Information Quality and Long-run Risks: Asset Pricing and Welfare Implications

Hengjie Ai[1]

Duke University

July, 2007

Abstract

I study asset pricing and welfare implications of the quality of public information about long-run risks in a general equilibrium model with a linear production technology and Kreps and Porteus (1978) utility. I show that learning could substantially improve the performance of the long-run risk asset pricing model in terms of its predictions for the risk premium and the volatility of the return to aggregate wealth. I quantify the effect of learning in a calibrated version of the model. I provide a novel decomposition of welfare gain of information and show that better information about long-run risks could lead to important welfare gain even if it does *not* improve the efficiency of resource allocation.

[1] I thank Ravi Bansal, V.V. Chari, James Choi, Bob Goldstein, Larry Jones, Narayana Kocherlakota, David Levine, Ellen McGrattan, Martin Schneider, Jan Werner, and Ofer Zeitouni for helpful comments. I especially thank Michele Boldrin for his continuous encouragement and support and Erzo Luttmer for numerous discussions on this topic. Financial support of the dissertation fellowship of University of Minnesota is gratefully acknowledged.

Introduction

In a recent study, Bansal and Yaron (2004) demonstrated the fundamental importance of long-run risks in providing a coherent account for many stylized empirical findings that have long challenged financial economists. However, long-run risks are difficult to identify statistically (Hansen and Sargent (2006)) and would therefore entail significant learning by investors. This paper studies the asset pricing and welfare implications of learning about long-run risks and makes two major contributions. First, I show that learning could substantially improve the performance of the long-run risk asset pricing model in terms of its predictions for the risk premium and the volatility of the return to aggregate wealth. Second, I provide a novel decomposition of welfare gain of information and show that better information about long-run risks could lead to important welfare gain even if it does *not* improve the efficiency of resource allocation.

In a general equilibrium model with linear production technology and Kreps and Porteus (1978) recursive preference, I obtain three main results regarding the effect of learning on equilibrium prices and quantities. First, lower information quality leads to *higher equity premium* if the representative agent's *relative risk aversion* (RRA) is higher than unity. This implies that learning increases the risk premium of aggregate wealth if RRA is higher than unity. The key to understanding this result is the fact that learning creates a positive covariance between the realized return and the expected return of the production technology (which in equilibrium equals the return to aggregate wealth): optimal Bayesian learning implies an upward revision of the agent's posterior belief about expected return whenever a high realized return is observed. The positive covariance induces a negative hedging demand when the agent's RRA is higher than unity, which in equilibrium must be met by an elevated risk premium (i.e., higher than the risk premium in the case without learning).

My second result is that learning decreases the *volatility of consumption growth* if the agent's *intertemporal elasticity of substitution* (IES) is higher than unity. To understand the intuition, notice that the variance of consumption growth can be decomposed into three components: the variance of the return to wealth; the variance of the consumption-to-wealth ratio; and the covariance of the return to wealth and the consumption-to-wealth ratio. While learning has negligible effects on the first two variance terms, it creates a negative covariance term when IES is higher than unity. The negative covariance can be explained by the consumption-to-wealth ratio effect as follows. When IES is higher than unity, the consumption-to-wealth ratio is a decreasing function of expected consumption growth (Bansal and Yaron (2004)), which is an increasing function of the expected return to wealth. Since learning induces a positive covariance between the realized and expected return to wealth, it induces a negative covariance between the return to wealth and the consumption-to-wealth ratio, decreasing the overall volatility of consumption growth.

My third result is that learning decreases the volatility of the risk-free interest rate. Fluctuations in the risk-free interest rate come from fluctuations in expected consumption growth. If the agent observes the persistent component of the return of the technology, then innovations in news about the persistent component of technology growth translate 100% into innovations in expected consumption growth. As information quality decreases, the information content in news becomes smaller, and the agent's belief about future consumption growth becomes less sensitive to innovations in news, making the risk-free rate less volatile.

I calibrate the model to the U.S. economy to assess the quantitative importance of each of the above channels and examine whether learning will reconcile the long-run risk asset pricing model with the empirical evidence on the statistics of the return to aggregate wealth provided in Lustig, Nieuwerburgh, and Verdelhan (2008). The long-run risk model without learning, although quite successful in accounting for many of the stylized facts in asset market data, implies too low a risk premium of aggregate wealth and too low a volatility of the return to aggregate wealth (Lustig, Nieuwerburgh, and Verdelhan (2008)). My calibrated benchmark model performs quite well along both dimensions. I show that learning could account for as much as 28% of the observed risk premium of aggregate wealth, and without learning, the volatility of consumption growth and the volatility of the risk-free interest rate would have been substantially higher than their historical levels.

In addition to contributing to the long-run risk literature by showing the importance of learning, my paper also contributes to two other large strands of literature. First is on the role of information quality and learning in general equilibrium asset pricing models.[2] In this regard, my paper is closely related to Veronesi (2000). In opposition to my finding, Veronesi (2000) finds that higher information quality leads to lower equity premium when RRA is higher than unity. The seemingly disparity results from two key differences between our models. First, Veronesi (2000) assumes CRRA utility function while I use the more general Kreps and Porteus (1978) utility function, which allows me to distinguish between the role of RRA and IES. I show that Veronesi (2000)'s result is driven by IES, and not RRA. Second, Veronesi (2000) studies a pure exchange economy, whereas I introduce production to the model. To see how adding production makes the difference, notice that, in general, learning affects the equilibrium prices and quantities through two channels: the *hedging demand* channel, the sign of which depends on the agent's *RRA*; and the *consumption-to-wealth ratio* channel, the sign of which depends on the agent's *IES*. In the linear production economy, learning affects equity premium through the hedging demand channel, and it affects the volatility of consumption growth through the consumption-to-wealth ratio channel. Therefore the relationship between information quality and equity premium depends on the RRA parameter. In pure exchange economies, quantity adjustment is not possible. The effect of learning through both channels must show up in price adjustment and affect equity premium. In the case of CRRA utility,

[2] A partial list includes Veronesi (2000), Brennan and Xia (2001), Brevik and D'Addona (2007), Gollier and Schlee (2006), Hansen and Sargent (2006), and Croce, Lettau, and Ludvigson (2007).

the consumption-to-wealth ratio channel dominates. Therefore, Veronesi (2000)'s result works through the consumption-wealth ratio channel and is determined by IES.

By providing a framework to assess the welfare consequences of information quality, my paper also relates to the literature on welfare cost of consumption risks pioneered by Lucas (1987). Tallarini (2000) and Croce (2006) discuss the welfare cost of consumption risks in economies with recursive utilities. The above studies focus on estimating the welfare gain from completely eliminating the uncertainty of the economy and, thus, provide an upper bound for the welfare gain from stabilization macroeconomic policies. However, consumption risks in these models are due to exogenous productivity shocks (or endowment shocks). The models themselves do not provide guidance for why and how these risks can be eliminated. In this paper, I ask a different, but related, question. What is the welfare gain of better information about the long-run risks of the economy? Realization of the welfare gain of information does not require the ability of the agents in the economy to change the nature of the risks, but only require a better understanding of the risks. The non-expected utility setting allows me to distinguish two sources of welfare gain of information: the welfare gain associated with *early resolution of uncertainty*, and the welfare gain associated with *resource allocation*. I provide precise definitions of the decomposition, and I show in the calibration that the total welfare gain of information can be more than 2.69% of permanent increase in annual consumption. Most of the welfare gain comes from early resolution of uncertainty, not from improvement on the supply side of the economy.

There is a tight connection between the asset pricing and the welfare implications of information about long-run risks. The high equity premium generated by long-run risks and the high welfare gain of information about long-run risks are both manifestations of the strong preference for early resolution of uncertainty assumed in long-run risk models. Therefore experimental evidence on individuals' preference for early resolution of uncertainty should be viewed as evidence on the plausibility of long-run risk models. My estimation of welfare gain from early resolution of uncertainty may be used in designing such experiments.

The paper is organized as follows: Section I describes the physical environment of the economy. Section II characterizes the optimal consumption allocation and studies the asset pricing implications of information quality. Section III studies the welfare implications of information about long-run risks. Section IV concludes.

I The Economy

Consider an infinite horizon economy populated by a continuum of investors with identical Kreps and Porteus (1978) SDU with constant RRA parameter γ, and constant IES parameter ψ. Specification of the SDU follows Duffie and Epstein (1992a) and Duffie and Epstein (1992b), who develop a continuous-time version of the recursive preference considered in Kreps and Porteus (1978), Epstein and Zin (1989) and Weil (1989). I use $\{C_t\}_{t\geq 0}$ to denote the consumption process and $\{V_t\}_{t\geq 0}$ to denote the utility process of the representative agent. I specify the SDU by a pair of aggregators $(f, \mathcal{A})$ such that:

$$dV_t = [-f(C_t, V_t) - \frac{1}{2}\mathcal{A}(V_t)||\sigma_V(t)||^2]dt + \sigma_V(t)dB_{Vt}, \tag{1}$$

for some standard Brownian motion $\{B_{Vt}\}_{t\geq 0}$. I adopt the convenient normalization $\mathcal{A}(V) = 0$, so that the aggregator $f(C, V)$ is given by:

$$f(C, V) = \frac{\beta}{1 - 1/\psi} \frac{C^{1-1/\psi} - ((1-\gamma)\, V)^{\frac{1-1/\psi}{1-\gamma}}}{((1-\gamma)\, V)^{\frac{1-1/\psi}{1-\gamma}-1}}. \tag{2}$$

In this paper, I focus on the case $\gamma \neq 1$. Extension to the limiting case $\gamma = 1$ is straightforward. However, I do allow $\psi = 1$ with the understanding that in this case,

$$f(C,V) = \beta(1-\gamma)V\left[\ln C - \frac{1}{1-\gamma}\ln[(1-\gamma)V]\right].$$

Under this convention, the date-t utility of the agent satisfies:

$$V_t = E_t\left[\int_t^\infty f(C_s, V_s)\,ds\right]. \tag{3}$$

Let $\{\Omega, \mathcal{F}^*, P\}$ be a complete probability space and $\mathcal{F} = \{\mathcal{F}_t\}_{t\geq 0}$ be the augmented, right-continuous filtration. Let $B_t = [B_{Kt}, B_{\theta t}, B_{et}]$ be a vector of standard Brownian motions adapted to $\mathcal{F}$. I allow B_K and B_θ to be correlated and denote

$$Corr(B_K, B_\theta) = \rho. \tag{4}$$

However, I assume without loss of generality that B_e is independent of $[B_K, B_\theta]$. Let $\{\theta_t\}_{t\geq 0}$ be an Ornstein-Uhlenbeck process adapted to $\mathcal{F}$; that is, $\{\theta_t\}_{t\geq 0}$ satisfies the following stochastic differential equation:

$$\forall t \geq 0, \quad d\theta_t = a\left(\overline{\theta} - \theta_t\right)dt + \sigma_\theta B_{\theta t}. \tag{5}$$

Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-filtration of $\mathcal{F}$ that represents the representative agent's information.

An aggregate consumption process $\{C_t\}_{t\geq 0}$ is *feasible* with respect to filtration $\mathcal{G}$ under the initial condition (k_0, θ_0) if it is $\mathcal{G}$ adapted; and there exists a capital stock process $\{K_t\}_{t\geq 0}$ such that $K_0 = k_0$, $\theta_0 = \theta_0$, $K_t \geq 0$ for all $t \geq 0$, and Equations (5) and (6) hold.

$$\forall t \geq 0, \quad dK_t = K_t\left[\theta_t dt + \sigma_K dB_{Kt}\right] - C_t dt. \tag{6}$$

Here θ_t is the expected return of the production technology. The realized return is determined by θ_t and a small noise represented by Brownian motion B_{Kt}, where $\sigma_K > 0$ is a constant.

There is an exogenous process $\{e_t\}_{t\geq 0}$ that carries noisy information about θ_t. $\{e_t\}_{t\geq 0}$ obeys the following ODE:

$$de_t = \theta_t dt + \sigma_e dB_{et}, \quad e_0 = 0. \tag{7}$$

I consider two alternative information structures: In a *completely observable economy*, $\mathcal{G} = \mathcal{F}$. In particular, this implies $\forall t \geq 0$, θ_t is observable to the agent. In a *partially observable economy*, $\mathcal{G} = \mathcal{G}^{K,e}$, where $\mathcal{G}^{K,e}$ is the information filtration generated by the $\{K_t, e_t\}_{t\geq 0}$ process. That is, $\mathcal{G}^{K,e} = \left\{\mathcal{G}_t^{K,e}\right\}_{t\geq 0}$, where $\forall t \geq 0$, $\mathcal{G}_t^{K,e}$ is the $\sigma-$field generated by the history $\{K_s, e_s\}_{s=0}^{t}$. In this case, the state variable θ_t is not directly observable by the agent. However, the agent does observe the $\{K_t, e_t\}_{t\geq 0}$ process and updates her belief about θ_t based on observations. In a partially observable economy, the agent's consumption policy must be $\mathcal{G}^{K,e}$ adapted, and cannot depend on θ. The parameter σ_e is a measure of information quality. If $\sigma_e = 0$, then the $\{e_t\}_{t\geq 0}$ process carries perfect information of $\{\theta_t\}_{t\geq 0}$ and therefore, observing $\{e\}_{t\geq 0}$ is equivalent to observing θ itself. If $\sigma_e > 0$, then $\{e_t\}_{t\geq 0}$ contains noisy information about $\{\theta_t\}_{t\geq 0}$. If $\sigma_e \to \infty$, $\{e_t\}_{t\geq 0}$ does not contain any information about $\{\theta_t\}_{t\geq 0}$, the agent can only update her belief about $\{\theta_t\}_{t\geq 0}$ from the observed $\{K_t\}_{t\geq 0}$ process.

The representative firm operates the technology described by Equations (5) and (6). Here I have a continuous-time version of the Bansal and Yaron (2004) economy with production. $\{\theta_t\}_{t\geq 0}$ is the source of long-run risks in this economy. Fluctuations of θ_t affect the productivity of the production technology, which generate persistent changes in consumption growth. The asset market is elementary. There is an equity, which is the claim to the representative firm. The date-t price of the equity is denoted by P_t. In the class of economies considered in this paper, the equity price satisfies:

$$dP_t = P_t \left[\mu_{Pt} dt + \sigma_{Pt} dB_{Pt}\right] \tag{8}$$

for some progressively measurable processes $\{\mu_{Pt}\}_{t\geq 0}$ and $\{\sigma_{Pt}\}_{t\geq 0}$, and some properly defined standard Brownian motion $\{B_{Pt}\}_{t\geq 0}$. There is a bond, which allows the agents in the economy to borrow from and lend to each other at a locally risk-free interest rate r_t. In the above economy, the equity pays aggregate consumption as its dividend. The cumulative return process of the equity is a diffusion process of the form:

$$\frac{dR_t}{R_t} = \mu_{Rt} dt + \sigma_{Pt} dB_{Pt}, \tag{9}$$

where $\mu_{Rt} = \ \mu_{Pt} + \frac{C_t}{P_t}$, and $\mu_{Rt} - r_t$ is the instantaneous risk premium of the stock.

The notion of competitive equilibrium is standard in this environment. In the following section, I deal directly with the Pareto optimality problem and appeal to the first and second fundamental welfare theorems to solve for the implied asset prices.

II Asset Pricing Implications

A The Pareto optimality Problem

If the representative agent observes $\{\theta_t\}_{t\geq 0}$, the Pareto optimality problem is to maximize the lifetime utility of the representative agent:

$$\begin{aligned} V(K,\theta) &= \max_{\{c_t\}_{t\geq 0}} E_0 \left[\int_0^\infty f\left(C_t, V_t\right) dt \right] \\ s.t.\ \ dK_t &= K_t[\theta_t dt + \sigma_K dB_{Kt}] - C_t dt,\ \ K_0 = K,\ \ K_t \geq 0,\ \ \forall t \\ d\theta_t &= a\left(\overline{\theta} - \theta_t\right) dt + \sigma_\theta B_{\theta t}, \qquad \theta_0 = \theta. \end{aligned} \tag{10}$$

Here the optimal consumption plan $\{C_t\}$ is chosen among all plans that are feasible with respect to $\mathcal{F}$ given initial condition (K, θ). In particular, this implies that the consumption plan is allowed to be a function of the history of $\{\theta_t\}_{t\geq 0}$. In the partially observable case, the Pareto optimality problem is still to maximize the same objective function as in (10), subject to the constraints (5) and (6). However I now require the optimal consumption plan to be feasible with respect to $\mathcal{G}^{K,e}$. I also assume that the agent's date-0 prior belief about θ_0 is a Gaussian distribution with mean m_0 and variance Q_0, denoted $N(m_0, Q_0)$.

The Pareto optimality problem in the partially observable economy can be solved by a two-step procedure.[3] The first step is a filtering problem, i.e., deducing the conditional distribution of θ given observations. If the prior distribution of θ_0 is Gaussian, then the conditional distribution of θ_t given $\mathcal{G}_t$ is Gaussian for all t. Therefore the conditional distribution of θ_t given $\mathcal{G}_t$ can be characterized by the first two moments, which I denote m_t and Q_t, where $m_t = E[\theta_t | \mathcal{G}_t]$ and $Q_t = Var[\theta_t | \mathcal{G}_t]$. m_t and Q_t satisfy the following (stochastic) differential equations (Liptser and Shiryaev (2001)):

$$dm_t = a\left(\overline{\theta} - m_t\right)dt + \left(\frac{1}{\sigma_K}Q_t + \rho\sigma_\theta\right)d\widetilde{B}_{Kt} + \frac{1}{\sigma_e}Q_t d\widetilde{B}_{et} \tag{11}$$

$$dQ_t = \left\{\sigma_\theta^2 - 2aQ_t - \left[\left(\rho\sigma_\theta + \frac{1}{\sigma_K}Q_t\right)^2 + \left(\frac{1}{\sigma_e}Q_t\right)^2\right]\right\}dt. \tag{12}$$

In Equation (11), $\widetilde{B}_{Kt}$ and $\widetilde{B}_{et}$ (usually called innovation processes) are defined recursively by:

$$d\widetilde{B}_{Kt} = \frac{1}{\sigma_K}\left[\frac{1}{K_t}(dK_t + C_t dt) - m_t dt\right]; \quad d\widetilde{B}_{et} = \frac{1}{\sigma_e}\left[de_t - m_t dt\right]. \tag{13}$$

In Equation (12), the law of motion of Q_t obeys a deterministic Riccati equation. As $t \longrightarrow \infty$, the posterior variance of θ_t, Q_t converges to a steady-state level, Q, where Q is

[3] A separation property applies here. The standard reference is Liptser and Shiryaev (2001).

given by:

$$Q = \frac{\left(1-\rho^2\right)\sigma_\theta^2}{\left(a+\rho\frac{\sigma_\theta}{\sigma_K}\right)+\sqrt{\left(a+\rho\frac{\sigma_\theta}{\sigma_K}\right)^2+\left(1-\rho\right)\sigma_\theta^2\left(\sigma_K^{-2}+\sigma_e^{-2}\right)}}. \tag{14}$$

I further assume that the conditional variance starts at its steady-state level, therefore $Q_t = Q$ for all t. Whenever necessary, I will use the notation $Q(\sigma_e)$ to emphasize the dependence of Q on the information quality parameter σ_e. The following properties of the $Q(\sigma_e)$ function will be useful:

$$Q(0) = 0; \quad \frac{\partial}{\partial \sigma_e} Q(\sigma_e) \geq 0. \tag{15}$$

It is convenient to denote

$$\sigma_m = \sqrt{\left(\frac{1}{\sigma_K}Q+\rho\sigma_\theta\right)^2+\left(\frac{1}{\sigma_e}Q\right)^2}, \tag{16}$$

and define

$$\tilde{B}_{mt} = \frac{1}{\sigma_m}\left[\left(\frac{1}{\sigma_K}Q+\rho\sigma_\theta\right)\widetilde{B}_{Kt}+\frac{1}{\sigma_e}Q\widetilde{B}_{et}\right]. \tag{17}$$

Therefore $\tilde{B}_m$ is a standard Brownian motion, and the law of motion of $\{m_t\}_{t\geq 0}$ can therefore be written as

$$dm_t = a\left(\overline{\theta}-m_t\right)dt+\sigma_m d\tilde{B}_{mt}.$$

The second step is a dynamic programming problem. By taking the posterior distribution of θ as state variables, the second-step problem can be made recursive and solved by standard dynamic programming techniques. Since I assume learning steady state, the conditional mean m_t is sufficient to keep track of the conditional distribution of θ_t, and, therefore, can serve as a state variable in the second stage dynamic programming problem. Using the definition of the innovation process in (13), the second stage optimization problem can now be written as:

$$
\begin{aligned}
V(K,m) &= \max_{c_t \geq 0} E_0 \left[\int_0^{\infty} f(C_t, V_t)\, dt \right] && (18) \\
s.t.\ dK_t &= K_t[m_t dt + \sigma_K d\tilde{B}_{Kt}] - C_t dt, \quad K_0 = K,\ K_t \geq 0. \\
dm_t &= a\left(\overline{\theta} - m_t\right) dt + \sigma_m d\tilde{B}_{mt}, \quad m_0 = m.
\end{aligned}
$$

The optimization problems in (10) and (18) are almost identical, except that the covariance between the innovations in the returns of the technology and the innovations in the state variables are different:

$$
Cov\left(\frac{dK_t}{K_t}, d\theta_t\right) = \rho\sigma_K\sigma_\theta \tag{19}
$$

$$
Cov\left(\frac{dK_t}{K_t}, dm_t\right) = \rho\sigma_K\sigma_\theta + Q. \tag{20}
$$

Intuitively, learning creates an additional positive covariance between the innovations in the realized return of the linear technology and the innovations in the expected return, m_t. Optimal Bayesian learning implies that whenever the realized return of the technology is high, the agent will revise upward her posterior belief about the expected return, creating a positive covariance between the two.

The similarity between (10) and (18) is the consequence of the observational equivalence result: for example, Proposition 6 in Veronesi (1999). Therefore, as far as its implications for asset prices and consumption volatility are concerned, the effect of learning could be interpreted as generating endogenously certain conditional distribution of the state variables. In the rest of this section, I will focus on the Pareto optimality problem in (18). Solution to the Pareto optimality problem of the completely observable economy can be obtained as the special case $\sigma_e = 0$. The following proposition characterizes the solution to the Pareto optimality problem:

Proposition 1 *The Pareto optimality Problem*

1) The value function of the Pareto optimality problem is of the form

$$V(K,m;\sigma_e) = H(m;\sigma_e)\frac{K^{1-\gamma}}{1-\gamma}, \tag{21}$$

where $H(m;\sigma_e)$ *satisfies the ODE in Equation (61) in Appendix A.* $H(m;\sigma_e)$ *is strictly increasing (decreasing) in* m *if* $\gamma < 1$ $(\gamma > 1)$.

2) The optimal consumption policy function is given by:

$$C(K,m;\sigma_e) = x(m;\sigma_e)K, \tag{22}$$

where $x(m;\sigma_e)$ *is the consumption-to-wealth ratio and satisfies*

$$x(m;\sigma_e) = \beta^{\psi} H(m;\sigma_e)^{\frac{1-\psi}{1-\gamma}}. \tag{23}$$

Furthermore, $x(m;\sigma_e)$ *is strictly increasing (decreasing) in* m *if* $\psi < 1$ $(\psi > 1)$.

Proof. See Appendix A. ■

The first part of the proposition implies that the value function is homogenous of degree $1-\gamma$. This is due to the fact that the utility function is homogeneous of degree $1-\gamma$, and that the constraint set is linearly homogenous in K. Since capital is the only factor of production in this economy, it also represents the aggregate wealth. Equation (6) implies that one unit of C_t can always be transformed freely into one unit of K_t. Therefore K_t equals aggregate wealth measured in terms of current period consumption goods, i.e. $P_t = K_t$, where P_t is the value of the stock (measured in current period consumption goods) that delivers aggregate consumption as its dividend payment, as defined in (8). The second part of Proposition 1 implies that the consumption-to-wealth ratio is an increasing (decreasing) function of the state if $\psi < 1$ $(\psi > 1)$. This has been pointed out by Bansal and Yaron (2004), among others, and can be explained by the interaction between the income effect and the substitution effect as follows: a higher value of m_t implies a higher expected consumption growth rate of the economy. This means that the agent is wealthier, and, other things being equal, will consume a greater proportion of her total wealth. This is the income effect and tends to increase the consumption-to-wealth ratio. On the other hand, a higher expected consumption growth rate implies a higher real interest rate; therefore, other things being equal, the agent will save more in order to raise future consumption. This is the intertemporal substitution effect and tends to decrease the consumption-to-wealth ratio. If $\psi < 1$, IES is small, and, therefore, the substitution effect is small and the income effect dominates. Consequently, the consumption-to-wealth ratio is increasing in expected consumption growth. By the same reason, if $\psi > 1$, the substitution effect dominates, and the consumption-to-wealth ratio is a decreasing function of expected consumption growth.

I obtain the value function for the case of observable θ by setting $\sigma_e = 0$ and $m = \theta$ in the above proposition. In this case the value function, denoted by $\bar{V}(\theta, K)$, and consumption policy function, denoted by $\bar{C}(K, \theta)$ satisfy:

$$\bar{V}(K,\theta) = H(\theta;0)\frac{K^{1-\gamma}}{1-\gamma}; \quad \bar{C}(K,\theta) = x(\theta;0)K, \tag{24}$$

where the functions $H(\theta;0)$ and $x(\theta;0)$ are given by Proposition 1.

B Asset Pricing Implications of Information Quality

The equity premium in the partially observable economy is given by the following proposition:

B.1 The Equity Premium

Proposition 2 *The Equity Premium*

The risk premium of the claim to aggregate wealth is given by:

$$\mu_{Rt} - r_t = \gamma \sigma_K^2 - \frac{H'(m_t;\sigma_e)}{H(m_t;\sigma_e)} \left(\rho \sigma_K \sigma_\theta + Q\right). \tag{25}$$

Proof. *See Appendix B.* ■

To see the effect of learning on equity premium, first consider a special case $\rho = 0$, that is, assume that the innovations in return and the innovations in the state variable θ are independent. If θ is observable, that is, $\sigma_e = 0$, then (15) implies the equity premium in this economy is[4]:

$$\mu_{Rt} - r_t = \gamma \sigma_K^2. \tag{26}$$

If, instead, θ is not observable, the risk premium is given by:

$$\mu_{Rt} - r_t = \gamma \sigma_K^2 - \frac{H'(m_t;\sigma_e)}{H(m_t;\sigma_e)} Q.$$

That is, learning creates an additional term $-\frac{H'(m_t;\sigma_e)}{H(m_t;\sigma_e)} Q$ in the equity premium. From Proposition 1, $H'(m;\sigma_e) < 0$ (> 0) if $\gamma > 1$ $(\gamma < 1)$, therefore learning *increases* equity premium if $\gamma > 1$, and *decreases* equity premium if $\gamma < 1$.

In general, the average equity premium in the economy is given by:

$$E\left[\mu_{Rt} - r_t\right] = \gamma \sigma_K^2 - \left(\rho \sigma_K \sigma_\theta + Q\right) E\left[\frac{H'(m;\sigma_e)}{H(m;\sigma_e)}\right], \tag{27}$$

where the expectation in (27) is taken with respect to the steady-state distribution of m. Numerical results indicate that the term $E\left[\frac{H'(m;\sigma_e)}{H(m;\sigma_e)}\right]$ almost does not change with

[4] Note that the Lucas (1978)-Breeden (1979) formula obtains in this case, despite the recursive prefer-ece. The intuition for this will be explained in section II.C of the paper.

σ_e. In fact, using the log-linear approximation method proposed in Campbell, Chacko, Rodriguez, and Viceira (2004), one can show (see Appendix B)

$$\frac{H'(m_t;\sigma_e)}{H(m_t;\sigma_e)} \approx \frac{1-\gamma}{a+\kappa_1}, \tag{28}$$

where κ_1 is a constant given in Appendix B. Therefore:

$$\frac{\partial}{\partial\sigma_e} E\left[\mu_{Rt} - r_t\right] \approx \frac{\gamma-1}{a+\kappa_1}\frac{\partial}{\partial\sigma_e} Q\left(\sigma_e\right). \tag{29}$$

That is, equity premium increases (decreases) as the noise contained in the $\{e_t\}_{t\geq 0}$ process increases if $\gamma > 1$ ($\gamma < 1$).

The above result can be illustrated as follows. From Merton (1992), the demand for equity of an agent with CRRA utility is determined by:

$$\frac{1}{\gamma\sigma_{Rt}^2}(\mu_{Rt} - r_t) + \textit{Hedging Demand}, \tag{30}$$

where μ_{Rt} and σ_{Rt} denote the expectation and volatility of the return of the equity, respectively. The first term in (30) is the myopic demand and the second term is the hedging demand. This result is still true in the case of Kreps and Porteus (1978) utility. Moreover, because aggregate wealth is the claim to the asset that pays the equilibrium consumption as dividend, the cumulative return process of aggregate wealth defined in (9), can be written as:

$$\frac{dR_t}{R_t} = \frac{dK_t}{K_t} + C_t dt = m_t dt + \sigma_K d\tilde{B}_{Kt}. \tag{31}$$

In particular, (31) implies $\sigma_{Rt} = \sigma_K$ holds true regardless of the information structure. In equilibrium, market clearing condition requires that the total demand has to sum up to 1:

$$1 = \frac{1}{\gamma\sigma_K^2}(\mu_{Rt} - r_t) + \textit{Hedging Demand}. \tag{32}$$

Consider first the case in which hedging demand is 0. This will happen if the realized return of the equity and the expected return of wealth are uncorrelated, i.e.,

$$Cov_t \left[\frac{dR_t}{R_t}, dm_t \right] = \rho \sigma_K \sigma_\theta + Q = 0.$$

In this case (32) implies the equity premium must be the same as in (26).

In general, the second term in (25) reflects the equilibrium price adjustment for hedging demand. Hedging demand of the equity depends on the covariance between the innovations in the expected return of the equity and the innovations in the return to aggregate wealth, i.e. $Cov_t \left[\frac{dR_t}{R_t}, dm_t \right]$. If RRA is higher (lower) than 1, the agent prefers a negative (positive) covariance between the two. The market clearing condition in (32) implies that risk premium paid by the equity must adjust to compensate for the hedging demand. If θ is observable, the equilibrium risk premium will contain a hedging demand adjustment term as long as $\rho \neq 0$:

$$\mu_{Rt} - r_t = \gamma \sigma_K^2 - \rho \sigma_K \sigma_\theta \frac{H'(\theta_t)}{H(\theta_t)}. \tag{33}$$

As shown in Equations (19) and (20), if θ is not observable, learning creates an additional positive covariance between innovations in the return to aggregate wealth and innovations in the expected return to aggregate wealth. The positive covariance induces a negative (positive) hedging demand if $\gamma > 1$ ($\gamma < 1$). In equilibrium, expected return of the equity has to adjust to equate supply and demand (See the market clearing condition in (32)); therefore, learning creates a higher (lower) equity premium than the completely observable case if $\gamma > 1$ ($\gamma < 1$), as shown in Equation (29).

B.2 Volatility of Consumption Growth and Volatility of Risk-free Interest Rate

The high equity premium, the low volatility of consumption growth, and the low volatility of the risk-free rate observed in the data are among the most important challenges to consumption-based asset pricing models. In Section II.B.1, I provide conditions under which learning increases or decrease the equity premium. This section is devoted to the analysis of the relationship between information quality and the volatility of consumption growth and the volatility of the risk-free interest rate.

First, learning decreases (increases) the conditional volatility of consumption growth if $\psi > 1$ ($\psi < 1$). Using Proposition 1,

$$C_t = \beta^{\psi} H(m_t; \sigma_e)^{\frac{1-\psi}{1-\gamma}} K_t. \tag{34}$$

By Ito's formula,

$$Var_t(\ln C_t) = \left(\frac{1-\psi}{1-\gamma}\frac{H'(m;\sigma_e)}{H(m;\sigma_e)}\right)^2 \sigma_m^2 + \sigma_K^2 + 2\left(\rho\sigma_K\sigma_\theta + Q\right)\frac{1-\psi}{1-\gamma}\frac{H'(m;\sigma_e)}{H(m;\sigma_e)}.$$

Using the definition of σ_m in (16), and Equation (14), I obtain:

$$\sigma_m^2 = \sigma_\theta^2 - 2aQ.$$

Using the log-linear approximation in (28) to simplify Equation (34), I have:

$$Var_t(\ln C_t) \approx \left(\frac{1-\psi}{a+\kappa_1}\right)^2 \left(\sigma_\theta^2 - 2aQ\right) + \sigma_K^2 + 2\left(\rho\sigma_K\sigma_\theta + Q\right)\frac{1-\psi}{a+\kappa_1}.$$

Therefore

$$\frac{\partial}{\partial\sigma_e} Var_t(\ln C_t) \approx 2(1-\psi)\frac{a\psi+\kappa_1}{(a+\kappa_1)^2}\frac{\partial}{\partial\sigma_e}Q(\sigma_e). \tag{35}$$

By (15), the conditional volatility of consumption growth is decreasing (increasing) in σ_e if $\psi > 1$ ($\psi < 1$).

The intuition of the above conclusion can be explained as follows. Since $x(m_t; \sigma_e)$ is the consumption-to-wealth ratio,

$$\ln C_t = \ln x(m_t; \sigma_e) + \ln K_t. \tag{36}$$

Therefore, I have the following variance decomposition:

$$Var_t (\ln C_t) = Var_t (\ln x(m_t; \sigma_e)) + Var_t (\ln K_t) + 2Cov_t (\ln x(m_t; \sigma_e),\ \ln K_t). \tag{37}$$

The effect of σ_e on consumption-to-wealth ratio function $x(m_t; \sigma_e)$ is virtually zero, and $Var_t (\ln K_t) = \sigma_K^2$ does not change with σ_e. Consequently, the information quality parameter σ_e affects the volatility of consumption growth primarily through the third term in (37), which captures the covariance between the consumption-to-wealth ratio and the return to wealth. As discussed earlier, learning creates a positive covariance between the innovations in return to wealth and the innovations in the posterior belief m_t. By Proposition 1, if $\psi > 1$ ($\psi < 1$), then the consumption-to-wealth ratio function $x(m_t; \sigma_e)$ is decreasing (increasing) in m. Consequently, learning decreases (increases) the term $Cov_t (\ln x(m_t; \sigma_e),\ \ln K_t)$ if $\psi > 1$ ($\psi < 1$).

Next, learning decreases the volatility of the risk-free interest rate. To understand this result, one can use Equations (25) and (31) to write the risk-free interest rate as:

$$r_t = m_t - \gamma \sigma_K^2 + \frac{H'(m; \sigma_e)}{H(m; \sigma_e)} (\rho \sigma_K \sigma_\theta + Q) \approx m_t - \gamma \sigma_K^2 + \frac{1-\gamma}{a + \kappa_1} (\rho \sigma_K \sigma_\theta + Q),$$

where the approximate equality uses the log-linearization approximation in (28). Therefore

$$Var_t\left(dr_t\right) \approx Var_t\left(dm_t\right) = \sigma_m^2 = \sigma_\theta^2 - 2aQ,$$

and consequently,

$$\frac{\partial}{\partial \sigma_e} Var_t\left(dr_t\right) = -2a \frac{\partial}{\partial \sigma_e} Q\left(\sigma_e\right) < 0.$$

Intuitively, fluctuations in the interest rate in this economy come from fluctuations in expected consumption growth, which in turn come from fluctuations in the posterior belief m_t. If information is imprecise, then the revision of the posterior belief m_t is small when new information arrives. In the extreme case where there is absolutely no new information, m_t would be a constant: the agent does not change her belief at all. In the model, the $\{K_t\}_{t\geq 0}$ process always carries some information about $\{\theta_t\}_{t\geq 0}$, therefore, the volatility of interest rate achieves its minimum when $\sigma_e = \infty$. When information quality gets better, the posterior mean becomes more sensitive to the arrival of new information. In fact, when θ_t is known, m_t moves one on one with movements in θ_t, and $Var_t\left(dr_t\right) \approx \sigma_\theta^2$. This is the case in which the volatility of m_t, and consequently the volatility of the risk-free interest rate is maximized.

To summarize, my results imply that in an economy with linear production technology and SDU, the direction in which learning affects equity premium is completely determined by the RRA parameter, while the direction in which learning affects volatility of consumption growth is completely determined by the IES parameter. I also show that learning decreases the volatility of the risk-free interest rate regardless of the preference parameters. My results are similar in spirit to Tallarini (2000)'s finding in the sense that the quantity implications of real business cycle models are primarily determined by IES, and the risk premium of these models are primarily determined by RRA.

B.3 Comparison with Pure Exchange Economies

The relationship between information quality and equity premium has been studied by Veronesi (2000), Brennan and Xia (2001), Brevik and D'Addona (2007), among others, in pure exchange economies. In pure exchange economies with CRRA utility, Veronesi (2000) establishes the result that learning leads to *lower* (higher) equity premium if RRA is *higher* (lower) than unity, which is the opposite of my result in Equation (29). Here I briefly compare the intuitions in my production economy with those in the pure exchange economies.

As discussed in Sections II.B.1 and II.B.2, learning affects both the agent's hedging demand for equity and her optimal consumption-to-wealth ratio. The direction of the first effect depends on RRA and the direction of the second effect depends on IES. In the linear production economy considered in this paper, the first effect shows up in equity premium, and the second effect shows up in volatility of consumption growth. In other words, learning affects equilibrium price through the hedging demand channel and affects equilibrium quantity through the consumption-to-wealth ratio channel.

In a pure exchange economy, however, quantity cannot adjust. Consequently learning affects the equity premium through *both* channels. Veronesi (2000)'s result works through the consumption-to-wealth ratio channel, and therefore it is driven by the agent's IES, not RRA. The Kreps and Porteus (1978) utility allows me to disentangle the agent's IES and RRA, and gain a better understanding of the effect of learning on equity premium.

To see the above more clearly, consider a pure exchange economy in which the representative agent has the Kreps and Porteus (1978) SDU as specified in (3). The endowment process of the economy is assumed to be

$$dC_t = C_t \left[\theta_t dt + \sigma_C dB_{Ct}\right], \tag{38}$$

where $\{\theta_t\}_{t\geq 0}$ follows the same process as described in (5). Again, assume

$$Corr\left(B_C, B_\theta\right) = \rho.$$

If the agent does not observe $\{\theta_t\}_{t\geq 0}$, she must update her belief based on the observed consumption process and an additional source of information denoted by $\{e_t\}_{t\geq 0}$, where

$$de_t = \theta_t dt + \sigma_e dB_{et}.$$

I again assume B_e is independent of $[B_C, B_\theta]$. The posterior mean of θ obeys the following SDE (again, assuming learning steady state for simplicity):

$$dm_t = a\left(\overline{\theta} - m_t\right)dt + \left(\frac{1}{\sigma_K}Q_E + \rho\sigma_\theta\right)d\widetilde{B}_{Ct} + \frac{1}{\sigma_e}Q_E d\widetilde{B}_{et},$$

where $\widetilde{B}_{Ct}$ and $\widetilde{B}_{et}$ are innovation processes defined by:

$$d\widetilde{B}_{Ct} = \frac{1}{\sigma_C}\left[\frac{dC_t}{C_t} - m_t dt\right], \quad d\widetilde{B}_{et} = \frac{1}{\sigma_e}\left[de_t - m_t dt\right].$$

and Q_E is the steady-state posterior variance of θ in the exchange economy. Using the logic in (30), the risk premium of aggregate wealth is given by:

$$\mu_{Rt} - r_t = \gamma\sigma_{Rt}^2 + Hedging\ Demand\ Adjustment. \tag{39}$$

In the above economy, learning affects both the myopic demand component and the hedging demand component of the equity premium.

First, learning increases (decreases) the myopic demand component (the first term in (39)) of equity premium by increasing (decreasing) the volatility of return to wealth if $\psi > 1$ ($\psi < 1$). To see this, using log-linear approximation around $\psi = 1$, one can show (details in Appendix C):

$$\frac{\partial}{\partial \sigma_e}\left[\gamma \sigma_{Rt}^2\right] \approx 2\gamma\left(1-\frac{1}{\psi}\right)\frac{(\varpi_1 + a/\psi)}{(a+\varpi_1)^2}\frac{\partial Q}{\partial \sigma_e}, \tag{40}$$

where $\varpi_1 > 0$ is a constant defined in Appendix C. By (15), it is clear that the myopic demand component of the equity premium is increasing in σ_e if $\psi > 1$.

The intuition for this result can be explained by the consumption-to-wealth ratio channel discussed above. Denote aggregate wealth by W_t, and consumption-to-wealth ratio by x_t. The definition of consumption wealth ratio, $x_t = C_t/W_t$, implies the following variance decomposition:

$$Var_t(\ln W_t) = Var_t(\ln C_t) + Var_t(\ln x_t) - 2Cov_t(\ln C_t,\ \ln x_t). \tag{41}$$

As in equation (37), changes in σ_e affect the variance of return to wealth mainly through the third term in (41). Learning creates a positive covariance between innovations in consumption growth and innovations in expected consumption growth m. If $\psi > 1$, then x is a decreasing function (Proposition 1), and, therefore, learning makes the term $Cov_t(\ln C_t,\ \ln x_t)$ negative. Consequently, the variance of the return to wealth increases because of learning. Of course if $\psi < 1$, the same argument implies that learning decreases the volatility of return to wealth. This explains the effect of learning on the myopic demand component of the equity premium in (40). In Equation (41), I use a variance decomposition different from Equation (37), because in pure exchange economies, volatility of consumption growth is exogenous and information quality affects only the volatility of return to wealth. In the linear production economy that I focus on in this paper, volatility of consumption growth is determined endogenously, and volatility of return to wealth is completely determined by the exogenous technology.

Second, learning increases the hedging demand component of equity premium if $\gamma > 1$

($\gamma < 1$). Again, using log-linear approximation around $\psi = 1$, one can show (see Appendix C for details.)

$$\frac{\partial}{\partial \sigma_e} \left[Heding\ Demand\ Component \right] \approx \frac{1}{\psi} \left(\gamma - 1 \right) \frac{\left[a \left(1/\psi - 1 \right) + \kappa_1 \right]}{\left(a + \varpi_1 \right)^2} \frac{\partial Q}{\partial \sigma_e}. \tag{42}$$

Therefore, for ψ close to 1, the hedging demand component of equity premium is increasing in σ_e if $\gamma > 1$. The intuition for this is the same as discussed earlier in Section II.B.1.

In the case of CRRA preference, $\gamma = \frac{1}{\psi}$, the effect on the myopic demand component of equity premium always dominates. (One can see this by letting $\gamma = \frac{1}{\psi}$ and comparing (40) with (42).) This is why Veronesi (2000) obtained the result that lower information quality decreases (increases) equity premium if $\gamma > 1$ ($\gamma < 1$). The effects of learning on equity premium are summarized in Table 1. In particular, the following observations are true:

1. In general, learning increases (decreases) the myopic demand component of equity premium if $\psi > 1$ ($\psi < 1$), and increases (decreases) the hedging demand component of equity premium if $\gamma > 1$ ($\gamma < 1$). The overall effect depends on the relative magnitude of the two.

2. CRRA utility (i.e. $\gamma = \frac{1}{\psi}$) is a special case in which the effect on the myopic demand component of equity premium always dominates. In this case, learning decreases equity premium if $\psi < 1$ or equivalently, $\gamma > 1$.

3. A sufficient condition for learning to increase equity premium in general is $\gamma > 1$ and $\psi > 1$.

(Insert Table 1 about here.)

In the case of CRRA, the first-order effect of learning works through the consumption-to-wealth ratio channel and affects the myopic demand component of equity premium.

In the general case of Kreps and Porteus (1978) utility, the direction of this effect is determined by the agent's IES parameter, not the RRA parameter. In this sense, Veronesi (2000)'s result is driven by the IES parameter, not the RRA parameter.

In general, both the pure exchange economy and the linear production economy can be viewed as special cases of models with adjustment cost. Pure exchange assumes infinite adjustment cost, and the linear production economy considered here has zero adjustment cost. Adding adjustment cost sacrifices analytical tractability of the model. The above discussion suggests that in models with adjustment cost, learning affects equity premium through both the hedging demand channel and the consumption-to-wealth ratio channel. Although in general the two channels might work in different directions and the total effect is ambiguous, result 3 above is still true. That is, learning will increase equity premium if both the RRA and IES parameter are higher than unity.[5] The simple setting of linear technology allows me to highlight the hedging demand channel and make unambiguous predictions about the effect of learning on equity premium.

C Calibration

The purpose of this section is to evaluate the quantitative importance of learning in accounting for some of the key asset pricing statistics in the data. Although the long-run risk model developed by Bansal and Yaron (2004) provides a successful explanation of many salient features of the asset market data, the risk premium and the volatility of the return to aggregate wealth implied by the model is too low relative to the empirical evidence presented in Lustig, Nieuwerburgh, and Verdelhan (2008). My benchmark model substantially improves the performance of the Bansal and Yaron (2004) model and matches the empirical findings of Lustig, Nieuwerburgh, and Verdelhan (2008) quite well. I set $\sigma_e = \infty$ in the benchmark model to maximize the impact of learning. I then de-

[5] Proof of this general case with adjustment cost is available from the author upon request.

crease σ_e and measure the quantitative changes in the model's predictions for the key asset pricing statistics.

I choose the preference parameters in line with the long-run risk literature. I calibrate the discount factor $\beta = 0.0215$. This is equivalent to an annual discount rate of 0.98 in discrete time models. I choose $\gamma = 9$, $\psi = 2.0$. These values are close to those used in the long-run risk literature.[6]

The technology parameters are chosen as follows:

σ_K	a	$\bar{\theta}$	σ_θ
0.053	0.15	0.043	0.0085

. (43)

I rely on the estimation presented in Lustig, Nieuwerburgh, and Verdelhan (2008) to discipline the choice of $\bar{\theta}$ and σ_K. The parameter $\bar{\theta}$ determines the mean of return to aggregate wealth. Lustig, Nieuwerburgh, and Verdelhan (2008)'s point estimate of the mean equity premium of aggregate wealth is 3.52% per annum.[7] This estimation, together with an historical average risk-free interest rate of 0.86% (Table 3) implies an annualized return on aggregate wealth of 4.38%. I choose $\bar{\theta} = 0.043$. Equation (31) implies:

$$\lim_{\Delta \to 0} E\left[\frac{R_{t+\Delta}}{R_t} - 1\right] = e^{\bar{\theta}} - 1 = 4.39\%. \tag{44}$$

The discrete-time annualized average return generated by my model is 4.38%, almost the same as the continuous-time limit in (44). The estimated volatility of return to aggregate wealth by Lustig, Nieuwerburgh, and Verdelhan (2008) is 6% at annual level. With $\sigma_K = 0.053$, Equation (31) implies that the instantaneous volatility of return to

[6]Bansal and Yaron (2004) use $\gamma = 10$ and $\psi = 1.5$. Bansal, Kiku, and Yaron (2007) argue γ is close to 10 and ψ is close to 2 based on an estimation using the Euler equation-GMM method.

[7]The estimations in Lustig, Nieuwerburgh, and Verdelhan (2008) are done at the quarterly frequency and are annualized in this paper.

aggregate wealth is:

$$\lim_{\Delta\to 0} Std\left[\frac{R_{t+\Delta}}{R_t}-1\right]=\sigma_K=5.3\%.$$

The annualized volatility of return to wealth is also affected by the fluctuations in θ and is, therefore, higher. This brings the volatility of return to wealth in my model to 5.69%, close to the estimation in Lustig, Nieuwerburgh, and Verdelhan (2008).

Because of the statistical difficulty of estimating the long-run risk parameters, there is little independent empirical evidence I could use to discipline the choice of a and σ_θ. Here I choose $a = 0.15$, so that the monthly autocorrelation of the long-run risk state variable is approximately $e^{-\frac{a}{12}} = 0.987$, close to the 0.979 used in the discrete time model in Bansal and Yaron (2004). I view σ_θ as a free parameter and choose σ_θ to maximize the model's ability to match the rest of the asset pricing and consumption growth statistics in Table 2, namely, the first two moments of annual consumption growth, the first two autocorrelation functions of annual consumption growth, and the first two moments of the risk-free interest rate. A priori, there is no apparent reason to believe that σ_θ could be chosen such that all the above moments could be matched. However, as shown in Table 2, my model comes remarkably close.

Table 2 compares the moments of consumption growth, risk-free interest rate, and return to wealth in the data, with those generated by Bansal and Yaron (2004)'s model and those generated by my benchmark model. Data moments of consumption growth and the risk-free interest rate are taken from Bansal and Yaron (2004). Data moments of the return to aggregate wealth are taken from the estimation in Lustig, Nieuwerburgh, and Verdelhan (2008). Table 2 shows that the consumption growth statistics and moments of risk-free interest rate in my model are very close to their empirical counterparts, and to those produced in the Bansal and Yaron (2004) model. However, my model produces a significantly higher equity premium and volatility of the return to aggregate wealth.

(Insert Table 2 about here.)

My procedure of calibrating the technology parameters differs from that used in most of the production-based asset pricing literature, for example, Rouwenhorst (1995), Jermann (1998) and Boldrin, Christiano, and Fisher (2001). It is a well-known fact that incorporating production presents additional challenges for general equilibrium models to generate a realistic equity premium (Rouwenhorst (1995)). One of the reasons is that the physical capital in the data is very smooth. In order to generate large fluctuations in the equity price, one has to rely on adjustment costs as in Jermann (1998) or real rigidity as in Boldrin, Christiano, and Fisher (2001). However, since I do not explicitly model labor in my model, K_t should *not* be interpreted as physical capital. It should be interpreted as aggregate wealth, the bulk of which is human capital. This is the reason I use the empirical estimates of moments of return to aggregate wealth to discipline my choice of the technology parameters.

The keys to understanding why my model delivers better results on the statistics of return to aggregate wealth are the role of learning and the choice of the parameter σ_θ. First, learning and a large σ_θ keep the volatility of consumption growth low. In my calibration, $\frac{\sigma_\theta}{\sigma_K} = 0.16$. The discrete-time, pure-exchange-economy counterpart of this number in Bansal and Yaron (2004) is 0.044, which is much lower. Note that I choose $\sigma_K = 0.53$ to deliver the high volatility of return to aggregate wealth. Without learning, this would imply a volatility of consumption growth of roughly the same magnitude, which is much higher than the 2.9% documented in Bansal and Yaron (2004). In fact, as shown in Table 3, the volatility of consumption growth in an otherwise identical model without learning is 5.9% per annum. As explained in Section II.B.2, large fluctuations in θ cause considerable movement in the consumption-to-wealth ratio, which offsets the large volatility of return to aggregate wealth because of learning, making consumption growth much less volatile.

Second, learning and a large σ_θ generate a large equity premium. The high volatility of θ implies that the continuation utility of the agent is volatile, and, therefore, so is the pricing kernel. Learning creates a high exposure of the return to aggregate wealth to fluctuations in the pricing kernel, creating a high level of equity premium (to be precise, a high level of premium for long-run risks, as I will explain later in this section).

Finally, learning dampens the fluctuations in expected consumption growth, and keeps the volatility of the risk-free rate low. Note that a high level of σ_θ implies that the predictable component of consumption growth is volatile, which, without learning, will lead to a counterfactually high volatility of the risk-free interest rate. As explained earlier, learning also helps smooth the risk-free interest rate, so the volatility of risk-free interest rate in my model remains low.

To quantify the importance of learning, I compare the predictions of the model for various choices of the information quality parameter σ_e. Table 3 documents the moments of consumption growth, the equity premium of aggregate wealth, and the risk-free interest rate generated by the model for various levels of σ_e. $\sigma_e = 0$ corresponds to the completely observable case, and $\sigma_e = \infty$ corresponds to the benchmark partially observable model.

(Insert Table 3 about here.)

As I decrease σ_e from ∞ to 0, the mean consumption growth rate increases slightly. The volatility of consumption growth increases from 3.15% to 5.9%. At the same time, the equity premium of aggregate wealth decreases from 3.50% to 2.53%, by 28%. Furthermore, the mean risk-free interest rate increases from 0.87% to 1.85%, which echoes the change in the risk premium. The standard deviation of the risk-free interest rate increases from 0.59% to 1.40%. These all confirm the theoretical results obtained in Sections II.B.1 and II.B.2.

To better understand the effect of information quality on equity premium, it is useful to decompose the equity premium into *premium for long-run risks* and *premium for short-run risks*. The decomposition follows Bansal and Yaron (2004) and Bansal (2008).[8]

[8] Bansal and Yaron (2004) and Bansal (2008) use discrete time models. Their decomposition relies on

I briefly explain the decomposition in the context of my continuous time model. The consumption process of the economy follows a diffusion process of the following form:

$$dC_t = C_t \left[\mu_{Ct} + \sigma_{Ct} dB_{Ct}\right], \tag{45}$$

where B_C is some appropriately defined Brownian motion. Using Proposition 1, one can write the continuation utility of the representative agent as

$$V_t = \frac{1}{1-\gamma} \beta^{\psi(\gamma-1)} H(m_t)^{\psi} C_t. \tag{46}$$

Using Equations (45), (46), and (62) in Appendix B, the state price process can be written as:

$$d\pi_t = \pi_t \left[-r_t dt - \gamma \sigma_{Ct} dB_{Ct} - \psi \frac{\gamma - \frac{1}{\psi}}{1-\gamma} \frac{H'(m_t;\sigma_e)}{H(m_t;\sigma_e)} \sigma_m d\widetilde{B}_{mt} \right],$$

where r_t is the risk-free interest rate. Therefore the risk premium of an asset with cumulative return process given in (9) is

$$\mu_{Rt} - r_t = -Cov_t\left(\frac{dR_t}{R_t}, \frac{d\pi_t}{\pi_t}\right) = \gamma \rho_{C,P} \sigma_{Ct} \sigma_{Pt} + \psi \frac{\gamma - \frac{1}{\psi}}{1-\gamma} \frac{H'(m_t;\sigma_e)}{H(m_t;\sigma_e)} \rho_{m,P} \sigma_m \sigma_{Pt}, \tag{47}$$

where

$$\rho_{C,P} = Corr\left(B_C, B_P\right); \quad \rho_{m,P} = Corr\left(\widetilde{B}_m, B_P\right).$$

Note the term $\gamma \rho_{C,P} \sigma_{Ct} \sigma_{Pt}$ in Equation (47) is the Lucas (1978)-Breeden (1979) formula for equity premium. That is, it is the equity premium that would obtain with expected utility. This term depends only on the instantaneous covariance between the innovations in consumption growth ($\frac{dC_t}{C_t}$) and the innovations in return ($\frac{dR_t}{R_t}$). I define this part of risk premium as *premium for short-run risks*. Intuitively, premium for short-run

the approximation results of Campbell and Shiller (1988). As shown in this paper, the decomposition is exact in continuous time models.

risks is the compensation for the asset's exposure to the risks associated with fluctuations in the short-run component of consumption growth, namely dB_{Ct} in Equation (45). Using (34) and (31), one can show that the premium for short-run risks is given by

$$\gamma\left[\sigma_K^2+\frac{1-\psi}{1-\gamma}\frac{H'\left(m_t;\sigma_e\right)}{H\left(m_t;\sigma_e\right)}\left(\rho\sigma_K\sigma_\theta+Q\right)\right]\approx\gamma\left[\sigma_K^2+\frac{1-\psi}{a+\kappa_1}\left(\rho\sigma_K\sigma_\theta+Q\right)\right], \tag{48}$$

where the approximate equality makes use of the log-linearization result in (28).

The term $\psi\frac{\gamma-\frac{1}{\psi}}{1-\gamma}\frac{H'(m_t;\sigma_e)}{H(m_t;\sigma_e)}\rho_{m,P}\sigma_m\sigma_{Pt}$ in Equation (47) is special to the Kreps and Porteus (1978) utility. It would vanish in the case of expected utility, i.e., $\gamma=\frac{1}{\psi}$. Note this term depends on the covariance between innovations in the persistent component of consumption growth (dm_t) and innovations in return ($\frac{dR_t}{R_t}$). I define this part of the risk premium as *premium for long-run risks*. It is the compensation for the asset's exposure to the risks associated with fluctuations in the long-run component of consumption growth, namely dm_t. Using (31) and (20), the premium for long-run risks is given by:

$$\psi\frac{\gamma-\frac{1}{\psi}}{1-\gamma}\frac{H'\left(m_t;\sigma_e\right)}{H\left(m_t;\sigma_e\right)}\left(\rho\sigma_K\sigma_\theta+Q\right)\approx\psi\frac{\gamma-\frac{1}{\psi}}{a+\kappa_1}\left(\rho\sigma_K\sigma_\theta+Q\right). \tag{49}$$

Now it is clear why the Lucas (1978)-Breeden (1979) formula in (26) obtains in the case $\sigma_e=\rho=0$ even in an economy with Kreps and Porteus (1978) utility. $\sigma_e=\rho=0$ implies that the innovations in the persistent component of consumption growth (dm_t) and the innovations in return ($\frac{dR_t}{R_t}$) are uncorrelated, therefore the premium for long-run risks vanishes even though $\gamma\neq\frac{1}{\psi}$.

Columns 5 and 6 of Table 3 show the decomposition of the equity premium in the benchmark model for various choices of the σ_e parameter. If $\sigma_e=0$, or equivalently, θ is observable, the risk premium of aggregate wealth comes completely from compensation for short-run risks. As discussed above, the Lucas (1978)-Breeden (1979) formula obtains despite $\gamma\neq\frac{1}{\psi}$. Note also, risk premium in this case, although quite high, is associated with a counterfactually high level of volatility of consumption growth.

Once the external source of information (e_t) is contaminated with noise, i.e., $\sigma_e > 0$, the agent can no longer distinguish completely the short-run and the long-run fluctuations in the return to aggregate wealth. Optimal learning implies that the covariance between innovations in the expected and the realized return to aggregate wealth (that is, the term $\rho\sigma_K\sigma_\theta + Q$ in Equations (48) and (49)) increases with σ_e. Equation (48) implies that the premium for short-run risks is decreasing in σ_e if $\psi > 1$, consistent with the results obtained in Section II.B.2: volatility of consumption growth is decreasing in σ_e if $\psi > 1$. Equation (49) implies that the premium for long-run risks is increasing in σ_e if $\gamma > 1$. One can show that the long-run risks channel always dominates by combining Equations (48) and (49). This confirms the result obtained in Section II.B.1: equity premium is increasing in σ_e if $\gamma > 1$. The calibration results are shown in Table 3. As σ_e increases from 0 to ∞, the premium for short-run risks decreases from 2.53% to 1.49%, while the premium for long-run risks increases from 0 to 2.01%. The fact that the volatility of aggregate consumption growth is small in the data implies that the premium for short-run risks generated by the model is necessarily small unless I use an implausibly high γ. As shown in Table 3, in the benchmark model with $\sigma_e = \infty$, the bulk of the total risk premium is compensation for long-run risks.

III Welfare Gain of Information

A Decomposition of Welfare Gain

In this section, I discuss the welfare gain of information about the long-run risk state variable $\{\theta_t\}_{t\geq 0}$ in the production economy considered above. In the case of expected utility, the agent is averse to fluctuations in consumption, but not to fluctuations in continuation utilities. In the case of recursive utility that features preference for early resolution of uncertainty, the agent is also averse to fluctuations in continuation utilities.[9] From the asset pricing perspective, this means any return correlated with fluctuations in continuation utilities must pay a premium, which is the premium for long-run risks discussed in the last section. From the welfare perspective, this means better information will lead to a welfare gain even if it does *not* change the allocation. I show that welfare gain from this channel could be quantitatively important.

I first provide a definition of the decomposition of welfare gain of information. I show that the welfare gain of information can come from two channels, welfare gain from *early resolution of uncertainty;* and welfare gain from *resource allocation.* I continue to consider the SDU specified in (1) with normalized aggregator $\mathcal{A}(V) = 0$. I will use M_0 to denote the certainty equivalent functional[10] with respect to the trivial σ field. Intuitively, the agent's date-0 utility V_0 may depend on the information available at date 0, and is therefore a random variable. $M_0(V_0)$ is the *unconditional* certainty equivalent of the stochastic utility V_0. Since I am working with the normalized aggregator, i.e. $\mathcal{A} = 0$, it can be shown that the certainty equivalent functional coincides with the expectation functional (Duffie and Epstein (1992b)), that is,

$$M_0(V_0) = E[V_0]. \tag{50}$$

For simplicity, I continue to assume learning steady state at date 0. Therefore the conditional distribution of θ_0 given m_0 is $N(m_0, Q)$, where Q is given in (14). I also assume at date 0, before any information is revealed, the prior distribution of m_0 is its steady-state distribution implied by (11), i.e. $N\left(\bar{\theta}, \frac{\sigma_m^2}{2a}\right)$.

At any date t, the history of the partially observable economy can be summarized

[9] Another way of interpreting preference for early resolution of uncertainty is aversion to persistence. This is emphasized in (Piazzesi and Schneider (2007)).

[10] To be accurate, the certainty equivalent functional associated with the aggregator $(f, \mathcal{A})$. The normalization $\mathcal{A} = 0$ implies that M_0 is the expectation functional.

by the triple of state variables (θ_t, m_t, K_t). Let $C(m, K)$[11] be the optimal consumption policy of the agent, and $V(m, K)$ be the value function as given in Proposition 1. Then $\forall t \geq 0$, the agent's date-t consumption C_t and date-t utility V_t satisfies:

$$C_t = C(m, K)\,; \quad V_t = V(m_t, K_t) = \frac{1}{1-\gamma} H(m_t) K_t^{1-\gamma}.$$

Now consider increasing the consumption of the agent permanently by a fraction of $\lambda \in (-1, \infty)$.[12] That is, $\forall t \geq 0$, the date-t consumption level is changed to $(1+\lambda) C(m_t, K_t)$. Denote the associated utility process by $\left\{V_t^\lambda\right\}_{t\geq 0}$, then V_t^λ is a time-invariant function of the state variables, which I denote as $V^\lambda(m_t, K_t)$. By homogeneity,

$$V_t^\lambda = V^\lambda(m_t, K_t) = (1+\lambda)^{1-\gamma} V(m_t, K_t).$$

That is, $V_t^\lambda = V^\lambda(m_t, K_t)$ is the date-t utility of a hypothetical agent in the partially observable economy whose consumption is increased permanently by a fraction of λ.

Similarly, in the completely observable economy, the consumption process $\bar{C}_t$ and utility process $\bar{V}_t$ are time-invariant functions of the state variables:

$$C_t = \bar{C}(\theta_t, K_t)\,; \quad \bar{V}_t = \bar{V}(\theta_t, K_t) = \frac{1}{1-\gamma} H(\theta_t) K_t^{1-\gamma}.$$

where the value function $\bar{V}(\theta, K)$ and the consumption policy function $\bar{C}(\theta, K)$ are defined in Equation (24). Presumably, knowing more information improves the agent's welfare:

$$M_0\left[\bar{V}(\theta_0, K_0)\right] \geq M_0\left[V(m_0, K_0)\right]. \tag{51}$$

The interpretation of the above expression is that, at date 0, before any information is

[11] For notational convenience, I suppress the dependence of the value functions and policy functions on σ_e.

[12] $\lambda > 1$ guarantees that the consumption in the economy is always strictly positive.

revealed, the agent will prefer a completely observable economy to a partially observable economy with the same initial condition. A sufficient condition for (51) is that the agent weakly prefers early resolution of uncertainty. Regardless of whether (51) holds or not, one can quantitatively measure the welfare gain of the information by the parameter λ that satisfies:

$$M_0\left[V^{\lambda}\left(m_0, K_0\right)\right] = M_0\left[\bar{V}\left(\theta_0, K_0\right)\right]. \tag{52}$$

Definition 1 *Welfare Gain of Information*

Welfare gain of information is measured by the parameter λ, which is the percentage of permanent consumption increase in a partially observable economy that is needed so that the representative agent is indifferent, at date 0, before any information is revealed, between the partially observable economy and a completely observable economy with the same initial condition.

In other words, the welfare gain of information is measured by the λ that equalizes the date-0 certainty equivalent in a partially observable economy with that in a completely observable economy as defined in Equation (52). By (50), Equation (52) is written as:

$$E\left[(1+\lambda)^{1-\gamma}\frac{1}{1-\gamma}H\left(m_0;\sigma_e\right)K_0^{1-\gamma}\right] = E\left[\frac{1}{1-\gamma}H\left(\theta_0\right)K_0^{1-\gamma}\right].$$

Therefore, λ can be solved as:

$$1+\lambda = \left\{\frac{E\left[H\left(\theta\right)\right]}{E\left[H\left(m;\sigma_e\right)\right]}\right\}^{\frac{1}{1-\gamma}}.$$

The welfare gain can be decomposed into *gain from early resolution of uncertainty* and *gain from resource allocation*. Before giving formal definitions to these concepts, I first introduce the notion of *completely observable economy with information-constrained allocation*. It is useful to consider the following thought experiment. Consider an agent in a partially observable economy who has already made a complete set of plan for her posterior belief and consumption for all future contingencies according to (6), (11)-(13) and (22). I will call the posterior belief and consumption of such an agent, as described above, the *information-constrained belief* and the *information-constrained consumption*, respectively. Suppose now at all $t \geq 0$, the true value of θ_t is revealed to the agent, however, the agent is not allowed to change her information-constrained consumption. (Presumably, the agent would like to reoptimize and choose a different consumption plan, because more information is available.) I define U_t as the date-t utility of such an agent whose information-constrained posterior belief about θ_t is m_t, who has capital stock K_t, and who will be informed of the true state of the economy at all future dates $s \geq t$. Then the recursion relation that defines the $\{U_t\}_{t\geq 0}$ process is:

$$\forall t \geq 0,\ U_t = E\left[\int_0^{\infty} f\left[C\left(m_s, K_s; \sigma_e\right), U_s\right] dt \left| K_t, \theta_t\right.\right]. \tag{53}$$

Definition 2 *Completely Observable Economy with Information-Constrained Allocation*

A completely observable economy with information-constrained allocation is an economy in which the representative agent's information filtration is $\mathcal{F}$, and in which the representative agent chooses the information-constrained consumption plan as determined by (6), (11)-(13) and (22).

In other words, the information set of the representative agent in a completely observable economy with information-constrained allocation is identical to that of an agent in a completely observable economy with the same initial condition; however, her consumption allocation is identical to that of an agent in a partially observable economy with the same initial condition. Her life time utility is given by (53). Appendix D shows that U_t is a time-invariant function of the state variables (m_t, θ_t, K_t) of the form:

$$U_t = U(\theta_t, m_t, K_t) = \frac{1}{1-\gamma} G(\theta_t, m_t) K_t^{1-\gamma}, \tag{54}$$

where $G(\theta, m)$ satisfy the PDE (89) in Appendix D. I also use U_t^{λ} to denote the date-t utility of such an agent whose information-constrained consumption is increased by a fraction of λ for all $t \geq 0$. $\forall t \geq 0$, U_t^{λ} is defined by the following recursion relation:

$$\forall t \geq 0,\ U_t^{\lambda} = E\left[\int_0^{\infty} f\left[(1+\lambda) C(m_s, K_s; \sigma_e), U_s^{\lambda}\right] dt \,|K_t, \theta_t\right].$$

Again, U_t^{λ} is a time-invariant function of the state variables. Homogeneity implies:

$$U_t^{\lambda} = U^{\lambda}(\theta_t, m_t, K_t) = (1+\lambda)^{1-\gamma} U(\theta_t, m_t, K_t).$$

The formal definitions of welfare gain from early resolution and welfare gain from resource allocation are stated below.

Definition 3 *Welfare Gain from Early Resolution of Uncertainty*

Welfare gain from early resolution of uncertainty is measured by the parameter λ_R, which is the percentage of permanent consumption increase in a partially observable economy that is needed so that the representative agent is indifferent, at date 0, before any information is revealed, between the partially observable economy and an economy with the same initial condition and the same information-constrained allocation, but with completely observable state. That is, it is the unique λ_R that satisfies

$$M_0\left[V^{\lambda_R}(m_0, K_0)\right] = M_0\left[U(\theta_0, m_0, K_0)\right]. \tag{55}$$

Definition 4 *Welfare Gain from Resource Allocation*

Welfare gain from resource allocation is measured by the parameter λ_A, which is the *percentage of permanent consumption increase in a completely observable economy with information-constrained allocation that is needed so that the representative agent is indifferent, at date* 0*, before any information is revealed, between the completely observable economy with information-constrained allocation and a completely observable economy with the same initial condition, but in which the agent chooses consumption allocation optimally given all available information. That is, it is the unique* λ_A *that satisfies*

$$M_0\left[U^{\lambda_A}\left(\theta_0, m_0, K_0\right)\right] = M_0\left[\bar{V}\left(\theta_0, K_0\right)\right]. \tag{56}$$

To understand the above definitions, let's consider the following two-stage experiment in a partially observable economy with the consumption process $\{C(m_t, K_t)\}_{t\geq 0}$, and the utility process $\{V(m_t, K_t)\}_{t\geq 0}$. In the first stage, we reveal the true state of the economy, θ_t to the agent at all $t \geq 0$. However, the agent is not allowed to update her consumption policy according to the new information. In particular, she will continue to use the information-constrained consumption policy at all times in the future. The agent's utility process $\{U_t\}_{t\geq 0}$ is given by (53) above. Suppose

$$M_0\left[U\left(\theta_0, m_0, K_0\right)\right] \geq M_0\left[V\left(m_0, K_0\right)\right], \tag{57}$$

that is, at date 0, before any information is revealed, the agent will prefer a completely observable economy with information-constrained allocation, to an otherwise identical partially observable economy. Note that the allocation associated with $U(\theta_0, m_0, K_0)$ and $V(m_0, K_0)$ are identical, so the welfare gain from $V(m_0, K_0)$ to $U(\theta_0, m_0, K_0)$ is completely due to the better information about future consumption. I term this component of the welfare gain *the welfare gain from early resolution of uncertainty.* For an expected utility maximizer, there is no welfare gain associated with more information as long as the information does not change the allocation. In fact, one can show for a recursive utility maximizer, (57) is true if and only if the agent prefers early resolution of uncertainty in the sense of Kreps and Porteus (1978). So the terminology is indeed appropriate. Regardless (57) holds or not, one can always increase the consumption in the partially observable economy by a fraction of $\lambda \in (-1, \infty)$, so that the agent is indifferent between the partially observable economy and the completely observable economy with information-constrained allocation. This is my measure of welfare gain from early resolution of uncertainty, λ_R. By definition, λ_R equalizes the data-0 certainty equivalent in a partially observable economy with that in an otherwise identical completely observable economy with information-constrained allocation, as defined in (55).

From Kreps and Porteus (1978), Epstein and Zin (1989),and Weil (1989), $\gamma > 1/\psi (\gamma < 1/\psi)$ corresponds to the case of preference for early (late) resolution of uncertainty. Therefore, under the parameter values in the calibration in Section II.C, the welfare gain from early resolution of uncertainty is positive.

In the second stage of the experiment, we let the agent in the completely observable economy with information-constrained allocation re-optimize the consumption-saving policy given all available information. The *welfare gain from resource allocation* is the welfare gain from re-optimizing the consumption-saving policy in the presence of better information. The welfare gain from resource allocation must be positive, since the agent can at least choose the information-constrained consumption allocation. Now consider increasing the information-constrained consumption permanently by a fraction λ_A. Since utility is strictly increasing in consumption, I can always choose λ_A so that the agent is just indifferent between the completely observable economy with information-constrained allocation, and the complete observable economy with optimal consumption allocation given all available information. In other words, λ_A equalizes the data-0 certainty equivalent in a completely observable economy with information-constrained allocation with that in an otherwise identical completely observable economy, but, in which the agent chose the optimal consumption allocation given all available information. This is the interpretation of Equation (56). By definition,

$$1+\lambda=(1+\lambda_R)(1+\lambda_A).$$

This is the sense in which the welfare gain from early resolution of uncertainty and the welfare gain from resource allocation sum up to the total welfare gain.

In the context of the production economy I consider in this paper, the welfare gain from resource allocation arises because of the presence of a non-trivial intertemporal saving technology. In pure exchange economies, of course, welfare gain from allocation is always 0. The following proposition provides explicit formulae for the decomposition of welfare gain:

Proposition 3 ***Welfare Gain of Information.***

The decomposition of welfare gain of information is given by:

$$1+\lambda=\left\{\frac{E\left[H\left(\theta\right)\right]}{E\left[H\left(m;\sigma_e\right)\right]}\right\}^{\frac{1}{1-\gamma}},$$

$$1+\lambda_R=\left\{\frac{E\left[G\left(\theta,m;\sigma_e\right)\right]}{E\left[H\left(m;\sigma_e\right)\right]}\right\}^{\frac{1}{1-\gamma}},$$

and

$$1+\lambda_A=\left\{\frac{E\left[H\left(\theta\right)\right]}{E\left[G\left(\theta,m;\sigma_e\right)\right]}\right\}^{\frac{1}{1-\gamma}}.$$

Proof. See Appendix D. ■

In the special case $\psi=1$, λ, λ_R, and λ_A can all be solved in closed form. Using arguments similar to those in Appendix B, one can show in this case,

$$H\left(\theta\right)=G\left(\theta,m\right)=e^{A+B\theta};\quad H\left(m;\sigma_e\right)=e^{A+Bm+B^2Q}$$

where $\not{A}$ and B are given in Equations (73) and (74) in Appendix B. Assume the ex ante belief about θ and m is the steady-state distribution of (θ, m), using Proposition 3, it is straightforward to show:

$$1 + \lambda = 1 + \lambda_R = \exp\left\{\frac{1}{2}\frac{\gamma - 1}{(a+\beta)^2}Q\right\}; \quad \lambda_A = 0 \tag{58}$$

The intuition for $\lambda_A = 0$ is that in the case $\psi = 1$, consumption-to-wealth ratio is constant, and, therefore, the optimal consumption policy does not depend on information about the state variable θ. Consequently, welfare gain from resource allocation is zero. In general, if $\psi \neq 1$ the optimal consumption policy will depend on the information and $\lambda_A > 0$.

B Quantitative Welfare Gain of Information

Having defined the decomposition of welfare gain of information, I can evaluate the quantitative importance of these channels of welfare gain in the calibrated economy in Section II.C. Table 4 summarizes the numerical results for the decomposition of welfare gain of information in the benchmark partially observable economy ($\sigma_e = \infty$) under various values of IES. All other parameters are the same as in the calibration in Section II.C.

(Insert Table 4 about here.)

First, there is a sizable welfare gain of information under all choices of the IES parameter. As shown in Table 4, the welfare gains of better information in the benchmark model with $\psi = 2.0$ is equivalent to 2.81% of permanent increase in annual consumption. The total welfare gain is above 2.69% for all choices of the IES parameter. Note even in the case $\sigma_e = \infty$, it is not true that the agent has no information about the long-run risks. In fact, the agent can still update her posterior beliefs about the long-run risks through observations of the return of the technology. Taking this into consideration, the welfare gain of more than 2.69% of permanent increase in annual consumption seems quite high.

Second, most of the welfare gain comes from the early resolution of uncertainty channel. From Table 4, the welfare gain from resource allocation is smaller at least by an order of magnitude than the welfare gain from early resolution of uncertainty in all cases. As discussed earlier, $\lambda_A = 0$ if $\psi = 1$. Fro $\psi = 1.5$, the optimal consumption-to-wealth ratio is not very sensitive to θ, therefore, welfare gain from resource allocation is very small. In fact λ_A is so close to 0 that it cannot be accurately computed numerically. Both components of the welfare gain increases with the IES parameter ψ. λ_R is increasing in IES because a higher IES is associated with a stronger preference for early resolution of uncertainty. λ_A is increasing in IES in the range $\psi \in [1, \infty)$ because a higher IES implies that the optimal consumption-to-wealth ratio is more sensitive to θ.[13] However, the welfare gain from resource allocation remains below 0.3% of permanent increase in consumption for all values of ψ. To summarize, the welfare gain of better information can be important, but most of the gain comes from preference for early resolution of uncertainty, not from improving the allocation of resources of the economy.

The high welfare gain of information from the early resolution of uncertainty channel found in this paper is consistent with the findings in Croce (2006) that the welfare gain of completely eliminating long-run risk is very high. Both of our results are outcomes of the agent's strong preference for early resolution of uncertainty under the calibrated preference parameters.

The asset pricing and welfare implications of information about long-run risks are both manifestations of the agent's strong preference for early resolution of uncertainty. Therefore experimental evidence on individuals' preference for early resolution of uncertainty should be viewed as evidence on the plausibility of the long-run risk asset pricing model. The intuition for the tight link between the asset pricing and the welfare results in this paper can be explained as follows. If $\gamma = \frac{1}{\psi}$, the agent is only averse to fluctuations in

[13] The same arguement implies that λ_A is decreasing in ψ in the range $\psi \in (0, 1]$. Although not shown here, numerical results indicate this is indeed the case.

consumption, but indifferent to fluctuations of continuation utility as long as the expected continuation utility is kept constant. In this case only short-run fluctuations of consumption enter the pricing kernel and affect the agent's welfare. Learning does not generate a premium for long-run risks, nor does it imply any welfare gain from early resolution of uncertainty. If $\gamma > \frac{1}{\psi}$, the agent is not only averse to fluctuations in consumption but also to fluctuations of continuation utility. Mathematically, $\gamma > \frac{1}{\psi}$ implies that the intertemporal utility aggregation functional is concave in continuation utility. A higher distance between γ and $\frac{1}{\psi}$ implies a higher curvature of the intertemporal utility aggregation functional with respect to the continuation utility, thus a higher degree of preference for early resolution of uncertainty. From the asset pricing perspective, this increases the volatility of the stochastic discount factor and, therefore, the equity premium for long-run risks. From the welfare perspective, this implies a higher welfare gain from early resolution of uncertainty. The decomposition of welfare gain of information in Proposition 3 and the quantitative results in Table 4 provide an intuitive measure of preference for early resolution of uncertainty and may be used in designing experiments on individuals' preference over timing of resolution of uncertainty, which will provide important insights on the plausibility of the long-run risk asset pricing model.

IV Conclusion

I study the asset pricing and welfare implications of information quality about the long-run growth rate of the economy in a simple production economy. I show that the relationship between information quality and equity premium depends on the RRA parameter, and the relationship between information quality and the volatility of consumption growth depends on the IES parameter. I show that under plausible parameter values, learning generates higher equity premium, lower volatility of consumption growth, and lower volatility of risk-free interest rate. All of the above channels are quantitatively significant and lead to substantial improvement of the Bansal and Yaron (2004) model in terms its predictions for the first two moments of the return to aggregate wealth. My welfare analysis implies that if one is to take the evidence from asset market data seriously, then the welfare gain of better information that comes from the early resolution of uncertainty channel can be more than 2.69% of permanent increase in annual consumption. This implies that a better understanding of the long-run risks of the economy has important welfare consequences even if it might not result in significant improvements in resource allocation.

Three elements in the model are essential for the asset pricing and welfare results: Kreps and Porteus (1978) utility, learning, and production. As explained in Sections II and III, the Kreps and Porteus (1978) utility is important for the asset pricing results because it achieves a separation between RRA and IES, and is important for the welfare analysis because it gives rise to a nontrivial preference for the timing of resolution of uncertainty.

Learning is important both from the asset pricing and the welfare perspective. It is well known that an economy with learning is observationally equivalent to another economy without learning, with appropriately redefined state variables.[14] The asset pricing results of this paper could be interpreted, therefore, as follows: the conditional distribution *endogenously* generated by learning improves the performance of the Bansal and Yaron (2004) model without learning and is quantitatively significant. The same asset pricing results can be obtained by assuming *exogenously* the conditional distribution of the state variables. The welfare analysis shows that the fact that the observed consumption data and equity price data come from an economy with learning has important welfare consequences: better information could lead to significant welfare gains, even if it does not affect the allocation of the economy.

[14] The precise statement of this claim can be found for example, in Veronesi (1999).

The assumption of production economy is not essential for delivering the asset pricing statistics, but is important for the two thought experiments conducted in the paper.

Although the Bansal and Yaron (2004) model is a pure exchange economy, and mine is a production economy, the similarity between the consumption process (Table 2) generated by the two models indicates that one could obtain the same asset pricing results by assuming, *exogenously*, a different covariance structure of the state variables in the Bansal and Yaron (2004) model.

However, one needs a production economy model to answer the question: "How would the asset prices be changed if information quality gets better?," and the question: "How would information quality improves economic welfare?" The assumption of production economy recognizes the fact that as information quality changes, the agent will change the consumption-saving policy given all available information. Changes in the allocation of the economy will affect, in general, both asset prices and economic welfare, the quantitative importance of which needs to be gauged in a well-specified production economy.

The assumption of linear production technology is employed for simplicity. The linear production technology probably does not capture all the important aspects of the real world that are needed to answer the quantitative asset pricing questions we posed. One needs a more sophisticated, real business cycle to check the robustness of the asset pricing results. However, it is reasonable to believe that the basic welfare result obtained in this paper will hold in much more general settings. My result implies that the welfare gain from early resolution of uncertainty is quantitatively important, while the welfare gain from resource allocation is negligible. In any production economy that endogenizes long-run risks, the fact that long-risks carry a large premium implies a non-trivial welfare gain associated with eliminating or mitigating the risk. As noted in Section III of the paper, this is the consequence of the strong preference of early resolution of uncertainty under the parameterization of the preference. The welfare gain from allocation is likely to be small. Information about long-run risks changes allocation by affecting the consumption-to-wealth ratio. The calibration in Section III implies that as long as the IES parameter is not too large, the optimal adjustment of the consumption-to-wealth ratio does not generate significant welfare even in a model without adjustment cost. Production-based asset pricing models typically assume adjustment cost or real rigidity, both of which are likely to make the welfare gain from allocation even smaller.

Appendices

A Proof of Proposition 1

The HJB equation for the Pareto optimality problem in (18) is written as:

$$\max_{C} \{f(C, V(m,K)) + \mathcal{L}V(m,K)\} = 0 \tag{59}$$

where $\mathcal{L}$ is the differential operator associated with the process $\{K_t\}$ and $\{m_t\}$, i.e.,

$$\begin{aligned} \mathcal{L}V(m,K) &= (Km - c) V_K + \frac{1}{2}K^2\sigma_K^2 V_{KK} \\ &\quad + a\left(\overline{\theta} - m\right) V_m + \frac{1}{2}\sigma_m^2 V_{mm} + (\rho\sigma_K\sigma_\theta + Q) K V_{mK} \end{aligned}$$

By homogeneity, V must be of the form in (21). Monotonicity of value function with respect to m implies $H(m;\sigma_e)$ is strictly increasing (decreasing) in m if $\gamma < 1 (\gamma > 1)$. Using (21) to simplify the HJB equation, and denoting $x = \frac{C}{K}$, one has:

$$\begin{aligned} &\max_{x} \left\{ \frac{\beta}{1-1/\psi} x^{1-\frac{1}{\psi}} H(m;\sigma_e)^{1-\frac{1-1/\psi}{1-\gamma}} + \left(m - x - \frac{1}{2}\gamma\sigma_K^2\right) H(m;\sigma_e) \right\} \\ &- \frac{\beta}{1-1/\psi} H(m;\sigma_e) + \left[\frac{1}{1-\gamma} a\left(\overline{\theta} - m\right) + (\rho\sigma_K\sigma_\theta + Q)\right] H'(m;\sigma_e) \\ &+ \frac{1}{2}\frac{1}{1-\gamma}\sigma_m^2 H''(m;\sigma_e) = 0 \end{aligned} \tag{60}$$

The first order condition in the HJB equation in (60) implies the optimal consumption policy function is of the form in (22), and $x(m;\sigma_e)$ is the consumption-to-wealth ratio given by (23). Using the optimal consumption policy, the HJB in (60) can be simplified to an ODE of $H(m;\sigma_e)$.

$$\frac{1}{\psi-1}\beta^{\psi}H(m;\sigma_e)^{\frac{1-\psi}{1-\gamma}}+\left(m-\frac{\beta}{1-1/\psi}-\frac{1}{2}\gamma\sigma_K^2\right) \tag{61}$$
$$+\left[\frac{1}{1-\gamma}a\left(\overline{\theta}-m\right)+\left(\rho\sigma_K\sigma_\theta+Q\right)\right]\frac{H'(m;\sigma_e)}{H(m;\sigma_e)}+\frac{1}{2}\frac{1}{1-\gamma}\sigma_m^2\frac{H''(m;\sigma_e)}{H(m;\sigma_e)}=0$$

Finally, monotonicity of $H(m;\sigma_e)$ implies $H(m;\sigma_e)^{\frac{1}{1-\gamma}}$ is always a strictly increasing function of m. This implies the consumption-to-wealth ratio function in (23) is strictly increasing (decreasing) in m if $\psi<1$ ($\psi>1$).

B Proof of Proposition 2

As shown by Duffie and Epstein (1992a), in an economy with recursive preference, the state price process, denoted $\{\pi_t\}_{t\geq 0}$ satisfies:

$$\frac{d\pi_t}{\pi_t}=df_C(C_t,V_t)+f_V(C_t,V_t)\,dt \tag{62}$$

Using results in Proposition 1, and applying Ito's formula, one can show π_t is a diffusion process of the form:

$$d\pi_t=\pi_t\left[-r_t dt-\gamma\sigma_K d\tilde{B}_{Kt}+\frac{H'(m_t;\sigma_e)}{H(m_t;\sigma_e)}\sigma_m d\widetilde{B}_{mt}\right] \tag{63}$$

where r_t is the risk-free interest rate.

Absence of arbitrage implies that the risk premium of the cumulative return process $\{R_t\}_{t\geq 0}$ defined in (9) must satisfy:

$$\mu_{Rt}-r_t=-cov_t\left(\frac{d\pi_t}{\pi_t},\frac{dR_t}{R_t}\right) \tag{64}$$

Using (31) and (63), one can obtain (25). The completely observable economy corresponds to the case $\sigma_e = 0$. In this case $Q = 0$ and the risk premium is given by (33).

I now derive a closed form expression for Equation (25) using the log-linearization method proposed by Campbell, Chacko, Rodriguez, and Viceira (2004). Note the only nonlinear term in (61) is $\beta^{\psi} H(m)^{\frac{1-\psi}{1-\gamma}}$, which is exactly the consumption-to-wealth ratio function $x(m; \sigma_e)$. Using log linear approximation of $x(m; \sigma_e)$, I have

$$x(m) \approx \kappa_0 + \kappa_1 \log x(m; \sigma_e) = \kappa_0 + \kappa_1 \psi \ln \beta + \kappa_1 \frac{1-\psi}{1-\gamma} \ln H(m; \sigma_e) \tag{65}$$

where

$$\kappa_0 = \kappa_1 (1 - \log \kappa_1) \tag{66}$$

and κ_1 can be chosen as the consumption-to-wealth ratio when m is equal to its unconditional mean $\bar{\theta}$

$$\kappa_1 = x\left(\bar{\theta}; \sigma_e\right)$$

Using (65), one can approximate the ODE in (61) as:

$$\begin{aligned} &\frac{1}{\psi - 1}\left[\kappa_0 + \kappa_1 \psi \log \beta + \kappa_1 \frac{1-\psi}{1-\gamma} \log H(m; \sigma_e)\right] + \left(m - \frac{\beta}{1 - 1/\psi} - \frac{1}{2}\gamma \sigma_K^2\right) \\ &+ \left[\frac{1}{1-\gamma} a\left(\bar{\theta} - m\right) + (\rho \sigma_K \sigma_\theta + Q)\right] \frac{H'(m; \sigma_e)}{H(m; \sigma_e)} + \frac{1}{2} \frac{1}{1-\gamma} \sigma_m^2 \frac{H''(m; \sigma_e)}{H(m; \sigma_e)} = 0 \end{aligned} \tag{67}$$

I guess $H(m; \sigma_e)$ is of the form

$$H(m; \sigma_e) = \exp(A + Bm) \tag{68}$$

Using method of undetermined coefficients, I have

$$B = \frac{1-\gamma}{a + \kappa_1} \tag{69}$$

and

$$A=\frac{1}{\kappa_1}\left\{\frac{1}{2}\sigma_m^2B^2+\left[(1-\gamma)\left(\rho\sigma_K\sigma_\theta+Q\right)+a\overline{\theta}\right]B+D\right\} \tag{70}$$

where

$$D=\frac{1-\gamma}{1-\psi}\left[-\kappa_0-\psi\kappa_1\log\beta+\psi\beta-\frac{1}{2}\gamma\left(1-\psi\right)\sigma_K^2\right]$$

Using (68),

$$\frac{H'\left(m;\sigma_e\right)}{H\left(m;\sigma_e\right)}=B=\frac{1-\gamma}{a+\kappa_1} \tag{71}$$

Therefore the risk premium can be approximated by

$$\mu_{Rt}-r_t\approx\gamma\sigma_K^2+\frac{\gamma-1}{a+\kappa_1}\left(\rho\sigma_K\sigma_\theta+Q\right) \tag{72}$$

This above approximation is exact if $\psi=1$, in which case

$$B=\frac{1-\gamma}{a+\beta} \tag{73}$$

$$A=\frac{1}{\kappa_1}\left\{\frac{1}{2}\sigma_m^2B^2+\left[(1-\gamma)\left(\rho\sigma_K\sigma_\theta+Q\right)+a\overline{\theta}\right]B+(1-\gamma)\left(\beta\log\beta-\beta-\frac{1}{2}\gamma\sigma_K^2\right)\right\} \tag{74}$$

C Equity Premium in the Pure Exchange Economy

Consider a pure exchange economy where the representative agent has the utility function given in (1), and (2), and the consumption process given in (38). Denote the representative agent's utility process in the pure exchange economy as $\{V_{E,t}\}$, then $V_{E,t}=V_E\left(m_t,C_t\right)$ must satisfy the following HJB equation (Duffie and Epstein (1992b)):

$$f\left(C,V_E\left(m,C\right)\right)+\mathcal{L}V_E(m,C)=0 \tag{75}$$

Homogeneity of the utility function implies

$$V(m,C) = \frac{1}{1-\gamma} H_E(m;\sigma_e) C^{1-\gamma} \tag{76}$$

for some smooth function H_E. Combine (75) and (76), $H_E(m;\sigma_e)$ must satisfy the following ODE:

$$\begin{aligned} &\frac{\beta}{1-1/\psi} H_E(m;\sigma_e)^{\frac{1-1/\psi}{1-\gamma}} + \left(m - \frac{\beta}{1-1/\psi} - \frac{1}{2}\gamma\sigma_C^2 \right) \\ &+ \left[\frac{1}{1-\gamma} a\left(\overline{\theta} - m\right) + \left(\rho\sigma_C\sigma_\theta + Q_E\right) \right] \frac{H_E'(m;\sigma_e)}{H_E(m;\sigma_e)} + \frac{1}{2}\frac{1}{1-\gamma}\sigma_m^2 \frac{H_E''(m;\sigma_e)}{H_E(m;\sigma_e)} = 0 \end{aligned} \tag{77}$$

where Q_E is the steady-state posterior variance of θ_t in the exchange economy. Using a log-linear approximation argument similar to that in Appendix B,

$$\frac{H_E'(m;\sigma_e)}{H_E(m;\sigma_e)} \approx \frac{1-\gamma}{a+\varpi_1} \tag{78}$$

where ϖ_1 is the steady-state consumption-to-wealth ratio in the exchange economy.

Using (62), the state price process of this economy is given by:

$$\frac{d\pi_t}{\pi_t} = -r_{Et}dt - \gamma\sigma_C dB_{C,t} + \frac{\frac{1}{\psi} - \gamma}{1-\gamma} \frac{H_E'(m_t;\sigma_e)}{H_E(m_t;\sigma_e)} \sigma_{E,m} d\tilde{B}_{E,mt} \tag{79}$$

where r_{Et} is the risk-free interest rate in the exchange economy, and

$$\sigma_{E,m} = \sqrt{\left(\frac{1}{\sigma_C}Q_E + \rho\sigma_\theta\right)^2 + \left(\frac{1}{\sigma_e}Q_E\right)^2} = \sigma_\theta^2 - 2aQ_E \tag{80}$$

The Brownian motion $\tilde{B}_{E,mt}$ is defined by:

$$\tilde{B}_{E,mt} = \frac{1}{\sigma_{E,m}} \left[\left(\frac{1}{\sigma_C}Q_E + \rho\sigma_\theta\right) \widetilde{B}_{Ct} + \frac{1}{\sigma_e} Q_E \widetilde{B}_{et} \right]$$

The aggregate wealth measured in terms of current period consumption is the present value of future consumption stream and is given by:

$$W_t = \frac{1}{\beta} H_E\left(m_t;\sigma_e\right)^{\frac{1-1/\psi}{1-\gamma}} C_t \tag{81}$$

Using Ito's formula, I have

$$\frac{dW_t}{W_t} = \mu_{Wt}dt + \sigma_C d\widetilde{B}_{Ct} + \frac{1-\frac{1}{\psi}}{1-\gamma}\frac{H_E'\left(m_t;\sigma_e\right)}{H_E\left(m_t;\sigma_e\right)}\sigma_{E,m}d\tilde{B}_{mt} \tag{82}$$

where μ_{Wt} is the instantaneous expected return of the aggregate wealth in the pure exchange economy. Therefore,

$$Var_t\left(\frac{dW_t}{W_t}\right) = \sigma_C^2 + \left(\frac{1-\frac{1}{\psi}}{1-\gamma}\right)^2\left(\frac{H_E'\left(m_t;\sigma_e\right)}{H_E\left(m_t;\sigma_e\right)}\right)^2\sigma_{E,m}^2 + 2\frac{1-\frac{1}{\psi}}{1-\gamma}\frac{H_E'\left(m_t;\sigma_e\right)}{H_E\left(m_t;\sigma_e\right)}\left(\rho\sigma_C\sigma_\theta + Q_E\right) \tag{83}$$

Using (79), (82) and the formula for equity premium (64), the equity premium on aggregate wealth is given by:

$$\begin{aligned}\mu_{Rt} - r_t &= \gamma\sigma_C^2 + \frac{\left(1-\frac{1}{\psi}\right)\left(\gamma-\frac{1}{\psi}\right)}{\left(1-\gamma\right)^2}\left(\frac{H_E'\left(m_t;\sigma_e\right)}{H_E\left(m_t;\sigma_e\right)}\right)^2\sigma_{E,m}^2 \\ &\quad + \left[\left(1+\gamma\right)\frac{1-\frac{1}{\psi}}{1-\gamma} - 1\right]\frac{H_E'\left(m_t;\sigma_e\right)}{H_E\left(m_t;\sigma_e\right)}\left(\rho\sigma_C\sigma_\theta + Q_E\right)\end{aligned} \tag{84}$$

The total equity premium in (84) can be decomposed into a myopic demand component and a hedging demand component as in (39). Denote the myopic demand component of the equity premium as MD_t, using (83) and (80),

$$MD_t = \gamma Var_t\left(\frac{dW_t}{W_t}\right)$$

$$= \gamma\left\{\sigma_C^2 + \left(\frac{1-\frac{1}{\psi}}{1-\gamma}\right)^2\left(\frac{H_E'\left(m_t;\sigma_e\right)}{H_E\left(m_t;\sigma_e\right)}\right)^2\left(\sigma_\theta^2 - 2aQ_E\right) + 2\frac{1-\frac{1}{\psi}}{1-\gamma}\frac{H_E'\left(m_t;\sigma_e\right)}{H_E\left(m_t;\sigma_e\right)}\left(\rho\sigma_C\sigma_\theta + Q_E\right)\right\}$$

The hedging demand adjustment term, denoted HD_t is therefore given by

$$HD_t = (\mu_{Rt} - r_t) - MD_t$$
$$= -\frac{1}{\psi}\left\{\frac{1-\frac{1}{\psi}}{1-\gamma}\left(\frac{H_E'(m_t;\sigma_e)}{H_E(m_t;\sigma_e)}\right)^2\left(\sigma_\theta^2 - 2aQ_E\right) + \frac{H_E'(m_t;\sigma_e)}{H_E(m_t;\sigma_e)}\left(\rho\sigma_C\sigma_\theta + Q_E\right)\right\}$$

Using (78) to simplify the above expressions, I have:

$$MD_t \approx \gamma\left\{\sigma_C^2 + \frac{\left(1-\frac{1}{\psi}\right)^2}{\left(a+\varpi_1\right)^2}\left(\sigma_\theta^2 - 2aQ_E\right) + 2\frac{1-\frac{1}{\psi}}{a+\varpi_1}\left(\rho\sigma_C\sigma_\theta + Q_E\right)\right\}$$

$$HD_t \approx -\frac{1}{\psi}\left\{\frac{\left(1-\frac{1}{\psi}\right)(1-\gamma)}{\left(a+\varpi_1\right)^2}\left(\sigma_\theta^2 - 2aQ_E\right) + \frac{1-\gamma}{a+\varpi_1}\left(\rho\sigma_C\sigma_\theta + Q_E\right)\right\}$$

Therefore,

$$\frac{\partial}{\partial\sigma_e}MD_t \approx 2\gamma\left(1-\frac{1}{\psi}\right)\frac{\left(\varpi_1 + a/\psi\right)}{\left(a+\varpi_1\right)^2}\frac{\partial Q}{\partial\sigma_e} \tag{85}$$

and

$$\frac{\partial}{\partial\sigma_e}HD_t \approx \frac{1}{\psi}\left(\gamma-1\right)\frac{\left[a\left(1/\psi - 1\right) + \kappa_1\right]}{\left(a+\varpi_1\right)^2}\frac{\partial Q}{\partial\sigma_e} \tag{86}$$

Note for ψ close to 1, the sign of (85) is determined by $1-\frac{1}{\psi}$, and the sign of (86) is determined by $\gamma - 1$. If I impose $\gamma = \frac{1}{\psi}$, then

$$\frac{\partial}{\partial\sigma_e}\left[\mu_{Rt} - r_t\right] \approx \left(1-\frac{1}{\psi}\right)\frac{\frac{1}{\psi}}{a+\varpi_1}\frac{\partial Q}{\partial\sigma_e} \tag{87}$$

Note (87) will have the same sign as (85), that is, the effect through the myopic demand component dominates.

D Proof of Proposition 3

Using the HJB equation for SDU, $U(\theta, m, K)$ must satisfy:

$$f\left(C\left(m, K; \sigma_e\right), U\left(\theta, m, K\right)\right) + \mathcal{L}U\left(\theta, m, K\right) = 0 \tag{88}$$

where $\mathcal{L}$ is the differential operator (with respect to the complete information filtration $\mathcal{F}$) of the associated diffusion processes. Note with respect to $\mathcal{F}$, $\{\theta_t, K_t\}$ are diffusion processes given in (5) and (6). Using (11) (13), and (**??**), $\{m_t\}_{t\geq 0}$ is a diffusion process (with respect to $\mathcal{F}$) that satisfies:

$$\begin{aligned} dm_t &= \left\{a\left(\bar{\theta} - m_t\right) + \left[Q\left(\sigma_K^{-2} + \sigma_e^{-2}\right) + \rho\frac{\sigma_\theta}{\sigma_K}\right]\left(\theta_t - m_t\right)\right\} dt \\ &\quad + \left(\frac{Q}{\sigma_K} + \rho\sigma_\theta\right) dB_{Kt} + \frac{Q}{\sigma_e} dB_{et} \end{aligned}$$

By homogeneity, I know $U(\theta, m, K)$ must be of the form in (56). Use the consumption policy function (22) and (23), Equation (88) can be simplified to the following PDE:

$$\begin{aligned} &\frac{1}{1-1/\psi}\beta^\psi H(m)^{\frac{1-\psi}{1-\gamma}}\left[\frac{G(\theta,m)}{H(m)}\right]^{-\frac{1-1/\psi}{1-\gamma}} \\ &-\beta^\psi H(m)^{\frac{1-\psi}{1-\gamma}} - \frac{\beta}{1-1/\psi} + \theta - \frac{1}{2}\gamma\sigma_K^2 \\ &+\frac{1}{1-\gamma}a\left(\bar{\theta}-\theta\right)\frac{G_\theta(\theta,m)}{G(\theta,m)} + \frac{1}{2}\frac{1}{1-\gamma}\sigma_\theta^2\frac{G_{\theta\theta}(\theta,m)}{G(\theta,m)} \\ &+\frac{1}{1-\gamma}\left[a\left(\bar{\theta}-m\right) + \left[Q\left(\sigma_K^{-2}+\sigma_e^{-2}\right)+\rho\frac{\sigma_\theta}{\sigma_K}\right](\theta - m) + (1-\gamma)\left(Q + \rho\sigma_\theta\sigma_K\right)\right]\frac{G_m(\theta,m)}{G(\theta,m)} \\ &\frac{1}{2}\frac{1}{1-\gamma}\sigma_m^2\frac{G_{mm}(\theta,m)}{G(\theta,m)} = 0 \end{aligned} \tag{89}$$

Because I am working with the normalized aggregator, the certainty equivalence functional is just the expectation functional. Using Equation (52), λ is given by:

$$E\left\{\frac{1}{1-\gamma}H\left(m;\sigma_e\right)\left[\left(1+\lambda\right)K\right]^{1-\gamma}\right\}=E\left\{\frac{1}{1-\gamma}H\left(\theta\right)K^{1-\gamma}\right\}$$

Simplify the above, I have:

$$1+\lambda=\left\{\frac{E\left[H\left(\theta\right)\right]}{E\left[H\left(m;\sigma_e\right)\right]}\right\}^{\frac{1}{1-\gamma}}$$

Similarly, using Equations (55) and (56), λ_R and λ_A are determined by:

$$1+\lambda_R=\left\{\frac{E\left[G\left(\theta,m;\sigma_e\right)\right]}{E\left[H\left(m;\sigma_e\right)\right]}\right\}^{\frac{1}{1-\gamma}}$$

and

$$1+\lambda_A=\left\{\frac{E\left[H\left(\theta\right)\right]}{E\left[G\left(\theta,m;\sigma_e\right)\right]}\right\}^{\frac{1}{1-\gamma}}$$

respectively.

References

Bansal, R. (2008). *Long-run Risks and Risk Compensation in Equity Markets*, Chapter 5, pp. 167–198. Elsevier.

Bansal, R., D. Kiku, and A. Yaron (2007, September). Risks for the long run: Estimation and inference.

Bansal, R. and A. Yaron (2004, August). Risk for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance 59*(4), 1481–1509.

Boldrin, M., L. Christiano, and J. Fisher (2001, March). Habit rersistence, asset returns, and the business cycle. *American Economic Review 91*(1), 149–166.

Breeden, D. T. (1979, September). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics 7*(3), 265–296.

Brennan, M. and Y. Xia (2001, April). Stock price volatility and equity premium. *Journal of Monetary Economics 47*(2), 249–283.

Brevik, F. and S. D'Addona (2007, March). Information quality and stock returns revisited.

Campbell, J. Y., G. Chacko, J. Rodriguez, and L. M. Viceira (2004, October). Strategic asset allocation in a continuous-time var model. *Journal of Economic Dynamics and Control 28*(11), 2195–2214.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies 1*(3), 195–228.

Croce, M. M. (2006, February). Welfare costs, long run consumption risk, and a production economy.

Croce, M. M., M. Lettau, and S. C. Ludvigson (2007, February). Investor information, long-run risk, and the duration of risk cash-flows.

Duffie, D. and L. G. Epstein (1992a). Asset pricing with stochastic differential utility. *Review of Financial Studies 5*(3), 411–436.

Duffie, D. and L. G. Epstein (1992b, March). Stochastic differential utility. *Econometrica 60*(2), 353–394.

Epstein, L. and S. E. Zin (1989, July). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica 57*(4), 937–969.

Gollier, C. and E. Schlee (2006). Information and equity premium.

Hansen, L. P. and T. J. Sargent (2006, May). Fragile beliefs and the price of model uncertainty.

Jermann, U. J. (1998, April). Asset pricing in production economies. *Journal of Monetary Economics 41*(2), 257–275.

Kreps, D. M. and E. L. Porteus (1978, January). Temporal resolution of uncertainty. and dynamic choice theory. *Econometrica 46*(1), 185–200.

Liptser, R. S. and A. N. Shiryaev (2001). *Statistics of Random Processes* (2 ed.). Stochastic Modelling and Applied Probability. Springer.

Lucas, R. E. (1978, November). Asset prices in an exchange economy. *Econometrica 46*(6), 1429–1445.

Lucas, R. E. (1987). *Models of Business Cycles*. Oxford: Basil Blackwell.

Lustig, H., S. V. Nieuwerburgh, and A. Verdelhan (2008, January). The wealth-consumption ratio: A litmus test for consumption-based asset pricing models.

Merton, R. C. (1992). *Continuous-Time Finance*. Oxford, U.K.: Basil Blackwell.

Piazzesi, M. and M. Schneider (2007). Equilibrium yield curves.

Rouwenhorst, K. G. (1995). Asset pricing implications of equilibrium business cycle models. In T. F. Cooley (Ed.), *Frontiers of Business Cycle Research*, Chapter 10, pp. 294–330. Princeton, NJ: Princeton University Press.

Tallarini, T. D. (2000, June). Risk sensitive real business cycles'. *Journal of Monetary Economics 45*(3), 507–532.

Veronesi, P. (1999, Winter). Stock market overreaction to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies 12*(5), 975–1007.

Veronesi, P. (2000, April). How does information quality affect stock returns? *Journal of Finance 55*(2), 807–837.

Weil, P. (1989, January). The equity premium puzzle and the riskfree rate puzzle. *Journal of Monetary Economics 24*(3), 401–422.

Table 1: **Information Quality and Equity Premium in Pure Exchange Economies**

	$\gamma < 1$	$\gamma > 1$
$\psi < 1$	$(-,-)$	$(-,+)$
$\psi > 1$	$(+,-)$	$(+,+)$

Notes: This table shows the effect of information quality on equity premium in pure exchange economies with Kreps-Porteus utility. The first sign is the effect of learning on equity premium through the consumption-to-wealth ratio channel, which depends on the IES parameter. The second sign is the effect of learning on equity premium through the hedging demand channel, which depends on the RRA parameter. In general the direction in which learning affects equity premium depends on the relative magnitude of the two effects. If RRA and IES are both greater than 1, the two effects reinforce each other, and learning increases equity premium. CRRA is a special case where the consumption-to-wealth ratio channel dominates, and consequently, learning increases equity premium if and only if $\psi > 1$(or, equivalently, $\gamma < 1$).

Table 2: **Asset Pricing Statistics of the Bench-mark Model**

	Data (%)	BY (%)	Model (%)
$E[g]$	1.8	1.8	1.80
$Std[g]$	2.93 (0.69)	2.72	3.15
$AC(1)$	0.49 (0.14)	0.48	0.45
$AC(2)$	0.15 (0.22)	0.29	0.22
$E[R_f]$	0.86	0.93	0.87
$Std(R_f)$	0.97	0.57	0.59
$E[R_W - R_f]$	3.52	1.60	3.50
$Std[R_W]$	6.00	3.28	5.69

Notes: This table compares the moments of annualized consumption growth, risk-free interest rate, and equity premium on aggregate wealth in the data, in Bansal and Yaron (2004)'s model, and in the benchmark partially observable economy. Rows 1-4 are the mean, the standard deviation and the first- and second-order autocorrelation of consumption growth. Data on these statistics are computed from BEA data 1929-1998 and are taken from Bansal and Yaron (2004) directly. Rows 5-6 are the mean and the standard deviation of the risk-free interest rate. The last two rows are the mean of risk-premium of aggregate wealth and the standard deviation of return to aggregate wealth. The first column shows the empirical estimates of these moments from data. Empirical estimates of moments of consumption growth and moments of risk-free interest rate are taken from Bansal and Yaron (2004), while the empirical estimates of risk premium and volatility of return to aggregate wealth are based on Lustig, Nieuwerburgh, and Verdelhan (2008). The second column displays these moments generated by Bansal and Yaron (2004)'s model. All moments are taken from Bansal and Yaron (2004), except for $E[R_W - R_f]$ and $Std[R_W]$, which are computed in Lustig, Nieuwerburgh, and Verdelhan (2008). The statistics in the third column are computed based on $1,000$ simulations of the benchmark model for 100 years. All simulations are done in the continuous-time model and aggregated to an annual level.

Table 3: **Asset Pricing Implications of Information Quality**

σ_e	g (%)		$E[\mu_{Rt} - r_t]$ (%)			r_t (%)	
	$E[g]$	$\sigma[g]$	Total	Short-Run	Long-Run	$E[r_t]$	$\sigma[r_t]$
0.00	1.92	5.9	2.53	2.53	0	1.85	1.40
0.01	1.89	5.89	2.86	2.16	0.70	1.52	1.39
0.02	1.84	5.07	3.07	1.92	1.15	1.31	1.17
0.05	1.80	3.82	3.32	1.63	1.69	1.06	0.81
0.10	1.79	3.33	3.41	1.54	1.87	0.97	0.65
∞	1.79	3.15	3.50	1.49	2.01	0.87	0.59

Notes: This table documents the effects of information quality on moments of consumption growth, equity premium of aggregate wealth, and risk-free interest rate for various values of σ_e. All quantities are computed at an annual level. Columns 2 and 3 are the means and the standard deviations of annual consumption growth. Columns 4-6 show the total risk premium, and the decomposition of risk premium into premium for long-run risks and premium for short-run risks. Columns 7-8 show the means and the standard deviations of the risk-free interest rate in the simulated economies for various values of the σ_e parameter. All statistics are computed based on $1,000$ simulations of the benchmark model for 100 years. All simulations are done in the continuous-time model and aggregated to an annual frequency.

Table 4: **Decomposition of Welfare Gain of Information**

ψ	$\lambda\,(\%)$	$\lambda_R\,(\%)$	$\lambda_A\,(\%)$
1.0	2.69	2.69	0
1.5	2.70	2.70	0
2.0	2.81	2.71	0.10
2.5	2.92	2.72	0.19
3.0	3.02	2.73	0.28

Notes: This table documents the decomposition of welfare gain of information in the benchmark partially observable model for various values of the IES parameter. The results are based on the calibrated parameter values used in Section II.C of the paper, except for the values of IES, which are listed in the first column. The welfare gain for case $\psi = 1$ is computed using the closed form solution in Equation (58). The welfare gain for all other cases are based on numerical solutions. For $\psi = 1.5$, the theoretical value of λ_A is strictly positive; however, the value of λ_A is too small to be numerically distinguished form 0.