^h}{\gamma}m=\frac{m}{1-\lambda}$. In the bad state, $\left(\gamma^h-\gamma\right)\downarrow\Rightarrow\lambda\downarrow\Rightarrow\tilde{m}\downarrow$. Thus, the wedge of two agents ambiguity aversion or the belief dispersion matters for household liquidity demand.

0.6. **Detection Error Probability Calibration.** Obtain the relative entropies from household and expert optimal robust problem,

$$g^h\left(x_t\right)=\sqrt{\mathcal{L}_t}=\theta^h\sigma_{W,t}^hV_w=\frac{\theta^h\sigma_{R,t}\varepsilon_t^h}{\rho^hW_t^h}=\frac{\theta^h}{\rho^h\gamma^h}\frac{\pi_{R,t}-k_t}{\sigma_{R,t}}=\frac{\gamma^h-1}{\gamma^h}\left(\frac{\pi_{R,t}}{\sigma_{R,t}}-\frac{k_t}{\sigma_{R,t}}\right).$$

In constraint case, from equation (0.12),

$$\frac{\pi_{R,t}}{\sigma_{R,t}}-\frac{k_t}{\sigma_{R,t}}=\frac{\tilde{m}\rho^h\sigma\gamma}{\left(1-\rho x_t\right)\left(m\rho^h+\rho\right)}.$$

In unconstrained case, $k_t=0$,

$$\frac{\pi_{R,t}}{\sigma_{R,t}}-\frac{k_t}{\sigma_{R,t}}=\frac{\sigma\gamma\gamma^h}{\rho\left(\gamma^h-\gamma\right)x_t+\gamma}.$$

$$\begin{aligned}
\Rightarrow g^h(x_t) &= \frac{\gamma^h - 1}{\gamma^h}\left[\frac{\gamma^h m\rho^h\sigma}{(1-\rho x_t)(m\rho^h+\rho)}\mathbf{1}_{x_t\in(0,x^c]} + \frac{\sigma\gamma\gamma^h}{\rho(\gamma^h-\gamma)x_t+\gamma}\mathbf{1}_{x_t\in(x^c,\frac{1}{\rho}]}\right]. \\
&= \sigma\gamma\left(\gamma^h-1\right)\left[\frac{m\rho^h}{\gamma(m\rho^h+\rho)(1-\rho x_t)}\mathbf{1}_{x_t\in(0,x^c]} + \frac{1}{\rho(\gamma^h-\gamma)x_t+\gamma}\mathbf{1}_{x_t\in(x^c,\frac{1}{\rho}]}\right]
\end{aligned}$$

$$g(x_t) = \sqrt{\mathcal{H}_t} = \theta\sigma_{W,t}J_w = -\frac{\theta\sigma_{R,t}\varepsilon_t}{\rho W_t} = -\frac{\theta}{\rho\gamma}\frac{\pi_{R,t}}{\sigma_{R,t}} = \frac{\gamma-1}{\gamma}\frac{\pi_{R,t}}{\sigma_{R,t}}$$

$$\Rightarrow g(x_t) = \sigma\gamma^h(\gamma-1)\left[\frac{1}{\gamma^h(m\rho^h+\rho)x_t}\mathbf{1}_{x_t\in(0,x^c]} + \frac{1}{\rho(\gamma^h-\gamma)x_t+\gamma}\mathbf{1}_{x_t\in(x^c,\frac{1}{\rho}]}\right]$$

where $\mathbf{1}$ denotes the indicator function.