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 are assumed to possess greater information processing capacity. Households purchase this capacity, and then delegate their investment decisions to intermediaries. As in He and Krishnamurthy (2012), the delegation contract is constrained by a moral hazard problem, which gives rise to a minimum capital requirement. Both agents have a preference for robustness, reflecting ambiguity about asset returns (Hansen and Sargent (2008)). We show that ambiguity aversion tightens the capital constraint, and amplifies its effects.

Indirect inference is used to calibrate the model’s parameters to the stochastic properties of asset returns. Detection error probabilities are used to discipline the degree of ambiguity aversion. The model can explain both the unconditional moments of asset returns and their state dependence, even with DEPs in excess of 15%.

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 financial intermediaries. Although these intermediaries may be investing on behalf of households, the contractual relationships between them are plagued by frictions. In HK12, asymmetric information produces a moral hazard problem that leads to a capital constraint, requiring the intermediary to maintain a minimum amount of 'skin in the game'. Due to these frictions, HK argue that it is better to relate asset prices to the marginal value of intermediary wealth than to the marginal utility of household consumption. The resulting stochastic discount factor becomes more volatile, and 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 arbitrage 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 empirical analysis to the market for mortgage-backed securities.

In this paper, we argue that intermediary asset pricing is indeed ‘the cake’. We respond to Cochrane’s skepticism in two ways. First, we operationalize the vague notion of ‘complexity’ by assuming agents face limits on their ability to process information, giving rise to Rational Inattention (Sims (2003)). Mean dividend growth is assumed to be stochastic and unobserved, so agents must solve a filtering problem. Intermediaries are assumed to have greater channel capacity, which enables them to solve this problem with less estimation risk. By delegating their portfolios to intermediaries, households can effectively purchase this additional capacity. Although households are free to invest for themselves, most choose not to do so.²

Just as Merton’s (1973) ICAPM can co-exist with the Lucas (1978) model, our model does not require that all agents delegate. What is important is that incentive constrained intermediaries are the marginal investor, making their marginal value of wealth a valid stochastic discount factor. It could well be that sophisticated private investors are also marginal. If so, their marginal utilities of consumption could also serve as stochastic discount factors. Unfortunately, we don’t observe their consumption. Hence, the case for Intermediary Asset Pricing is really a pragmatic and empirical one, based on the conjecture that relatively well measured financial intermediary data is more closely related to asset prices than poorly measured consumption data.³

²Kacperczyk, Nosal, and Stevens (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. 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’, which arises from loss aversion.

³Mankiw and Zeldes (1991) note that the consumption-CAPM works better using data on households that own stocks. There are still significant shortcomings, however, which perhaps reflect the poor quality of PSID consumption data. Adrian, Etula, and Muir (2014) and He, Kelly, and Manela (2017) show that leverage is a significantly priced factor in traditional cross-sectional regressions. Adrian, Moench, and Shin (2016) show that leverage also has time-series predictive ability.
Our second response to Cochrane is to suppose that investment is subject to Knightian Uncertainty, or equivalently, ambiguity. Our approach is based on the work of Hansen and Sargent (2008). Agents are assumed to have a correctly specified benchmark model of 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.\(^4\)

We show that ambiguity aversion tightens the capital constraint, and amplifies its effects. With ambiguity, incentive constraints can bind in otherwise normal times. A financial crisis isn’t required. This makes intermediary asset pricing more than just ‘frosting on the cake’.\(^5\) The reason ambiguity tightens the constraint is subtle. In contrast to Maenhout (2004), we do not scale the entropy penalty parameter by the value function. A constant entropy penalty produces horizon effects in portfolio choice. In particular, the effective degree of ambiguity aversion depends on an agent’s 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 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 (Yan (2008)).\(^6\) Amplification occurs because the relative pessimism of specialists increases once the constraint binds. As in HK12, when the constraint binds the relative risk exposure of specialists increases. Since subjective uncertainty ‘hides behind’ objective risk, the increasing leverage of specialists makes them endogenously more pessimistic. Since they are the marginal investor, their relative pessimism increases risk premia and the price of risk.\(^7\)

Although we are able to derive explicit expressions for asset prices and the distribution of wealth, these processes are highly nonlinear, and feature an occasionally and endogenously

\(^4\)There is a vast literature on asset pricing and Knightian Uncertainty. Epstein and Schneider (2010) survey the early literature. Brenner and Izhakian (2018) provide a more recent example. Our contribution is to combine Knightian Uncertainty with intermediation frictions and heterogeneous agents.

\(^5\)Caballero and Krishnamurthy (2008) also develop a model that combines scarce intermediary capital and Knightian Uncertainty. However, they focus on the role of central bank intervention rather than on asset pricing.

\(^6\)In contrast, the model in HK12 does not possess a nondegenerate stationary equilibrium, which makes it difficult to evaluate empirically. 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.

\(^7\)The state dependence in the degree of robustness here bears some resemblance to the business cycle model of Bidder and Smith (2012). However, the stochastic volatility in their model is exogenous. Here it arises endogenously via equilibrium portfolio policies. It also bears some resemblance to the heterogeneous risk aversion models of Chan and Kogan (2002) and Gărlăneanu and Panageas (2015). Once again, the difference here is that heterogeneity is endogenous.
binding constraint, which makes the model difficult to fit to the data using conventional methods. In response, we use the simulation-based methodology of ‘indirect inference’ (Gourieroux, Monfort, and Renault (1993)). We assess fit using two types of auxiliary functions. The first consists of simple unconditional means and variances of asset returns, price/dividend ratios, and risk premia. Since there are many frictionless models that can match unconditional moments, our second type of auxiliary functions are designed to assess the more challenging task of matching the persistence and cyclicity of asset returns. In particular, we compare model-implied and empirically estimated autoregressive processes for risk premia and price/dividend ratios. The free parameters consist of dividend volatility, rates of time preference, and entropy penalty parameters designed to capture the degree of robustness in control and filtering.

With the exception of the price/dividend ratio, we find that the model does well at matching unconditional moments. This is perhaps not too surprising given the results of Barillas, Hansen, and Sargent (2009), who show that model uncertainty can explain the equity premium and risk-free rate puzzles, although it is noteworthy that we can do so with higher detection error probabilities. The model does somewhat less well at capturing state dependence. Although it can replicate the persistence of the equity premium and price/dividend ratio, it generates only a 100-150 basis point increase in the equity premium during ‘crises’, and understates movements in the price/dividend ratio. We suspect that part of the problem here is that our model-implied relative wealth variable is a poor proxy for financial sector capital. To check this we use data from He, Kelly, and Manela (2017). They construct market equity capital ratios for approximately two dozen New York Fed primary dealers for the period 1970-2012. These are firms like JP Morgan, Goldman Sachs, and Citigroup. With this data, the model generates greater state dependence in risk premia and the price/dividend ratio.

The remainder of the paper is organized as follows. The next section briefly discusses the ongoing trend toward delegated wealth management, and compares our approach to the already voluminous literature on this topic. Section 3 outlines the information and market structure, and the objective functions of households and specialists. Section 4 imposes market-clearing, and solves for equilibrium asset prices. As in HK12/13, the key endogenous state variable is the distribution of wealth. We show that ambiguity induced heterogeneous beliefs are the crucial frosting on the Intermediary Asset Pricing cake. Section 5 outlines our indirect inference estimation strategy, and compares model predictions to US asset market data. Section 6 describes the simulation methodology we use to compute detection error probabilities. We show that our agents’ doubts are empirically plausible. Finally, Section 7 provides a few concluding remarks and offers extensions for future research. A technical Appendix contains proofs and derivations.

2. Delegated Wealth Management

Between 1980 and 2006 the share of US equities held by institutional investors (e.g., mutual funds and hedge funds) increased from 32% to 68% (Lewellen (2011)). Although precise data do not exist, it is widely suspected that institutional investors hold even a higher share of derivative securities. Growth in wealth management is even present in value-added GDP statistics. During the same time period, the share of GDP accounted
for by financial services grew from 4.9% to 8.3% (Greenwood and Scharfstein (2013)). (It has since declined a bit, to 7.5% in 2018).

Not surprisingly, this growth has attracted the attention of financial economists. Much of the early work was empirical. It asked whether professional money managers earn positive (risk-adjusted) returns. There is by now a consensus that they do not. For example, a widely cited calculation by French (2008) suggests that the average investor could increase his annual returns by 67 basis points by following a simple passive, market-portfolio, investment strategy.

More recently, financial economists have turned their attention to the equilibrium pricing implications of delegated wealth management. The Intermediary Asset Pricing literature is a direct descendant of this literature. Early work by Chevalier and Ellison (1997) showed that induced payoff convexity caused by the relationship between fund performance and customer flows can create potential conflicts of interest between funds and their customers. Kaniel and Kondor (2012) introduce portfolio delegation and a convex flow/performance function into a Lucas-tree economy. They show that portfolio delegation can generate increased Sharpe ratios, and can explain why on average funds outperform the market during downturns, while underperforming during booms. Basak and Pavlova (2013) study the equilibrium pricing implications of another commonly observed feature of delegated wealth management, i.e., ‘benchmarking’, wherein funds are evaluated based on their returns relative to a benchmark index (e.g., the S&P500). They show that benchmarking can generate countercyclical Sharpe ratios and increase market volatility. Compared to this literature, the HK12/13 models are relatively simple, in that customer flows do not depend on performance. Instead, the HK models focus on the nonlinearity that is induced by an occasionally binding incentive constraint. Consequently, they are more designed to address the asset pricing implications of financial crises.

As noted by Cochrane (2017), a potential criticism of the HK12/13 models is that they do not explain why households delegate. The absence of measurable excess returns from delegation suggests that other motivations must be at work. Our model motivates delegation by introducing heterogeneity in information processing. Mean returns are unobserved, and households delegate because financial institutions reduce estimation risk. Recent work by Gennaioli, Shleifer, and Vishny (2015) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) provides similar motivations. Gennaioli, Shleifer, and Vishny (2015) argue that delegation is based on ‘trust’, not performance, where trust is defined to be a reduction in the subjective perception of risk. Our model simply makes this explicit by introducing filtering and estimation risk. Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) is even more closely related, as it is explicitly based on information processing constraints and Rational Inattention. However, instead of focusing on asset prices, they focus on the cyclical allocation of attention. Their model predicts that fund managers focus on aggregate shocks during recessions, while focusing more on idiosyncratic shocks during booms.

\footnote{Berk and Green (2004) argue that in equilibrium the net-of-fees excess return \textit{should} be zero.}
3. Information, Market Structure, and Preferences

3.1. The HK12 Model. Our contribution is to introduce robust control and filtering into the HK12 model. Before doing that, it is useful to begin by briefly summarizing their model. Consider an infinite horizon continuous-time Lucas (1978) endowment economy populated by two types of agents, specialists and households. There are two assets: one risky and one risk-free. Only specialists can invest in the risky asset. However, in contrast to traditional segmented-markets models (e.g., Basak and Cuoco (1998)), households can indirectly invest in the risky asset by delegating some of their wealth to the specialist. At every $t$, households invest in intermediaries run by specialists. The relationship is short term, i.e. only lasts from $t$ to $t + dt$. Household’s cannot observe the portfolio choices of the specialist, nor can they observe its ‘effort’ level. These unobserved choices produce a moral hazard problem. Figure 1 depicts the economy’s market structure. The intermediary sector is indicated in the middle block.

![Figure 1: Market Structure and Intermediation Relationship](image1)

Specialist wealth is $W_t$, and household wealth is $W^h_t$. Households allocate $T^h_t$ to purchase intermediary equities, and the remaining fraction is used to buy riskless bonds. Intermediaries absorb in total $T^I_t$ dollars, $T^h_t$ from households and $T_t$ from their own wealth. They then allocate a fraction $\alpha_t$ to the risky asset and $1 - \alpha_t$ to the riskless bond. Assuming there is no short-selling constraint for the intermediary, we expect $\alpha_t$ 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 $(\alpha_t - 1)T^I_t$ from the bond market. The total risky asset exposure of the intermediary is $\varepsilon^I_t$. Through an affine contract developed by HK12, $\beta_t \in [0, 1]$ is the share of returns going to specialist, while $1 - \beta_t$ goes to households. Thus, at time $t$, the specialist bears a total risk exposure of $\varepsilon_t = \beta_t \varepsilon^I_t$, and the household is offered an exposure of $(1 - \beta_t) \varepsilon^I_t$.

In practice, wealth management typically involves two fees - a one-time fixed cost, and an ongoing variable cost, typically expressed as a percentage of profits. HK12 only considers the variable cost, denoted $K_t$. It is an endogenous variable. Later we model the...
fixed cost $\bar{K}$ using filtering and information-processing constraints. It will be determined by differences in channel capacity.

The population measures of households and specialists are normalized to one. Both are infinitely lived and have log preferences over consumption. Denote households (specialists) consumption rate as $C_h^t$ ($C_t$). The household’s objective is to:

$$\max_{\{C_t, \varepsilon_t^h\}} E \left[ \int_0^\infty e^{-\rho_h t} \ln C_h^t dt \right],$$

while the specialist’s objective is to:

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

where $\rho_h$ and $\rho$ denote the time discount rates for households and specialists, respectively.

The budget constraints are

$$dW_h^t = \varepsilon_h^t (dR_t - r_t dt) - k_t \varepsilon_h^t dt + W_h^t r_t dt - C_h^t dt,$$

and

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

The endogenous risky asset return is defined as:

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

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 return $r_t$. The risk premium is defined as: $\pi_{R,t} \equiv \mu_{R,t} - r_t$. The equilibrium per-unit-of-exposure intermediation fee is $k_t$, and

$$q_t \equiv \frac{K_t}{W_t} = \left(1 - \frac{\beta_t}{\beta_t} \right) \frac{\pi_{R,t}}{\sigma_{R,t}} k_t,$$

are total variable intermediation fees, expressed as a share of specialist wealth.

Households obtain an exposure $\varepsilon_t^h$ from the intermediary with an excess return indicated by the first term in the budget constraint, i.e., $\varepsilon_t^h (dR_t - r_t dt)$. Specialists bear a risky exposure $\varepsilon_t$, by investing 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 in the household’s budget constraint represents this fee. The specialist chooses the optimal contract share $\beta_t$ to maximize the intermediation fee. The third term is the risk-free interest earned 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 $k_t = 0$. The full-information rational expectations solutions for the above two maximization problems are:

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

and

$$\varepsilon_t^{h*} = \frac{\pi_{R,t} - k_t}{\sigma_{R,t}^2} W_t^h,$$
The key endogenous state variable in HK12 is the scaled relative wealth of the specialist: \( x_t \equiv W_t / D_t \). It is governed by the following stochastic process:

\[
\frac{dx_t}{x_t} = \mu_{x,t} dt + \sigma_{x,t} dZ_t,
\]

where \( \mu_{x,t} \) and \( \sigma_{x,t} \) are endogenously determined drift and diffusion coefficients. They capture the nonlinear dynamics of the model. (See Appendix 7.2 for the derivations of \( \mu_{x,t} \) and \( \sigma_{x,t} \)).

### 3.2. Capacity Constrained Robust Filtering

As noted earlier, our model combines robust control and robust filtering. A key benefit of our log preference specification is that we can separate these two problems. Here we focus on the robust filtering problem, which we later use to motivate the portfolio delegation decision. Interestingly, we shall see that even with log preferences, agents care about estimation risk when they distrust their priors.

The analysis in this section can be interpreted as combining the backward-looking robust filtering approach of Hansen and Sargent (2005) with the channel capacity-constrained Merton model of Turmuhambetova (2005).

Letting \( D_t \) be the exogenous dividend process, the capacity constrained robust filtering problem can be written as follows:

\[
\frac{dD_t}{D_t} = g_t dt + \sigma dZ^0_t, \tag{3.7}
\]

\[
dg_t = \rho_g (\hat{g} - g_t) dt + \sigma_g dZ^g_t, \tag{3.8}
\]

\[
ds_t = g_t dt + \sigma_s dZ^s_t, \tag{3.9}
\]

where \( Z^0_t, Z^g_t, \) and \( Z^s_t \) are independent standard Brownian motions. In contrast to HK12, we assume that mean dividend growth, \( g_t \), is both stochastic and unobserved. A preference for robustness arises because agents distrust the benchmark model for mean dividend growth (3.8). At the same time, the channel capacity constraint imposes an upper bound on the precision of the signal \( 1/\sigma_s \) in equation (3.9).

First consider the capacity constraint. Let \( \kappa \) denote the agent’s channel capacity:

\[
\lim_{T \to \infty} \sup \frac{1}{T} \mathcal{I}(g^T; s^T) \leq \kappa, \tag{3.10}
\]

where \( \mathcal{I}(g^T; s^T) \) is the Shannon mutual information between the true state process \( g^T = \{g_t: 0 \leq t \leq T\} \) and the noisy signal process \( s^T = \{s_t: 0 \leq t \leq T\} \). Mutual information measures the average reduction in uncertainty per unit of time over an infinite horizon. It follows from Duncan (1970) that in the stationary case the constraint becomes

\[
\lim_{T \to \infty} \sup \frac{1}{T} \mathcal{I}(g^T; s^T) = \frac{1}{2} \frac{Q}{\sigma_s^2} \leq \kappa,
\]

Using data on CDS credit spreads, Boyarchenko (2012) shows that state uncertainty and robust filtering were important in the early phase of the financial crisis. She also provides evidence that the importance of model uncertainty and robust control increased relative to state uncertainty as the crisis unfolded.
where $\hat{g}_t = \mathbb{E}_t [g_t | Z_t]$ and $Q_t = \mathbb{E}_t [(g_t - \hat{g}_t)^2 | Z_t]$ are the conditional mean and variance of $g_t$, respectively, and $Q = \mathbb{E} [Q_t]$ is the steady state variance.

Now let’s turn to the robustness part of the problem. Following Hansen and Sargent (2005), the robust filter can be implemented by introducing a drift distortion into the Kalman filtering equation for the conditional mean, and another distortion into the Ricatti equation for the conditional variance.\(^{10}\)

The robust filtering equations for the conditional mean and variance can be written as:\(^{11}\)

\[d\hat{g}_t = \frac{\rho_g (\bar{g} - \hat{g}_t)}{\sigma} dt + \frac{Q_t}{\sigma^2} d\tilde{Z}^{s}_t,\]  
\[dQ_t = \left[\sigma^2 - 2\rho_g Q_t - \frac{Q_t^2}{\sigma^2} \left(1 + \frac{1}{\sigma^2} - \theta_2\right)\right] dt,\]  
\[i = \{s, h\}.\]

where $d\tilde{Z}^s = dZ_t - \omega_t dt$, $dZ_t = \frac{1}{\sigma} \left(\frac{dD_t}{dt} - \hat{g}_t dt\right)$ and $d\tilde{Z}^{s}_t = \frac{1}{\sigma^2} (ds_t - g_t dt)$ are innovations corresponding to (3.7) and (3.9), respectively. The parameter $\theta_2$ penalizes the evil agent’s distortions, $\omega_t$, to the conditional mean. Since the capacity constraint will always bind, we can finally write the capacity constrained robust filter as:

\[d\hat{g}_i^t = \frac{\rho_g (\bar{g} - \hat{g}_i)}{\sigma} \left(\frac{Q_i}{\sigma^2} d\tilde{Z}^{i}_t + \sqrt{2\kappa^i Q_i} d\tilde{Z}^{s}_t\right),\]  
\[dQ_i^t = \left[\sigma^2 - 2\rho_g Q_i^t - \left(1 - \frac{1}{\sigma^2} - \theta_2^i\right) Q_i^{12} - 2\kappa^i Q_i^t\right] dt, i = \{s, h\}.\]

Note that a large channel capacity, $\kappa$, accelerates learning (i.e., causes $Q_t$ to decrease faster). Although in general, households and specialists have their own filtering equations, in a delegated equilibrium, only specialists filter.

### 3.3. Robust Control

In addition to having doubts about the unobserved state, $g_t$, agents also have doubts about the dividend process, (3.7), and equilibrium asset returns. In response, they formulate robust portfolio policies (Anderson, Hansen, and Sargent (2003), Maenhout (2004)). Ambiguity about asset returns implies that the budget constraint in (3.3) is viewed as merely a useful approximating model, associated with a benchmark probability measure $\mathbb{P}$. To ensure robustness, the agent surrounds the approximating model by a convex set of unstructured alternatives, and then optimizes against the worst-case model within the set. The agent recognizes the worst-case model depends on his own actions. If we let $Q$ denote the probability measure of the worst-case model, then Girsanov’s Theorem implies that the conditional relative entropy (or Kullback-Leibler distance) between the benchmark and worst-case models is given by a drift distortion.

\(^{10}\)Note that we are assuming the evil agent can commit to previous distortions of the conditional mean. Hansen and Sargent (2007) develop a forward looking robust filter that does not permit the evil agent to commit. In this case, the Ricatti equation for the conditional variance is unaltered.

\(^{11}\)See Appendix 7.1 for the proof.
\[ \nu_t^{12} \int \log \left( \frac{d\mathbb{Q}_t}{d\mathbb{P}_t} \right) d\mathbb{Q}_t = \frac{1}{2} \mathbb{E}^\mathbb{Q} \int_0^t (\nu_t^i)^2 ds. \]

The \( \nu_t^i \) process conveniently parameterizes the alternative models. A hypothetical evil agent chooses \( \nu_t^i \) subject to a relative entropy cost. Using this distortion, we can define a change of measure \( dZ_t = dZ_t^0 - \nu_t^i dt \), where \( dZ_t \) is a Brownian motion under \( \mathbb{Q} \).

Under the distorted probability measure, the household’s problem becomes

\[ V(\hat{x}^h_t, Q^h_t, W^h_t, Y^h_t) = \sup_{\{C^h_t, \nu_t^h\} \{\nu_t^h, \omega_t^h\}} \inf \mathbb{E}^\mathbb{Q} \left[ \int_0^\infty e^{-\rho h t} \left( \ln C_t^h + \frac{1}{2\theta h} (\nu_t^h)^2 + \frac{1}{2\theta h^2} (\omega_t^h)^2 \right) dt \right] \]

subject to (3.13), (3.14), and

\[ dW^h_t = \left[ \varepsilon_t^h (\pi_{R,t} - k_t) + r_t W^h_t - C_t^h \right] dt + \sigma_t^{hW} \left( \nu_t^h dt + dZ_t \right), \quad (3.15) \]

where \( \sigma_t^{hW} \equiv \sigma_{R,t} \varepsilon_t^h \). Note that the household’s dynamic programming problem features four state variables: \( (\hat{x}^h_t, Q^h_t) \) summarize his current beliefs, \( W^h_t \) is his current wealth, and \( Y^h_t \) captures the endogenous dynamics of equilibrium asset prices. As in HK12, \( Y^h_t \) is determined by a second order ODE in \( x_t \). Following Hansen and Sargent (2007), we assume that state and model uncertainty are constrained by separate relative entropy penalties, \( \theta_2^h \) and \( \theta^h \).

Using Itô’s Lemma, the HJB equation is:

\[ \rho^h V = \sup_{\{C^h_t, \varepsilon_t^h\} \{\nu_t^h, \omega_t^h\}} \inf \left[ \ln C_t^h + DV + \nu_t^h \sigma_{W,t}^h V_w + \omega_t^h Q_t^h V_g + \frac{1}{2\theta h} (\nu_t^h)^2 + \frac{1}{2\theta h^2} (\omega_t^h)^2 \right] \]

where \( D[\cdot] \) is the Dynkin operator,

\[ DV = V_w \left[ \varepsilon_t^h (\pi_{R,t} - k_t) + r_t W^h_t - C_t^h \right] + \frac{1}{2} V_{ww} \left( \varepsilon_t^h \right)^2 \sigma_{R,t} \]

\[ + V_{wg} \frac{\sigma_{R,t}}{\sigma} \varepsilon_t^h Q_t^h + V_g \rho_g \left( \hat{g} - \bar{g}_t^h \right) + \frac{1}{2} V_{gg} \left[ \left( \frac{Q_t^h}{\sigma} \right)^2 + 2\kappa^h Q_t^h \right] \]

\[ + V_Q \left[ \sigma_{g}^2 - 2\rho_g Q_t^h - \left( \frac{Q_t^h}{\sigma} \right)^2 - 2\kappa^h Q_t^h \right] + \mu_t^{hY}, \]

and where \( \mu_t^{hY} \) is the endogenously determined drift of \( Y_t^h \).

Solving first the infimization part and substituting back into the HJB equation gives:

\[ \rho^h V = \sup_{\{C^h_t, \varepsilon_t^h\}} \left[ \ln C_t^h + DV - \frac{\theta h}{2} \sigma_{W,t}^h V_w^2 - \frac{\theta h}{2} \left( \frac{Q_t^h}{\sigma} V_g \right)^2 \right]. \quad (3.16) \]

Note that ambiguity makes the agent dislike variance of continuation utility. The following proposition summarizes the solution:

---

12 Absolute continuity between \( \mathbb{P} \) and \( \mathbb{Q} \) requires their diffusion coefficients to be the same.

13 Note that in doing this, we are combining the T1/T2 operator approach of Hansen and Sargent (2007) with the commitment filter of Hansen and Sargent (2005).
Proposition 3.1. The consumption policy is:
\[ C_t^{h*} = \rho^h W_t^h, \quad (3.17) \]
and the optimal risk exposure is:
\[ \varepsilon_t^{h*} = \frac{\sigma_{R,t}}{\gamma^h R_t^2} W_t^h, \quad (3.18) \]
and the optimal entropy constrained drift distortions of the control and filtering problems are:
\[

\begin{align*}
\nu_t^{h*} &= \frac{-\theta^h \varepsilon_t^h \sigma_{R,t}}{W_t^h h}, \\
\omega_t^{h*} &= -\frac{\theta^h}{\rho^h (\rho^h + \rho_y)} Q_t^h
\end{align*}
\]
where \( \gamma^h = 1 + \theta^h / \rho^h \) is the household’s effective degree of ambiguity aversion and \( \theta^h, \theta_2^h \) summarize the degree of model and state uncertainty.

Given log preferences, the household’s value function takes the additively separable form
\[ V(\hat{g}_t^h, Q_t^h, W_t^h, Y_t^h) = \frac{1}{\rho^h} \ln W_t^h + F^h(\hat{g}_t^h, Q_t^h) + Y_t^h, \]
where \( F^h(\hat{g}_t^h, Q_t^h) \) is determined by the following PDE
\[

\begin{align*}
\rho^h F^{i} &= \hat{g}_t^i - \bar{g}^i + F_g \left[ \rho_y (\bar{g}^i - \hat{g}_t^i) - \frac{\theta^i Q_t^i}{2 \sigma^2} F_r^2 \right] + \frac{1}{2} F_{g g} \left( \frac{Q_t^i}{\sigma^2} + 2 \kappa^i Q_t^i \right) \\
&\quad + F_{Q} \left( \sigma_{Q, Q}^2 - 2 \rho_y Q_t^i - \frac{Q_t^i}{\sigma^2} - 2 \kappa^i Q_t^i \right),
\end{align*}
\]
and \( Y_t^h \) (a function of aggregate state \( x_t \)) solves the following ODE
\[ \ln \rho^h - 1 + \left( Y_t^h \right)' \mu_{x,t} x_t + \frac{1}{2} \left( Y_t^h \right)'' \sigma_{x,t}^2 = \rho^h Y_t^h - \frac{(\pi_{R,t} - k_t)^2}{2 \rho^h \gamma^h R_t^2} - \hat{r}_t^i, \quad (3.22) \]
where \( \hat{r}_t^i = r_t - (\hat{g}_t^i - \bar{g}) \), \( i = h \).

Proof. See Appendix 7.3 for the derivations. \qed

We now turn to the specialist’s problem. The specialist’s problem can be written as:
\[
J(\hat{g}_t, Q_t, W_t; Y_t) = \sup_{\{C_t, \xi_t, \{\nu_t, \omega_t\} \}} \inf \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( \ln C_t + \frac{1}{2 \theta_1^2} \nu_t^2 + \frac{1}{2 \theta_2^2} \omega_t^2 \right) dt \right]
\]
subject to (3.13), (3.14), and
\[ dW_t = [\varepsilon_t \pi_{R,t} + (q_t + r_t) W_t - C_t] dt + \sigma_{W,t} (\nu_t dt + dZ_t), \quad (3.23) \]
where \( \sigma_{W,t} = \sigma_{R,t} \varepsilon_t \). The HJB equation is:
\[ \rho J = \sup_{\{C_t, \xi_t, \{\nu_t, \omega_t\} \}} \left[ \ln C_t + DJ + \nu_t \sigma_{W,t} J + \omega_t Q_t \frac{J_g}{\sigma} + \frac{1}{2 \theta_1} \nu_t^2 + \frac{1}{2 \theta_2} \omega_t^2 \right], \]
where
\[
D J = J_w [\varepsilon_t \pi R,t + (q_t + r_t) W_t - C_t] + \frac{1}{2} J_{ww} \varepsilon_t^2 \sigma_{R,t}^2 + J_{wg} \frac{\sigma_{R,t}}{\sigma} \varepsilon_t Q_t^s + J_g \rho_g (g - \hat{g}_t^s)
\]
\[
+ \frac{1}{2} J_{gg} \left( \frac{Q_t^2}{\sigma^2} + 2 \kappa Q_t^s \right) + J_Q \left( \sigma_{g}^2 - 2 \rho_g Q_t^s - \frac{Q_t^2}{\sigma^2} - 2 \kappa Q_t^s \right) + \mu_{Y,t}.
\]

Solving first the infimization part and substituting back into the HJB equation gives:
\[
\rho J = \sup_{\{C_t, \varepsilon_t\}} \left[ \ln C_t + DJ - \frac{\theta}{2} (\sigma_{W,t} J_w)^2 - \frac{\theta_2}{2} \left( \frac{Q_t^s}{\sigma} J_g \right)^2 \right]
\]
(3.24)

The following proposition summarizes the solution:

**Proposition 3.2.** The specialist’s consumption policy is:
\[
C_t^* = \rho W_t,
\]
(3.25)
and the optimal risk exposure is:
\[
\varepsilon_t^* = \frac{\pi_{R,t}}{\gamma \sigma_{R,t}^2} W_t,
\]
(3.26)
and the optimal entropy constrained drift distortions of the control and filtering problems are:
\[
\nu_t^* = -\frac{\theta \varepsilon_t \sigma_{R,t}}{\rho W_t},
\]
(3.27)
\[
\omega_t^* = -\frac{\theta_2}{\rho} \frac{Q_t^s}{\sigma},
\]
(3.28)
where \( \gamma = 1 + \theta/\rho \) is the specialist’s effective degree of ambiguity aversion and \( \theta, \theta_2 \) summarize the degree of model and state uncertainty.

The specialist’s value function takes the additively separable form
\[
J(\hat{g}_t^s, Q_t^s, W_t; Y_t) = \frac{1}{\rho} \ln W_t + F(\hat{g}_t^s, Q_t^s) + Y_t,
\]
where \( F(\hat{g}_t^s, Q_t^s) \) is determined by the PDE (3.21) (where \( i = s \)) and \( Y_t \) solves the following ODE
\[
\ln \rho - 1 + Y't' \mu_{x,t} x_t + \frac{1}{2} Y''t'' \sigma_{x,t}^2 x_t^2 = \rho Y_t - \frac{q_t + \hat{r}_t^i}{\rho} - \frac{\pi_{R,t}^2}{2 \rho^2 \gamma \sigma_{R,t}^2}. \]
(3.29)

**Proof.** See Appendix 7.3 for the derivations. \( \square \)

### 3.4. Portfolio Delegation

Portfolio delegation is based on the fact that if \( \kappa > \kappa^h \), then \( J > V \). This value function difference motivates trade in channel capacity. Households are willing to pay to use the specialists’ channel.\(^\text{14}\) We assume the decision is made once at time-0, and so is based on an ex ante expected value calculation. The probability measure used to compute this expectation is the probability measure associated with a delegated competitive equilibrium, reflecting our assumption that households make

\(^{14}\)Note that in contrast to the investment decision, we assume that the filtering efforts of the specialist are not subject to any incentive constraints.
unilateral delegation decisions. As a result, only the \( F(\hat{g}_i^h, Q_i^h) \) component of \( V \) influences the delegation decision. In the steady state, \( dQ_i^* = 0 \), which from equation (3.14) implies:

\[
(1 - \theta^2 \sigma^2) Q_i^{*2} + 2\sigma^2 (\kappa^i + \rho_g) Q_i^* - (\sigma_g \sigma)^2 = 0.
\]

Solving this quadratic,

\[
Q_i^* = \frac{-(\kappa^i + \rho_g) + \sqrt{(\kappa^i + \rho_g)^2 + (1 - \theta^2 \sigma^2) (\sigma_g / \sigma)^2}}{1/\sigma^2 - \theta^2}.
\]  

(3.30)

It is straightforward to then show

\[
dQ_i^* = \frac{\sigma^2}{1 - \theta^2 \sigma^2} \left[ -1 + \frac{\kappa^i + \rho_g}{\kappa^i + \rho_g + (Q_i^{*2} / \sigma^2) (1 - \theta^2 \sigma^2)} \right].
\]

It is clear that \( \frac{dQ_i^*}{d\kappa^i} < 0 \) when \( \theta^2 \) is sufficiently small. That is, a higher channel capacity reduces the steady state conditional variance. In the steady state, (3.21) reduces to a second order ODE:

\[
\rho^i F_i = \hat{g} - \bar{g} + \rho_g (\bar{g} - \hat{g}) - \frac{1}{2} \theta^2 \sigma^2 (\kappa^i + \rho_g) + \frac{1}{2} \sigma^2 \rho_g \sigma^2,
\]

where \( \sigma^2 = \sigma_g^2 - 2\rho_g Q_i^* \). The solution to this ODE can be written as follows:

\[
F^i (\hat{g}) = A^i (\hat{g} - \bar{g}) + B^i,
\]

where \( A^i = \frac{1}{\rho^i (\rho^i + \rho_g)} \) and \( B^i = -\frac{1}{\rho^i} \left[ \frac{\theta^2}{2} \left( \frac{Q_i}{\sigma} \right)^2 \frac{1}{\rho^2 (\rho^i + \rho_g)} \right] \). Evaluating at \( \hat{g} = \bar{g} \), this implies the household’s expected value function difference arising from delegation is given by

\[
F^h (\kappa) - F^h (\kappa^h) = \frac{\theta^2 (Q_h^2 - Q_s^2)}{2\sigma^2 \rho^h (\rho^h + \rho_g)^2}.
\]  

(3.31)

Note that what matters is the interaction between state uncertainty \( \theta^2 \) and channel capacity \( \kappa^i \). Capacity differences become more important when state uncertainty is greater. In fact, without state uncertainty, \( \theta^2 = 0 \), estimation variance does not matter, as you would expect given log preferences. Clearly, \( F^h (\kappa) > F^h (\kappa^h) \) as long as \( \theta^2 > 0 \), and \( \kappa > \kappa^h \). One can view equation (3.31) as determining a maximum delegation fee, \( \tilde{K} (\kappa, \kappa^h, \theta_2, \theta^2_h) \). Using our benchmark parameter values we find that this fee turns out to be 2.23% of consumption.\(^{15}\)

4. Market Equilibrium

We now combine the policy functions derived in the previous section with market clearing conditions and derive equilibrium asset prices. We shall see that ambiguity aversion tightens the capital constraint and amplifies its effect.

\(^{15}\)See the Appendix for the derivation.
4.1. Capital Constraint and Exposure Fee. The key assumption in HK12 is that specialists can shirk. Assume the specialist’s flow benefit from doing so is $B$. Assume the flow cost to the intermediary from the specialist’s shirking is $X$. The cost incurred by the specialist depends on his exposure to the risky asset. If we define $\beta_t$ as the intermediary’s exposure, then to prevent shirking, $\beta_t > B/X$. Alternatively, if we define $B/X \equiv \frac{1}{1+m}$, then we can write the incentive constraint as follows.

$$\varepsilon^h_t \leq m \varepsilon_t$$

(4.32)

where $\varepsilon_t, \varepsilon^h_t$ are the specialist’s and household’s exposures given in Propositions (3.1) and (3.2). The constraint in (4.32) is the key ingredient in the HK12 model. It implies that the household is less willing to supply capital to the intermediary when the specialist is less exposed. $m$ can be interpreted as an inverse measure of the severity of the agency problem. The lower is $m$, the more severe the agency problem. If we substitute the optimal exposure policies given in Propositions (3.1) and (3.2) into the incentive constraint, we can write the constraint in terms of wealth.

$$W^h_t \leq \tilde{m} W_t,$$

where $\tilde{m} \equiv \frac{\gamma^h}{\gamma} m$ is the effective capital constraint. From this, we have the following result:

**Proposition 4.1.** If $\theta/\rho > \theta^h/\rho^h$, then ambiguity aversion tightens the capital constraint.

The intuition is as follows. If $\theta/\rho > \theta^h/\rho^h$, then the specialist is effectively more ambiguity averse than the household. Agents worry more about model uncertainty when they are more patient. Since the constraint binds when the household wants to invest while the specialist does not, ambiguity tightens the constraint since this makes the specialist less willing to invest compared to the household.

The exposure fee $k_t$ is determined by supply and demand. The specialist’s exposure supply is a step function:

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

In contrast, the household’s exposure demand depends negatively on $k_t$, and is $\varepsilon^*_t = \frac{\pi_{R,t} - k_t}{\gamma^h \sigma^2_{R,t}} W^h_t$. Notice that both the exposure supply and demand curves are influenced by ambiguity aversion.

Figure 2 depicts the equilibrium intermediary fee. There are two cases. The left panel depicts the case when the capital constraint binds, and $k_t > 0$, reflecting the scarcity of intermediary capital. The right panel depicts the case when the constraint is slack, intermediary capital is plentiful, so that $k_t = 0$.

---

16As in HK12, the incentive constrained financial relationship between households and specialists can be supported via equity trade. Households purchase a share of the intermediary’s outstanding equity subject to a constraint $T^h_t \leq \tilde{m} W_t$.

17Note that if we had scaled the entropy penalty by the value function as in Maenhout (2004), this linkage between time preference and ambiguity aversion would disappear (See Maenhout (2004) Appendix B for details).
4.2. Equilibrium Definition. Here we provide a detailed definition of market equilibrium in our model economy:

**Definition 4.2.** An equilibrium for the economy is a set of progressively measurable price processes \( \{P_t, r_t, R_t, k_t\} \), and households’ decisions \( \{C^{h*}_t, \varepsilon^{h*}_t\} \), and specialists’ decisions \( \{C^*_t, \varepsilon^*_t, \beta^*_t\} \) such that:

1. Given the price processes, agents consumption and portfolio decisions solve the objective functions (3.1) and (3.2).
2. The intermediation market reaches equilibrium with risk exposure clearing condition:
   \[
   \varepsilon^{h*}_t = \frac{1 - \beta^*_t}{\beta^*_t} \varepsilon^*_t. \tag{4.33}
   \]
3. The stock market clears:
   \[
   \varepsilon^*_t + \varepsilon^{h*}_t = P_t. \tag{4.34}
   \]
4. The goods market clears:
   \[
   C^*_t + C^{h*}_t = D_t. \tag{4.35}
   \]
5. Transversality conditions satisfy:
   \[
   \lim_{t \to \infty} E \left[ \exp \left( -\rho h t \right) V(W^{h}_t, t) \right] = 0 \quad \text{and} \quad \lim_{t \to \infty} E \left[ \exp \left( -\rho t \right) J(W_t, t) \right] = 0.
   \]

4.3. Equilibrium Asset Prices. The following propositions summarize the influence of ambiguity and the capital constraint on the dynamics of asset prices.

(a) 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,

\[
W^{h}_t + W_t = P_t. \tag{4.36}
\]

In equilibrium, from the goods market clearing condition (4.35) and the optimal consumption rules of households and specialists, we have

\[
\rho W_t + \rho^h W^h_t = D_t. \tag{4.37}
\]
Proposition 4.3. The equilibrium price/dividend ratio is given by:

\[
\frac{P_t}{D_t} = \frac{1}{\rho^h} + \left(1 - \frac{\rho}{\rho^h}\right)x_t = \frac{1 + \Delta \rho x_t}{\rho^h},
\]

(4.38)

where \(\Delta \rho \equiv \rho^h - \rho\).

Notice that ambiguity aversion only affects the price/dividend ratio through its influence on \(x_t\). Since specialists are relatively patient, the price/dividend ratio falls as their relative wealth decreases. Therefore, since crises are characterized by declines of specialist’s wealth, the model generates procyclical movements in the price/dividend ratio.

When the risk sharing constraint just starts to bind, the threshold level of the state \(x^c\) can be written as:

\[
\varepsilon^h_t = m\varepsilon_t,
\]

which means that \(P_t - W_t = mW_t\). Together with the equilibrium price/dividend ratio, this allows us to restate the capital constraint in terms of the specialist’s scaled wealth:

\[
x^c = \frac{1}{m \rho^h + \rho}.
\]

(4.39)

When \(x_t \leq x^c\), the economy is within the constrained region; otherwise, when \(x_t > x^c\), the economy is unconstrained. Notice that the ambiguity aversion of both households and specialists influence the critical level of \(x^c\). As the relative ambiguity of the specialist increases, the threshold wealth level increases.

(b) Specialist’s portfolio share. The specialist invests all his wealth into the intermediary. The specialist then makes a portfolio choice to invest a share \(\alpha_t\) of the total firm wealth \(T^I_t = W_t + T^h_t\) into the risky asset, and the rest into the riskless bond. Thus, the intermediary’s total exposure is:

\[
\varepsilon^I_t = \alpha_t \varepsilon^I_t,
\]

or in other words,

\[
\varepsilon^*_t + \varepsilon^h_t = \alpha_t(W_t + T^h_t).
\]

(4.40)

Given the specialist’s optimal \(\varepsilon^*_t\) choice, this can be interpreted as a constraint on the intermediary’s portfolio. It requires the specialist to choose \(\alpha_t\) in order to reach the optimal risk exposure \(\varepsilon^*_t\). Similarly, households obtain their desired exposure \(\varepsilon^h_t\) by choosing how much of their wealth \(T^h_t\) to contribute to the intermediary. Notice that when \(\alpha_t > 1\), the intermediary is leveraged.

Later we shall see that the specialist’s share of the intermediary’s return plays an important role in the dynamics of asset prices. The following proposition characterizes this share:

Proposition 4.4. In the unconstrained region, the specialist’s share of the intermediary’s return is

\[
\beta^U_t = \left[1 + \frac{\gamma}{\gamma^h} \left(\frac{1}{\rho^h x_t} - \frac{\rho}{\rho^h}\right)\right]^{-1}.
\]

(4.41)

In contrast, in the constrained region,

\[
\beta_t = \frac{1}{1 + m}.
\]

(4.42)
Proof. See Appendix 7.4 for the proof. □

In other words, within the unconstrained region, the specialist’s share of the intermediary’s return declines as his wealth share declines. However, once the constraint binds, his share remains fixed at $\beta = \frac{1}{1+m}$.

Above it was noted that the households’ and specialists’ desired exposures can be restated in terms of the intermediary’s portfolio choice. The following proposition summarizes this relationship.

Proposition 4.5. In the unconstrained region, the specialist’s desired risk exposure and the implied optimal portfolio choice of the intermediary are

$$\varepsilon^U_t = \left[1 + \frac{\gamma}{\gamma^h} \left( \frac{1}{\rho^h x_t} - \frac{\rho}{\rho^h} \right) \right]^{-1} P_t$$

and $\alpha^U_t = 1$, respectively. In the constrained region,

$$\varepsilon^*_t = \frac{1}{1+m} P_t$$

and $\alpha^*_t = \frac{1}{\varepsilon_t} + \frac{\rho^h - \rho}{(1+m)p^h}$. (4.43)

Proof. See Appendix 7.4 for the proof. □

Notice that when the constraint binds, the intermediary becomes leveraged ($\alpha_t > 1$), and its exposure increases as the specialist’s scaled wealth declines. Also notice that when specialists are effectively more ambiguity averse than households, so that $\hat{m} < m$, ambiguity aversion magnifies leverage.

(c) Return volatility of the risky asset.

Return volatility is a central feature of financial crises. In general, the effects of the specialist’s scaled wealth on return volatility are subtle. However, the following proposition shows that return volatility unambiguously increases once the constraint binds.

Proposition 4.6. In the unconstrained region return volatility is,

$$\sigma^U_{R,t} = \frac{1}{P_t/D_t} \frac{\sigma}{\rho^h - \Delta \rho \beta_t^U}.$$  (4.44)

In the constrained region it is,

$$\sigma_{R,t} = \frac{1}{P_t/D_t} \frac{\sigma}{\rho^h - \Delta \rho \beta^*_t}.$$  (4.45)

Proof. See Appendix 7.4 for the proof. □

When the constraint is not binding, declines in $x_t$ have offsetting effects on $\sigma_{R,t}$. The induced decline in $P_t/D_t$ increases $\sigma_{R,t}$. However, the decline in the specialist’s return share offsets this. Once the constraint binds and the return share remains fixed, only the $P_t/D_t$ effect remains, and $\sigma_{R,t}$ unambiguously increases as $x_t$ falls further.

(d) The risk premium and financial constraint.

In addition to increases in volatility, financial crises are also accompanied by large increases in risk premia. The following proposition shows that risk premia increase during crises, and that ambiguity amplifies this increase.
Proposition 4.7. In the unconstrained region the risk premium is given by,

\[
\pi_{R,t}^U = \frac{\sigma_\gamma \gamma^h}{(1 + \Delta \rho x_t)} \left[ \frac{\left( \rho^h \gamma^h - \rho \right) x_t + \gamma}{\rho \left( \gamma^h - \gamma \right) x_t + \gamma} \right]^2. \tag{4.46}
\]

In the constrained region,

\[
\pi_{R,t} = \frac{\sigma_\gamma \rho^h \gamma}{x_t (1 + \Delta \rho x_t)} \left( 1 + \frac{1 + m (m \rho^h + \rho)}{(m \rho^h + \rho)^2} \right). \tag{4.47}
\]

Proof. See Appendix 7.5 for the proof. \qed

Notice, the risk premium is highly state dependent. In the appendix we derive the following comparative static results, which show that ambiguity amplifies the risk premium.

Corollary 4.8.

\[
\frac{d\pi_{R,t}^U}{d\theta} > 0, \quad \frac{d\pi_{R,t}^U}{d\theta^h} \geq 0, \quad \frac{d\pi_{R,t}}{d\theta} > 0, \quad \frac{d\pi_{R,t}}{d\theta^h} = 0.
\]

Proof. See Online Appendix for the proof. \qed

The intuition is straightforward, as agents become more ambiguity averse, they must be compensated with a higher equilibrium risk premium. It is interesting to notice that \( \theta \) positively changes the risk premium both in the unconstrained and constrained region, while \( \theta^h \) also has a positive impact, but only in the unconstrained region. This reflects the fact that when the constraint binds, the specialist is the only marginal investor.

(c) The market price of risk and uncertainty.

The market price of risk is defined as the Sharpe ratio. In our model, the conventional Sharpe ratio measures a combination of risk aversion and ambiguity aversion. Ambiguity increases it unambiguously. Barillas, Hansen, and Sargent (2009) call this induced increase ‘the price of model uncertainty’. Propositions 4.6 and 4.7 summarize the results:

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

\[
\frac{\pi_{R,t}^U}{\sigma_{R,t}^U} = \frac{\sigma \gamma^h}{\rho (\gamma^h - \gamma) x_t + \gamma}. \tag{4.48}
\]

In the constrained region,

\[
\frac{\pi_{R,t}}{\sigma_{R,t}} = \frac{\gamma}{m \rho^h + \rho} \frac{1}{x_t}. \tag{4.49}
\]

Proof. The proof is straightforward by combining Equations (7.81), (7.82), (4.46), and (4.47). \qed

The following corollary summarizes the comparative static effects of ambiguity on the Sharpe ratio.
Corollary 4.10.  

\[
\frac{d}{d\theta} \left( \frac{\pi_R^U}{\sigma_R^U} \right) > 0, \quad \frac{d}{d\theta} \left( \frac{\pi_U^R}{\sigma_U^R} \right) \geq 0; \\
\frac{d}{d\theta} \left( \frac{\pi_R^R}{\sigma_R^R} \right) > 0, \quad \frac{d}{d\theta^h} \left( \frac{\pi_R^R}{\sigma_R^R} \right) = 0.
\]

Proof. See Online Appendix for the proof. \[\square\]

As with the risk premium, in the constrained region only the specialist’s ambiguity has a direct effect on the Sharpe ratio. Again, this reflects the central argument in the intermediary asset pricing literature that specialists are the marginal investors.

(g) The interest rate.

The real challenge of asset pricing is to explain both a high equity premium and a low average riskless interest rate. Recursive preference can do this. However, implausibly high risk version coefficients are still needed. Barillas, Hansen, and Sargent (2009) show that an empirically plausible ‘model uncertainty premium’ can substitute for high risk aversion. Our model shares this advantage, and at the same time captures observed state dependence in these returns. The following proposition characterizes the equilibrium interest rate.

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

\[
r_t^U = \rho^h + \hat{g}_t - \rho \Delta \rho x_t + \sigma^2 \frac{\gamma^h (\gamma^h - \gamma) + \gamma}{(\gamma^h - \gamma) x_t + \gamma}. \tag{4.50}
\]

In the constrained region,

\[
r_t = \rho^h + \hat{g}_t - \rho \Delta \rho x_t - \sigma^2 \frac{(1 - \rho x_t) \left[ \rho (1 + \gamma m) + \rho^h m^2 \gamma^h \right] + \rho^h m^2 (\rho^h x_t - \gamma^h)}{(1 - \rho x_t) (\rho + m \rho^h)^2 x_t}. \tag{4.51}
\]

Proof. See Appendix 7.5 for the proof. \[\square\]

This yields the following comparative statics.

Corollary 4.12. When \( \gamma^h \leq \gamma \) and \( \theta = \theta^h = \bar{\theta} \), then

\[
\frac{dr_t^U}{d\theta^U} \leq 0, \quad \frac{dr_t}{d\theta} < 0.
\]

Proof. See Online Appendix for the proof. \[\square\]

The intuition is that, as ambiguity increases, the specialists want to borrow less and households want to save more, which drives down the interest rate.

(f) The intermediation fee.

The information processing superiority of the specialist allows him to earn two kinds of fees. The one time participation fee characterized in section 3.4, and an on-going portfolio management fee \( k_t \), which is state-dependent and varies with the specialist’s scaled wealth. The following proposition characterizes the equilibrium depicted in Figure 2.
Proposition 4.13. In the unconstrained region, the per-unit exposure price \( k_t^U = 0 \). In the constrained region,

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

(4.52)

Proof. See Appendix 7.5 for the proof.

This leads immediately to the following comparative static results.

Corollary 4.14. The intermediation fee increases when the specialist’s ambiguity increases but decreases when the household’s ambiguity decreases:

\[
\frac{dk_t}{d\theta} > 0, \quad \frac{dk_t}{d\theta^h} < 0.
\]

Proof. See Online Appendix for the proof.

This reflects simple supply and demand. The fee increases when the household wants to invest but the specialist does not.

Since a picture is worth a thousand words (or equations), the following figure depicts the intermediary’s optimal portfolio policy along with equilibrium asset prices. Each panel plots the results as a function of the specialist’s scaled wealth \( x_t \). We use the benchmark parameter values estimated in the following section, which are contained in Table 1. In all cases, we assume that \( \theta = \theta^h \). Since \( \rho < \rho^h \), this implies that specialists are effectively more ambiguity averse in all plots. To illustrate the effects of ambiguity, each panel contains four lines, pertaining to four alternative values of \( \theta = \theta^h \). For comparison, the blue solid line in each plot pertains to the no ambiguity case of HK12.

The top left plot shows that leverage increases when the constraint binds and that ambiguity both amplifies the effect and causes it to occur at higher levels of specialist wealth. The top right panel shows that when the constraint binds, return volatility increases. Interestingly, it also shows that ambiguity depresses return volatility away from the constraint. This comes from a wealth reallocation to the more ambiguity averse specialist, who would bear less risk exposure and this reduces the equilibrium return volatility. The middle plots show the risk premium and Sharpe ratio. Evidently, the effects of \( x_t \) and \( \theta, \theta^h \) on them are very similar. Both the risk premium and Sharpe ratio increase as \( x_t \) decreases within the constrained region, as in HK12. More importantly, ambiguity increases the effect. For example, at the HK12 constraint, the equity premium rises from only 2% in the HK12 model to more than 10% with our benchmark parameters \( \theta = \theta^h = 0.03 \). Also notice that its mild increase in the unconstrained region reflects the fact that wealth is being redistributed toward the relatively ambiguity averse agent. A similar mechanism is at work in Garleanu and Panageas (2015). The lower left plot depicts the specialist’s state-dependent intermediation fee. At the HK12 constraint, the variable intermediation fee is approximately 3%, which is close to estimates provided by Greenwood and Scharfstein (2013). Finally, the lower right plot depicts the equilibrium interest rate. Notice that the interest rate is insensitive to the specialist’s scaled wealth when the constraint does not bind, but it drops sharply when the constraint binds.
Figure 3: Policy Functions and Asset Prices

This graph plots the policy functions of the specialist’s portfolio share for risky asset \( \alpha_t \), risky asset volatility \( \sigma_{R,t} \), risk premium \( \pi_{R,t} \), intermediation exposure fee \( k_t \) and risk free rate \( r_t \) for different ambiguity parameters (\( \theta = \theta^h \)). The threshold value \( x^c \) (vertical line) separates the constrained (left) and unconstrained (right) regions. The solid blue line shuts down ambiguity. All other parameter values are from Table 1.
4.4. The Stationary Wealth Distribution. To map the previous plots into empirical predictions, we need the stationary distribution of \( x_t \), which is endogenous. In fact, without ambiguity, a stationary distribution does not exist, because when \( \rho^h > \rho \), households are eventually driven out of the economy, the specialists accumulate all the wealth, and the constraint never binds. However, when the specialist is effectively more ambiguity averse, the greater impatience of the households can be offset by the greater ambiguity of the specialists (Yan (2008)). The following result characterizes this distribution.

**Proposition 4.15.** The specialist’s scaled wealth process is given by the following diffusion:

\[
\frac{dx_t}{x_t} = \mu_{x,t} dt + \sigma_{x,t} dZ_t. \tag{4.53}
\]

In the unconstrained region the drift and volatility are given by

\[
\mu_{x,t}^U = \Delta \rho (1 - \rho x_t) + \sigma^2 + \sigma^2 \frac{A_0^U x_t^2 + A_1^U x_t + A_2^U}{(1 - \rho x_t) [\rho (\gamma^h - \gamma) x_t + \gamma]^2}, \tag{4.54}
\]

\[
\sigma_{x,t}^U = \sigma \left[ \frac{1}{\rho (\gamma^h - \gamma) x_t + \gamma} - 1 \right], \tag{4.55}
\]

where \( A^U_0 = \rho^2 (\gamma^h - \gamma) (2\gamma^h + \gamma) \), \( A^U_1 = \rho (2\gamma^2 - 3(\gamma^h)^2 + 2\gamma^h) \), and \( A^U_2 = (\gamma^h)^2 - \gamma^2 - \gamma^h \).

In the constrained region they are given by

\[
\mu_{x,t} = \Delta \rho (1 - \rho x_t) + \sigma^2 + \sigma^2 \frac{A_0 x_t^2 + A_1 x_t + A_2}{(m \rho^h + \rho)^2 (1 - \rho x_t) x_t^2}, \tag{4.56}
\]

\[
\sigma_{x,t} = \frac{1}{(m \rho^h + \rho) x_t} - 1 \tag{4.57}
\]

where \( A_0 = \rho^h (\rho \gamma^h - \rho^h) m^2 + \rho (\rho \gamma + \rho^h) m + 2 \rho^2 \), \( A_1 = -\rho^h \gamma^h m^2 - (\rho^h + 2\rho \gamma) m - 3\rho \), and \( A_2 = \gamma m + 1 \).

**Proof.** See Appendix 7.2 for the proof. \( \square \)

Whether this diffusion process yields a stationary distribution depends on the boundary properties of \( \mu_{x,t} \) and \( \sigma_{x,t} \). Intuitively, \( \mu_{x,t} \) should be negative for sufficiently large \( x_t \) and positive for sufficiently small \( x_t \). Assuming it exists, let \( f(x) \) be the stationary density associated with the diffusion in equation (4.53). Since \( x_t \) is univariate, we can obtain the following expression for the solution to the steady state Kolmogorov Fokker-Planck (KFP) equation, which characterizes this density.

**Proposition 4.16.** The stationary density of specialist scaled wealth satisfies

\[
f(x) = \frac{C_1}{\sigma^2_x(x)} \exp \left( \int_0^x \frac{2\mu_x(s)}{\sigma^2_x(s)} ds \right) + \frac{C_2}{(\sigma^U_x(x))^2} \exp \left( \int_{x^c}^{x} \frac{2\mu_x^U(s)}{(\sigma^U_x(s))^2} ds \right), \tag{4.58}
\]

where \( C_1 \) and \( C_2 \) satisfy:

(1) \( \int f(x) dx = 1 \);

(2) continuous at \( x^c \).
Proof. See Appendix 7.5 for the proof. □

Existence rests on the convergence of these integrals. A formal proof and explicit solution would be challenging, since the drift and diffusion coefficients are non-linear and change at the constraint. Instead, we simply use Monte-Carlo simulations to informally check whether the stationary distribution exits. Figure 4 depicts the empirical distribution of $x_t$ for the benchmark parameter values in Table 1. It is computed from 5000 sample paths, each consisting of 200 years.

![Figure 4: Stationary Distribution](image)

This figure plots a histogram of the specialist’s scaled wealth using the benchmark parameters. The vertical red line separates the constrained (left) and unconstrained (right) regions.

5. Quantitative Implications

5.1. Indirect Inference. Due to the endogenously binding capital constraint, our model would be difficult to estimate using traditional methods. However, it is relatively easy to generate sample paths from the model. Hence, it is a good candidate for the simulation-based estimation methodology of indirect inference (Gourieroux, Monfort, and Renault (1993), Smith Jr (1993)). The basic idea behind indirect inferences is to specify a set of (easily estimated) reduced form ‘auxiliary functions’, which are designed to capture features of the data of interest, i.e., comovement, volatility and persistence. The reduced-form parameters are estimated twice - once using the actual data, and again using repeated samples generated from the structural model. The model is then evaluated based on how close (averaged) estimates from the simulated data are to those from the actual data. If

---

18See Karlin and Taylor (1981).
19The figure is not changed with longer simulations. We also simulated 5000 sample paths of 5000 years. The distribution is roughly the same.
there are more auxiliary function parameters than model parameters, then some sort of minimum distance/weighting matrix must be specified. The minimum distance can be used to test the null that the structural model is correctly specified. Note, the auxiliary functions can be misspecified.

We use two sets of auxiliary functions. The first captures the model’s ability to explain unconditional moments. Because there are many models capable of explaining unconditional moments, we also want to assess our model’s ability to explain state dependence in risk prices and return premia. This is more challenging. Our second set of auxiliary functions are therefore based on our model’s ability to match simple autoregressions of equity premia and price/dividend ratios.

Several of our model’s parameters can be calibrated to outside data. For example, $m$ relates to the specialist’s return share. Following HK12, we set $m = 4$ throughout, reflecting the assumption that specialists maintain a 20% return share. HK12 cite evidence in support of this. Mean dividend growth is calibrated to match US dividend data. This implies $\bar{g} = 0.02$, $\sigma_g = 0.03$, and $\rho_g = 0.055$. The channel capacities are set so that the implied delegation fee is roughly equal to the data, which suggest that upfront management fees are around 1.5-2.5% of invested assets (Greenwood and Scharfstein (2013))). This leaves $(\sigma, \rho, \rho^h, \theta, \theta^h, \theta_2, \theta_2^h)$ as the free parameters. The first set of auxiliary functions consists of the following six unconditional moments: the mean equity premium, the mean Sharpe ratio, the mean risk-free rate, the mean price/dividend ratio, and the standard deviations of the risky and risk-free returns. The second set of auxiliary functions consists of the intercept and autoregressive coefficient in simple univariate AR(1) processes for the equity premium and price/dividend ratio. For the equity premium data we use estimates from Gagliardini, Ossola, and Scaillet (2016). They fit a 4-factor conditional linear factor model to US monthly stock returns for the period 1964-2009. We use the risk premium on the market factor for the period 1970-2009. The price/dividend data are from Shiller’s webpage, again for the period 1970-2009. The weighting matrix used to match the 6 parameters to the 10 targets is described in the appendix.

5.2. Estimates. Table 1 contains calibrated and estimated benchmark parameter values. The grid search constrained $\theta = \theta^h$, $\theta_2 = \theta_2^h$, and $\rho^h > \rho$. This ensures a stationary wealth distribution and imposes the restriction that, absent discounting effects, the ambiguity aversion of households and specialists is the same. For each candidate parameter value we simulate 5000 observations, and use the final 40 years to compute moments and regressions. We then average across 20 repetitions to compute distances.

The data moments and model estimates are summarized in Table 2. Estimates are based on averages across 5000 simulations using the benchmark parameter values. To assess the importance of model uncertainty, we also compute moments in the HK12 model. The HK12 model has the ambiguity aversion parameter values close to zero. The results are in the column labeled ‘Non-Robust’. The column labeled ‘Model 1’ contains results when the two agents are ambiguity averse ($\theta = \theta^h = 0.03$).

There are several points to notice. First, without ambiguity, the HK12 model cannot explain the Equity Premium Puzzle. It only produces a 1.94% equity premium and a 14.02% Sharpe ratio. Both are significantly lower than in the data. However, when

\[20\text{In order to ensure non-degenerate distribution, we set ambiguity aversion as 0.0001 instead of 0.}\]
Table 1: Benchmark Parameters

<table>
<thead>
<tr>
<th>Panel A. Preferences and Intermediation</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$ Specialist Time Discount Rate</td>
<td>0.007</td>
</tr>
<tr>
<td>$\rho^h$ Household Time Discount Rate</td>
<td>0.014</td>
</tr>
<tr>
<td>$\theta$ Specialist Model Uncertainty Preference</td>
<td>0.03</td>
</tr>
<tr>
<td>$\theta^h$ Household Model Uncertainty Preference</td>
<td>0.03</td>
</tr>
<tr>
<td>$\theta_2$ Specialist State Uncertainty Preference</td>
<td>1</td>
</tr>
<tr>
<td>$\theta^2_2$ Household State Uncertainty Preference</td>
<td>1</td>
</tr>
<tr>
<td>$\kappa$ Specialist Information Channel Capacity</td>
<td>0.01</td>
</tr>
<tr>
<td>$\kappa^h$ Household Information Channel Capacity</td>
<td>0.00999</td>
</tr>
<tr>
<td>$m$ Intermediation Multiplier</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Equity Market</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{g}$ Mean Dividend Growth Rate</td>
<td>0.02</td>
</tr>
<tr>
<td>$\sigma$ Dividend Growth Volatility</td>
<td>0.1389</td>
</tr>
<tr>
<td>$\sigma_g$ Unobservable Dividend Growth Volatility</td>
<td>0.03</td>
</tr>
<tr>
<td>$\rho_g$ Mean Reversion Rate of Unobservable Dividend Growth</td>
<td>0.055</td>
</tr>
<tr>
<td>$Q_0$ Initial Value of Posterior Variance</td>
<td>0.005%</td>
</tr>
</tbody>
</table>

Table 2: Measurements and Estimates

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Non-Robust</th>
<th>Model 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_2$</td>
<td>0.0001</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\theta^2_2$</td>
<td>0.0001</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.0001</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>$\theta^h$</td>
<td>0.0001</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.007</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>$\rho^h$</td>
<td>0.014</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>1.014</td>
<td>5.29</td>
<td></td>
</tr>
<tr>
<td>$\gamma^h$</td>
<td>1.007</td>
<td>3.14</td>
<td></td>
</tr>
<tr>
<td>Risk Premium (%)</td>
<td>5.91</td>
<td>1.94</td>
<td>6.27</td>
</tr>
<tr>
<td>Sharpe Ratio (%)</td>
<td>31.91</td>
<td>14.02</td>
<td>48.12</td>
</tr>
<tr>
<td>Interest Rate (%)</td>
<td>1.29</td>
<td>1.26</td>
<td>1.24</td>
</tr>
<tr>
<td>Interest Rate Volatility (%)</td>
<td>2.25</td>
<td>2.34</td>
<td>2.33</td>
</tr>
<tr>
<td>Return Volatility (%)</td>
<td>17.30</td>
<td>13.87</td>
<td>13.04</td>
</tr>
<tr>
<td>P/D mean</td>
<td>41.31</td>
<td>93.00</td>
<td>87.09</td>
</tr>
<tr>
<td>$k$ (%)</td>
<td>2</td>
<td>0</td>
<td>0.23</td>
</tr>
<tr>
<td>DEP Specialist (%)</td>
<td>49.25</td>
<td>24.98</td>
<td></td>
</tr>
<tr>
<td>DEP Household (%)</td>
<td>49.36</td>
<td>25.06</td>
<td></td>
</tr>
<tr>
<td>Prob. of Sharpe Ratio Exceed 105bps (%)</td>
<td>0</td>
<td>2.86</td>
<td></td>
</tr>
<tr>
<td>Prob. of Constraint Binds (%)</td>
<td>0</td>
<td>11.57</td>
<td></td>
</tr>
</tbody>
</table>

This table reports unconditional moments. We simulate 200 years and 5000 sample paths using our benchmark parameters. To correspond to our empirical sample, 1970-2010, we report means from the final 40 periods of each sample paths. distribution.
agents are concerned about model misspecification, the model’s performance improves significantly. It now produces a 6.27% equity premium and 48.12% Sharpe ratio. These are both close to the data, although the Sharpe ratio is a little high. Second, our model produces a mean interest rate of 1.24% compared to 1.29% in the data. Hence, we avoid the risk-free rate puzzle. Although the model’s interest rate volatility may appear too low, the reported value from the data pertains to the ex post real rate, which is considerably larger than the ex ante real rate. Third, and perhaps most important, the Non-Robust HK12 model does not produce any crises. The constraint never binds because the relative impatience of households causes the specialist to accumulate most of the economy’s wealth. In contrast, in our model the constraint binds more than 10% of the time, and when it does, it often produces significant increases in the price of risk. For example, 2.86% of the time the Sharpe ratio exceeds its long-run mean by more than 105 basis points.

Table 3 shows that our model also captures the persistence of the equity premium and price/dividend ratio.

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Non-Robust</th>
<th>Model 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta )</td>
<td>0.0001</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>( \theta^h )</td>
<td>0.0001</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Risk Premium: AR(1) Coefficient</td>
<td>0.9556</td>
<td>1.0007</td>
<td>0.9978</td>
</tr>
<tr>
<td>Risk Premium: Intercept</td>
<td>0.2741</td>
<td>-0.0014</td>
<td>0.0136</td>
</tr>
<tr>
<td>PD: AR(1) Coefficient</td>
<td>0.9964</td>
<td>1.0004</td>
<td>0.9995</td>
</tr>
<tr>
<td>PD: Intercept</td>
<td>0.1838</td>
<td>-0.0247</td>
<td>0.0435</td>
</tr>
</tbody>
</table>

This table reports AR(1) coefficients and constants estimated from risk premium and price/dividend ratio monthly data (1970-2010) as well as the model implied coefficients and constants. We simulate 40 years and average across 5000 sample paths with implied monthly frequency.

The autoregressive coefficients are quite close to those in the data, and the implied long-run means are also close.

5.3. Sample Paths. Although it is encouraging that the model matches the unconditional moments of asset returns, there are many other models that can do this as well. The real value-added of the HK12/13 models is to explain the highly nonlinear, state-dependent dynamics of asset returns. For example, a significant fraction of the average equity premium is generated during infrequent crisis episodes when it rises dramatically. Generating strong counter-cyclicality in risk prices and risk premia is more challenging. Perhaps the best way to assess this ability is to simply look at sample paths produced by the model. Figure 5 displays a representative sample path using our benchmark parameters. The time unit is a quarter. The red rectangular areas indicate times when the capital constraint binds. The top left panel plots the specialist’s scaled wealth, which is the model’s key endogenous state variable. The top right panel plots the drift distortions of the specialist and household, while the middle left panel plots their difference. The middle right and bottom left panels plot the risk premium and Sharpe ratio, respectively.
The blue horizontal lines depict sample means. Finally, the bottom right panel plots the interest rate.

Figure 5:
This figure plots a representative sample path using the benchmark parameters. Red rectangular regions indicate the constrained region. The light blue horizontal lines in the first two figures are unconditional means for risk premium and Sharpe Ratio, respectively.
These plots nicely reveal the model’s key forces. As in HK12/13, the top left plot shows that crises occur when the specialist’s wealth declines. The top right and middle left plots depict the new mechanism produced by model uncertainty. Note that as the specialist’s wealth declines, he becomes relatively pessimistic, i.e., his relative drift distortion increases. This occurs because he endogenously becomes more exposed to the risky asset. This endogenous belief heterogeneity amplifies the rise in risk premia and risk prices, as depicted in the middle right and bottom left plots. The risk premium rises by about 100 basis points, from about 6.2% to 7.2%. It is interesting to observe that the constraint can easily bind for a decade or more, although the peak itself tends to be rather short-lived. In contrast, HK13 emphasize that most crisis models have difficulty capturing the duration of crises. Finally, note that these crisis episodes contribute significantly to average equity premia and Sharpe ratios. As you would expect from Figure 3, when the constraint is slack, the equity premium and Sharpe ratio are nearly constant.

5.4. Implications From Observed Capital Ratios. A 100-150 basis point increase in the risk premium is significant, but still less than the increases observed during actual crises, where spreads often increase by several hundred basis points. We suspect that the failure of our model to fully match the magnitude of the increase arises from the fact that our model-generated wealth process produces less variation than observed in the data. We can check this by using data from He, Kelly, and Manela (2017). They collect data on market-value capital ratios for the New York Fed’s primary dealers. These institutions actively trade in most, if not all, asset markets. Although there is heterogeneity across dealers, they compute a simple value-weighted average to correspond with our model’s assumed ‘representative’ specialist. We can then use this observed capital ratio in place of $x_t$ in the equilibrium pricing equations of our model. Figure 6 depicts the results.

21 There is also considerable heterogeneity across assets. However, their empirical results suggest that the model’s assumption of a single risky asset might not be a bad approximation, since the estimated price of risk is quite similar across asset classes.
This figure plots the model implied risk premium and price/dividend ratio by imputing He, Kelly, and Manela (2017) data against the risk premium from Gagliardini, Ossola, and Scaillet (2016) and real price/dividend ratio from Robert Shiller’s website.

The left panel displays the risk premium and the right panel displays the price/dividend ratio. For comparison, the blue lines plot the data. As before, for the equity premium we use the market factor from Gagliardini, Ossola, and Scaillet (2016), and the price/dividend ratio from Robert Shiller’s website. Perhaps not surprisingly given Figure 3, during normal times the model generates little variation in the equity premium and price/dividend ratio. Notice, however, that during the recessions the model generates significant increases. For example, during the financial crisis of 2008-09, the model generates an increase in the equity premium of about 10 percentage points. Still less than in the data, but much closer.

6. Detection-Error Probabilities

We have seen that the ambiguity parameters $(\theta, \theta^h, \theta_2, \theta^h_2)$ play an important role in our model’s ability to fit the unconditional moments and state dependent dynamics of asset returns. It is possible to view these parameters as solely a reflection of preferences and allow them to be unrestricted. However, following Hansen and Sargent (2008), we prefer to interpret them as reflecting both the presence of model uncertainty and the agent’s preference for robustness. Under this interpretation, it is important to discipline the magnitude of $(\theta, \theta^h, \theta_2, \theta^h_2)$. In particular, we do not want to allow agents to hedge against empirically implausible alternative models.

In discrete time models, it is possible to distinguish the presence of ambiguity from the agent’s aversion to ambiguity (see e.g., Klibanoff, Marinacci, and Mukerji (2005) and Ju and Miao (2012)). However, in the continuous time limit, one must rescale $\theta$ to maintain ambiguity aversion (see Hansen and Miao (2018) for details). Given the hidden state in our model, we could in principle distinguish ambiguity from ambiguity aversion. We leave this for future work.

---

Figure 6: Model Implied Risk Premium and Price/Dividend Ratio
Plausibility is quantified using detection error probabilities (DEPs). Agents are viewed as statisticians, who attempt to discriminate among models using likelihood ratio statistics. When likelihood ratio statistics are large, DEPs are small, and models are easy to distinguish. DEPs will be small when models are very different, or when there is a lot of data. DEPs are based on an equally-weighted average of Type I and Type II errors:

$$\text{DEP} = \frac{1}{2} \text{Prob} (H_A|H_0) + \frac{1}{2} \text{Prob} (H_0|H_A).$$

Hence, a DEP is analogous to a p-value. We compute DEPs using the Monte-Carlo simulation strategy outlined in Chapter 9 of Hansen and Sargent (2008).

Since ex post only specialists are filtering, the null/approximating model is:

$$\frac{dD_t}{D_t} = \dot{g}_t dt + \sigma dZ_t,$$

$$d\dot{g}_t = \rho g (\bar{g} - \dot{g}_t) dt + \frac{Q_t}{\sigma} dZ_t + \sqrt{2\kappa Q_t} d\dot{Z}_t, \quad (6.59)$$

while the distorted/alternative model is:

$$\frac{dD^i_t}{D^i_t} = (\dot{g}_t + \sigma \nu^i_t) dt + \sigma d\dot{Z}^i_t,$$

$$d\dot{g}^i_t = \left( \rho g (\bar{g} - \dot{g}_t) + \frac{Q_t}{\sigma} \omega_t \right) dt + \frac{Q_t}{\sigma} d\dot{Z}_t + \sqrt{2\kappa^2 Q_t} d\dot{Z}^i_t. \quad (6.60)$$

Let $L_P$ and $L_Q$ be the likelihood of null model and alternative model, the log-likelihood ratios are $l|P = \log \left( \frac{L_P}{L_Q} \right)$ and $l|Q = \log \left( \frac{L_Q}{L_P} \right)$, respectively. When the null model generates the data,

$$l|P = \frac{1}{T} \sum_{t=1}^{T} \left( \frac{1}{2} w^n_t w^n_t - w^n_t \epsilon^n_t \right), \quad (6.63)$$

$$(l|P)^h = \frac{1}{T} \sum_{t=1}^{T} \left( \frac{1}{2} w^n_t w^n_t - w^n_t \epsilon^n_t \right), \quad (6.64)$$

for specialists and households, respectively, where $w^n_t = \left[ \nu^n_t \omega_t \right]$, $w^{nh}_t = \left[ \nu^{nh}_t \omega_t \right]$, $\epsilon^n_t = \left[ \begin{array}{c} Z_t \\ \bar{Z}_t \end{array} \right]$, $\epsilon^{nh}_t = \left[ \begin{array}{c} Z^h_t \\ \bar{Z}^h_t \end{array} \right]$ and $Z_t$, $\bar{Z}_t$, and $Z^h_t$ are i.i.d. If the alternative model is the data-generating process,

$$l|Q = -\frac{1}{T} \sum_{t=1}^{T} \left( \frac{1}{2} w^a_t w^a_t + w^a_t \epsilon^a_t \right),$$

$$(l|Q)^h = \frac{1}{T} \sum_{t=1}^{T} \left( \frac{1}{2} w^a_t w^a_t - w^a_t \epsilon^a_t \right). \quad (6.65)$$
where $\epsilon_t^a = \epsilon_t^{ah} = \left[ Z_t \right]$. The relative entropy computed from the household’s and specialist’s robust control problems can be written as:

$$23 \nu_t^h = \sigma \gamma (1 - \gamma^h) \left[ \frac{m \rho^h}{\gamma (m \rho^h + \rho) (1 - \rho x_t)} 1_{x_t \in (x_{\min}, x^c)} + \frac{1}{\rho (\gamma^h - \gamma) x_t + \gamma} 1_{x_t \in (x^c, 1/\rho)} \right],$$

$$\nu_t = \sigma \gamma^h (1 - \gamma) \left[ \frac{1}{\gamma^h (m \rho^h + \rho) x_t} 1_{x_t \in (x_{\min}, x^c)} + \frac{1}{\rho (\gamma^h - \gamma) x_t + \gamma} 1_{x_t \in (x^c, 1/\rho)} \right],$$

(6.67) (6.68)

where $x_{\min} = 1/(m \rho^h + \rho)$ and 1 denotes the indicator function. The negative drift distortion $\omega_t$ comes from the robust filtering solution, (3.28).

Figure 7 depicts the results using our benchmark parameters.

As in Hansen and Sargent (2007), each agent has two robustness parameters, one pertaining to the control problem, $\theta^i$, and one pertaining to the filtering problem, $\theta^i_2$. The left panel of Figure 7 plots DEPs as a function of the robust control parameter, $\theta^i$, holding constant the filtering parameter at its benchmark value. The right panel of Figure 7 plots DEPs as a function of the filtering problem, $\theta^i_2$, holding constant the control parameter. From the left panel, we can see that DEPs remain above 10% for values of $\theta^i$ as large as 0.07. However, such a large value of $\theta^i$ produces counterfactually large equity premia and Sharpe ratios. The right panel shows that DEPs actually increase with $\theta^i_2$.

---

23See Appendix 7.6 for the proof.

24It is worth noting that Barillas, Hansen, and Sargent (2009)’s frictionless model could only attain the Hansen-Jaganathan bound with DEPs below 0.05.
This reflects our use of the commitment filter of Hansen and Sargent (2005). In this case, as $\theta_2^i$ increases, the agent’s state estimation variance $Q$ increases (see equation (3.30)). Notice that as $\theta_2^i \to 0$, DEPs become quite small. Hence, state uncertainty is important. In section 3 we saw that when $\theta_2^i = 0$, there would be no portfolio delegation. Here we see that $\theta_2^i \neq 0$ is also important in generating plausible DEPs.

7. Conclusion

This paper has argued that Intermediary Asset Pricing offers more than 2nd-order market micro-structure corrections to existing asset pricing models. We’ve done this by introducing two new elements into the influential model of He and Krishnamurthy (2012, 2013). First, we formalize the notion of ‘complexity’ and motivate portfolio delegation by assuming agents face information processing constraints. Portfolio delegation allows households to purchase the higher channel capacity of financial intermediaries. Second, we assume that households and specialists are ambiguity averse. We show that ambiguity aversion tightens incentive constraints and amplifies their effects.

Although Lucas-style models are convenient for studying asset pricing, they have the drawback of eliminating feedbacks from asset prices to the real economy. It is widely believed that the financial crisis of 2007-08 featured such feedbacks. Therefore, a useful extension would be to introduce ambiguity and information processing constraints into the production-based asset pricing model of He and Krishnamurthy (2019). In addition to helping us understand asset prices during the crisis, perhaps the combination of ambiguity and financial constraints can help explain the depth and duration of the ensuing recession.

Another extension would be to combine our analysis with the Complex Asset Markets model of Eisfeldt, Lustig, and Zhang (2017). They also study the pricing implications of a model that combines ‘experts’ and ‘nonexperts’. In contrast to our model, where channel capacity differences are exogenous, agents in their model can choose whether to become experts. Becoming an expert is beneficial because it reduces idiosyncratic investment risk. The key mechanism in their model is endogenous entry and exit, and their induced selection effects. However, in their model funds cannot be reallocated across investors, and there is no trade in expertise. As a result, there is no moral hazard problem or capital constraint. Combining endogenous expertise, equilibrium entry and exit, portfolio delegation, and moral hazard would obviously be challenging, but also potentially quite fruitful.
References


Appendix

7.1. Proof for Robust Filtering Problem. The basic equations in the main text are (3.7) (3.8) and (3.9). Following the robust filtering literature, applying the change of measure from the approximating model $\mathbb{P}$ to the distorted model $\mathbb{Q}$ to the state transition and observation equations yields the following distorted filtering model:

\[
dg_t = [\rho_g (\tilde{g} - g_t) + \sigma_g v_{1,t}] dt + \sigma_g d\tilde{Z}_t^g, \tag{7.69}
\]
\[
d\frac{D_t}{D_t} = (g_t + \sigma v_{2,t}) dt + \sigma d\tilde{Z}_t^0, \tag{7.70}
\]
\[
ds_t = (g_t + \sigma s v_{3,t}) dt + \sigma_s d\tilde{Z}_t^s, \tag{7.71}
\]

where $\tilde{Z}_t^g$, $\tilde{Z}_t^0$, and $\tilde{Z}_t^s$ are Wiener processes that are related to the corresponding approximating processes:

\[
\tilde{Z}_t^g = Z_t^g - \int_0^t v_{1,s} ds, \quad \tilde{Z}_t^0 = Z_t^0 - \int_0^t v_{2,s} ds, \quad \text{and} \quad \tilde{Z}_t^s = Z_t^s - \int_0^t v_{3,s} ds
\]

and $v_{1,t}$, $v_{2,t}$, and $v_{3,s}$ are distortions to the conditional means of the three shocks, $\tilde{Z}_t^g$, $\tilde{Z}_t^0$, and $\tilde{Z}_t^s$, respectively. As shown in Pan and Başar (1996), Ugrinovskii and Petersen (2002), and Kasa (2006), a robust filter can be characterized by the following dynamic zero-sum game:

\[
L_t = \inf_{\{m_j\}} \sup_{\{Q\}} \left\{ \lim_{T \to \infty} \sup_{\{P\}} E^Q \left[ F - \theta_2^{-1} H_\infty (Q | P) \right] \right\}, \tag{7.72}
\]

subject to (7.69), (7.70), and (7.71), where $F \equiv \frac{1}{2} \int_0^T (g_j - \tilde{g}_j)^2 dj$ is the loss function and $H_\infty$ is the relative entropy and is bounded from above:

\[
H_\infty (Q | P) = \lim_{T \to \infty} \sup_{\{Q\}} E^Q \left[ \frac{1}{2T} \int_0^T (v_{1,t}^2 + v_{2,t}^2 + v_{3,t}^2) dt \right] \leq \eta_0, \tag{7.73}
\]

where $\eta_0$ defines the set of models that the consumer is considering, and $\theta_2^{-1}$ is the Lagrange multiplier on the relative entropy constraint, (7.73). As shown in Dai Pra, Meneghini, and Runggaldier (1996), the entropy constrained robust filtering problem, (7.72), is equivalent with the following risk-sensitive filtering problem:

\[
\frac{1}{\theta_2} \log \left( \int \exp (\theta_2 F (g, \tilde{g})) d\mathbb{P} \right) = \sup_{Q} \left\{ \int F (g, \tilde{g}) dQ - \theta_2^{-1} H_\infty (Q | \mathbb{P}) \right\}. \tag{7.74}
\]

The following proposition summarizes the results for this robust filtering problem:

**Proposition 7.1.** When $\theta_2 < 1/\sigma^2 + 1/\sigma_s^2$, there is a unique solution for the robust filtering problem, (7.74):

\[
d\hat{g}_t = \rho_g (\tilde{g} - \hat{g}_t) dt + \frac{Q_t}{\sigma} (\omega_t dt + d\tilde{Z}_t) + \frac{Q_t}{\sigma_s} d\tilde{Z}_t^s, \tag{7.75}
\]
\[
d\frac{Q_t}{dt} = \sigma_g^2 - 2\rho_g Q_t - Q_t^2 \left( \frac{1}{\sigma^2} + \frac{2\kappa}{Q_t} - \theta_2 \right), \tag{7.76}
\]
where \( Q_t = \mathbb{E}_t \left[ (g_t - \hat{g}_t)^2 \right] \) is the conditional variance of \( g_t \), \( d\tilde{Z}_t = dZ_t - \omega_t dt \), \( dZ_t = \frac{1}{\sigma} \left( \frac{dD_t}{D_t} - \hat{g}_t dt \right) \) and \( d\tilde{Z}_s = \frac{1}{\sigma_s} (ds_t - g_t dt) \) are innovations. In the steady state, the conditional variance converges to

\[
Q = - \left( \rho_g + \kappa \right) \sigma^2 + \sigma \sqrt{\left( \rho_g + \kappa \right)^2 \sigma^2 + \left( 1 - \theta_2 \sigma^2 \right) \sigma_\gamma^2} > 0. \tag{7.77}
\]

Proof. The right-hand side of (7.74) is an entropy constrained filtering problem in terms of the scaled quadratic objective function \( F \), while the left-hand side is a risk-sensitive filtering problem in terms of the same function \( F \), with risk-sensitivity parameter, \( \theta_2 \). The solution of this risk-sensitive filtering problem is just a special case of Theorem 3 of Pan and Başar (1996). Here we also use the result that \( Q_t / \sigma_\gamma^2 = 2 \kappa \) to replace \( 1 / \sigma_\gamma^2 \) with \( 2 \kappa / Q_t \).

![Figure 8: Steady state posterior variance and channel capacity](image)

7.2. 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, \( x_t = W_t / D_t \). Using Ito’s formula, we have

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

Substituting the dividend process (3.7) and derived specialist wealth process (7.84) into the above equation yields:

\[
\frac{dx_t}{x_t} = \left( \sigma^2 - \hat{g}_t - \rho + q_t + r_t + \frac{1}{\gamma^2} \left( \frac{\pi_{R,t}}{\sigma R_{t}} - \frac{\sigma \pi_{R,t}}{\gamma \sigma R_{t}} \right) \right) dt + \left( \frac{\pi_{R,t}}{\gamma \sigma R_{t}} - \sigma \right) dZ_t.
\]
The drift and diffusion coefficients of the aggregate state process, $\mu_{x,t}$ and $\sigma_{x,t}$, can be written as:

\[
\begin{align*}
\mu_{x,t} &= \sigma^2 \bar{g}_t - \rho + q_t + r_t + \left( \frac{\pi_{R,t}}{\gamma \sigma_{R,t}} \right)^2 - \sigma \left( \frac{\pi_{R,t}}{\gamma \sigma_{R,t}} \right) \\
&= \sigma^2 + \rho^h - \rho - \rho \left( \rho^h - \rho \right) x_t + (1 - \rho x_t) q_t + \frac{\left( \frac{\pi_{R,t}}{\gamma \sigma_{R,t}} \right)^2 (1 - 2 \rho x_t) + (3 \rho x_t - 1) \sigma \left( \frac{\pi_{R,t}}{\gamma \sigma_{R,t}} \right) - \sigma^2}{1 - \rho x_t} \\
\sigma_{x,t} &= \left( \frac{\pi_{R,t}}{\gamma \sigma_{R,t}} \right) - \sigma.
\end{align*}
\]

7.3. Solving the Agents’ Optimization Problems. Proof of Proposition 3.1. Solving first the infimization part yields $\nu_t^{h*} = -\theta^h \sigma_{W_t}^h V_w$ and $\omega_t^{h*} = -\theta^h \frac{Q_t^h}{\sigma} V_g$.

Optimal household consumption and portfolio rule under robustness are

\[ C_t^h = \frac{1}{V_w} \quad (7.78) \]

and

\[ \varepsilon_t^h = \frac{-V_w}{V_{ww} - \theta^h V_{w}^2} \frac{\left( \frac{\pi_{R,t}}{\gamma \sigma_{R,t}} - k_t \right)}{\sigma_{R,t}} \quad (7.79) \]

respectively. Guess the value function takes the following form:

\[ V \left( \hat{g}_t^h, Q_t^h, W_t^h, Y_t^h \right) = \frac{1}{\rho^h} \ln W_t^h + F^h \left( \hat{g}_t^h, Q_t^h \right) + Y_t^h, \]

where $Y_t^h$ is a function of aggregate state variable $x_t$. Now define $Y_t^h$ and $Y_t$ as a function of $x_t$:

\[
\begin{align*}
\frac{dY_t^h(x_t)}{dt} &= \mu_{Y,t}^h dt + \sigma_{Y,t}^h dZ_t, \\
\frac{dY_t(x_t)}{dt} &= \mu_{Y,t} dt + \sigma_{Y,t} dZ_t.
\end{align*}
\]

Using Ito’s formula, we have

\[
\begin{align*}
\mu_{Y,t}^h &= \left( Y_t^h(x_t) \right)' \mu_{x,t} x_t + \frac{1}{2} \left( Y_t^h(x_t) \right)'' \sigma_{x,t}^2 x_t^2, \\
\sigma_{Y,t}^h &= \left( Y_t^h(x_t) \right)' \sigma_{x,t} x_t, \\
\mu_{Y,t} &= Y_t'(x_t) \mu_{x,t} x_t + \frac{1}{2} Y_t''(x_t) \sigma_{x,t}^2 x_t^2, \\
\sigma_{Y,t} &= Y_t'(x_t) \sigma_{x,t} x_t.
\end{align*}
\]

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 expressions into the FOCs, (7.78) and (7.79), gives the consumption and portfolio rules in the main text.
7.4. Proofs of Propositions 4.4 and 4.5. Proof of Proposition 4.4. In the unconstrained case, per-unit exposure price is zero. Recall that the share of return contract \( \beta_t \equiv \varepsilon_t^I / \varepsilon_t^I \). Since the robust concern distorts the specialist’s desired risk exposure \( \varepsilon_t^I \), the choice of share contract turns into

\[
\beta_U^t = \frac{W_t}{W_t + \frac{\gamma}{\gamma} W_h^t} \quad \text{and} \quad k_t = 0.
\]

Now the specialist and household no longer hold the equity claims according to their wealth contributions as the benchmark case, but with a distortion term \( \frac{\gamma}{\gamma} \) which equals the inverse of distortion on the financial constraint. Note that although agency friction \( m \) doesn’t enter \( \beta_U^t \) in unconstrained region, both robustness parameters distort the contract share alternatively. Replacing \( W_h^t \) with asset market clearing condition (4.36) yields:

\[
\beta_U^t = 1 \frac{1}{1 + \frac{\gamma}{\gamma} \left( \frac{1}{\rho^h x_t} - \frac{\rho^h}{\rho^h} \right)}.
\]

By the imposed assumption that \( 0 < \beta_U^t \leq 1 \), \( x_t \) 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

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

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

\[
\beta_t = \frac{1}{1 + m} \quad \text{and} \quad k_t > 0.
\]

Proof of Proposition 4.5. In the unconstrained region, \( T^h_t = W^h_t \), both households and specialists put all their wealth into the intermediation, such that the total risk exposure equals \( \varepsilon_t^I = \alpha_t (W_t + W^h_t) \). The equilibrium conditions (4.34) and (4.36) yield \( \alpha_U^t = 1 \). Given that \( \varepsilon_t^* + \varepsilon_t^{h*} = W_t + W^h_t \), the risk exposure for the specialist can be derived as follows:

\[
\varepsilon_t^* + \frac{1 - \beta_U^t}{\beta_U^t} \varepsilon_t^* = P_t,
\]

which means that \( \varepsilon_t^{U*} = \frac{1}{1 + \frac{\gamma}{\gamma} \left( \frac{1}{\rho^h x_t} - \frac{\rho^h}{\rho^h} \right)} P_t \). From Figure 4.3, in the unconstrained region, \( \alpha_t = 1 \) such that the specialist invests all of the intermediary’s equity capital into the risky asset. Once the constraint is binding, \( \alpha_t > 1 \) means the specialist holds above 100% of the total equity and borrows \((\alpha_t - 1) (W_t + T^h_t)\) riskless bonds.

In the constrained region, the specialist holds \( \beta_t = \frac{1}{1 + m} \) share of risk. Hence, the specialist’s risk exposure is:

\[
\varepsilon_t^* = \beta_t \varepsilon_t^I = \frac{1}{1 + m} P_t.
\]
Furthermore, the specialist's portfolio share is:
\[
\alpha_t^* = \frac{\varepsilon_t^I}{W_t + T_t} = \frac{P_t}{(1 + \tilde{m})W_t} = \frac{1}{\varepsilon_t^I} + \rho_h - \rho.
\]

7.5. Solving for the Key Moments of Asset Prices. The return volatility. The return volatility can be derived from matching the diffusion terms of equation (3.7), (3.5), and (4.38) that
\[
\sigma_{R,t} = \frac{\sigma D_t}{\rho_h P_t - (\rho_h - \rho)\varepsilon_t^I}.
\]
Using Proposition 4.5 and 4.4, we have
\[
\sigma_{U,R,t} = \frac{\sigma}{\rho_h} \frac{1}{P_t/D_t} \frac{\rho(\gamma_h - \gamma) x_t + \gamma}{(\rho_h \gamma - \gamma)x_t + \gamma}.
\]
Thus,
\[
\sigma_{U,R,t} = \frac{\sigma}{1 + \Delta x t} \frac{\rho(\gamma_h - \gamma) x_t + \gamma}{\rho(\gamma_h - \gamma)x_t + \gamma}. \tag{7.81}
\]

The risk premium. Given that
\[
\pi_{R,t} = \frac{\gamma \sigma_{R,t}^2 \varepsilon_t^I}{W_t} = \frac{\gamma \sigma_{R,t}^2 \beta_t P_t}{W_t} = \frac{\gamma \sigma_{R,t}^2 \beta_t (P_t/D_t)}{x_t},
\]
we have
\[
\pi_{U,R,t} = \frac{\gamma \sigma_{U,R,t}^2 \beta_t (P_t/D_t)}{x_t} = \frac{\sigma^2 \gamma_h^2}{(1 + \Delta x t)} \frac{[\rho(\gamma_h - \gamma) x_t + \gamma]}{[\rho(\gamma_h - \gamma)x_t + \gamma]^2},
\]
in the unconstrained region. By contrast, in the constrained region,
\[
\pi_{R,t} = \frac{\gamma \sigma_{R,t}^2 \beta_t (P_t/D_t)}{x_t} = \frac{\sigma^2 \gamma_h^2}{x_t (1 + \Delta x t)} \frac{1 + m}{(m \rho_h + \rho)^2}. \tag{7.82}
\]

Solving the exposure price and intermediation fee. In the constrained region, \(k_t \geq 0\). When household desired exposure demand (3.18) equals specialist exposure supply (3.26), i.e., \(\varepsilon_t^{h*}(k_t) = m \varepsilon_t^I\), we have
\[
\pi_{R,t} - k_t \frac{\gamma_h \sigma_{R,t}^2 W_t}{\gamma^2 \sigma_{R,t}^2 W_t} = m - \frac{\pi_{R,t}}{\gamma^2 \sigma_{R,t}^2 W_t},
\]
which gives
\[
k_t = \left(1 - \frac{\tilde{m} \rho_h x_t}{1 - \rho x_t}\right) \pi_{R,t} = \frac{\sigma^2 (1 + m)}{\rho (m \rho_h + \rho)^2} \left(\frac{\gamma - \rho_h \gamma m x_t}{1 - \rho x_t}\right) \frac{\rho_h}{(1 + \Delta x t) x_t},
\]
where we use the fact that \(\tilde{m} = (\gamma_h / \gamma) m\).

Solving for the risk free rate. Using the household’s consumption function, we have
\[
\frac{dC_t^{h*}}{C_t^{h*}} = \frac{d(\rho_h W_t)}{\rho_h W_t^2} = \frac{d(P_t - W_t)}{P_t - W_t}. \tag{7.83}
\]
Combining it together with the distorted model (3.24) yields:

$$\frac{dW_t}{W_t} = (\phi_{W,t} + r_t) \, dt + \frac{\pi_{R,t}}{\gamma_{R,t}} \, dZ_t,$$

(7.84)

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

From Equations (4.36) and (4.37), we have:

$$\frac{d (P_t - W_t)}{P_t - W_t} = \frac{(g_t \, dt + \sigma dZ_t)D_t - \rho dW_t}{1 - \rho x_t} = \frac{(g_t \, dt + \sigma dZ_t) - \rho dW_t}{1 - \rho x_t} \left(\frac{\sigma - \rho x_t}{\gamma_{R,t}}\right)^2 dt,$$

(7.85)

where we use the facts that \( P_t - W_t = \frac{D_t}{\rho^n} - \frac{\rho^n}{\rho^n} W_t \) and \( P_t - W_t = \frac{(g_t \, dt + \sigma dZ_t)D_t - \rho dW_t}{\rho^n} \).

Taking conditional expectations and variances on both sides of the above equation gives:

$$\mathbb{E}_t \left[ \frac{d(P_t - W_t)}{P_t - W_t} \right] = \frac{\hat{g}_t - \rho x_t \left(\phi_{W,t} + r_t\right)}{1 - \rho x_t} \, dt \text{ and } \text{var}_t \left[ \frac{d(P_t - W_t)}{P_t - W_t} \right] = \left(\frac{\sigma - \rho x_t}{\gamma_{R,t}}\right)^2 dt.$$

(7.86)

Substituting Expressions (7.83) and (7.86) into the household’s Euler equation under the distorted model,

$$r_t \, dt = \rho^h \, dt + \mathbb{E}_t \left[ \frac{dC^h_t}{C^h_t} \right] - \text{var}_t \left[ \frac{dC^h_t}{C^h_t} \right],$$

yields the expression for the risk-free rate:

$$r_t = \rho^h + \frac{\hat{g}_t - \rho x_t \left(\phi_{W,t} + r_t\right)}{1 - \rho x_t} - \left(\frac{\sigma - \rho x_t}{\gamma_{R,t}}\right)^2 = r_t = \rho^h + \frac{\hat{g}_t - \rho \left(\phi_{W,t} + r_t\right)}{1 - \rho x_t} - \rho x_t \left(\frac{\xi_{R,t}}{\gamma_{R,t}}\right)^2.$$

Using the expressions for \( \pi_{R,t}/\gamma_{R,t} \) and \( q_t \) in the constrained and unconstrained regions by Propositions 4.9 and 4.13, we have

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

From Equation (4.36),

$$\frac{dW_t^h}{W_t^h} = \frac{d(P_t - W_t)}{P_t - W_t} = \frac{(g_t \, dt + \sigma dZ_t) - \rho \frac{dW_t}{W_t}}{1 - \rho x_t} = \left(1 - \frac{1}{1 - \rho x_t}\right) \frac{dW_t}{W_t} + \frac{1}{1 - \rho x_t} \frac{dD_t}{D_t},$$

which gives:

$$\frac{dW_t^h}{W_t^h} - \frac{dW_t}{W_t} = -\frac{1}{1 - \rho x_t} \left(\frac{dD_t}{D_t} + \frac{\sigma \pi_{R,t}}{\gamma_{R,t}} dt - \sigma^2 dt\right).$$
7.6. **Detection Error Probabilities.** The relative entropy computed from the household’s robust control problem can be written as:

\[
\nu_t^h = -\frac{\theta_t^h \sigma_{R,t} \varepsilon_t^h}{\rho^h W_t^h} = -\frac{\theta_t^h}{\rho^h \gamma^h} \frac{\pi_{R,t} - k_t}{\sigma_{R,t}} = 1 - \frac{\gamma^h}{\gamma} \left( \frac{\pi_{R,t} - k_t}{\sigma_{R,t}} \right). \tag{7.87}
\]

In the constraint case, we have:

\[
\frac{\pi_{R,t}}{\sigma_{R,t}} - \frac{k_t}{\sigma_{R,t}} = \tilde{m}_{\rho^h} \sigma_{\gamma}
\]

In the unconstrained case in which \(k_t = 0\), we have

\[
\frac{\pi_{R,t}}{\sigma_{R,t}} - k_t = \frac{\sigma_{\gamma}}{\rho (\gamma^h - \gamma)} x_t + \gamma.
\]

Substituting these expressions into (7.87) yields:

\[
\nu_t^h = \frac{1 - \gamma^h}{\gamma^h} \left[ \frac{\gamma^h m_{\rho^h} \sigma}{(1 - \rho) x_t} (m_{\rho^h} + \rho) \mathbf{1}_{x_t \in (x_{min}, x_c]} + \frac{\sigma_{\gamma}}{\rho (\gamma^h - \gamma)} x_t + \gamma \mathbf{1}_{x_t \in (x_c, \frac{1}{\rho})} \right].
\]

where \(\mathbf{1}\) denotes the indicator function.

Similarly, for the specialist’s problem, we have

\[
\nu_t = -\frac{\theta_{\rho^h} \pi_{R,t} \varepsilon_t}{\rho W_t} = -\frac{\theta_{\rho^h}}{\rho^h \gamma} \frac{\pi_{R,t}}{\sigma_{R,t}} = \frac{1 - \gamma}{\gamma} \frac{\pi_{R,t}}{\sigma_{R,t}},
\]

which implies that

\[
\nu_t = \sigma_{\gamma} (1 - \gamma) \left[ \frac{1}{\gamma^h (m_{\rho^h} + \rho) x_t} \mathbf{1}_{x_t \in (x_{min}, x_c]} + \frac{1}{\rho (\gamma^h - \gamma)} x_t + \gamma \mathbf{1}_{x_t \in (x_c, \frac{1}{\rho})} \right].
\]