Learning and Asset-Price Jumps

Ravi Bansal
Ivan Shaliastovich *

November 2010

*Bansal (email: ravi.bansal@duke.edu, tel: 919-660-7758, fax: 919-660-8038) is affiliated with the Fuqua School of Business, Duke University, Durham, NC 27708 and NBER, and Shaliastovich (email: ishal@wharton.upenn.edu, tel: 215-746-0005, fax: 215-573-9419) is at the Wharton School of Business, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104. We would like to thank Hengjie Ai, Tim Bollerslev, Mikhail Chernov, Riccardo Colacito, Janice Eberly, Steve Heston, Neil Shephard, George Tauchen and the participants of 2009 WFA meeting, 2008 AEA meeting and CCAPM Workshop in Aarhus, Denmark for their comments. The usual disclaimer applies.

Learning and Asset-Price Jumps

Abstract

We develop a general equilibrium model in which income and dividends are smooth, but asset prices are subject to large moves (jumps). A prominent feature of the model is that the optimal decision of investors to learn the unobserved state triggers large asset-price jumps. We show that the learning choice is critically determined by preference parameters and the conditional volatility of income process. An important prediction of the model is that the conditional volatility of income predicts future jump periods, while the level of income growth does not. We find that indeed in the data large moves in returns are predicted by consumption volatility, but not by the changes in consumption level. We show that the model can quantitatively capture these novel features of the data.

1 Introduction

A prominent feature of financial markets is infrequent but large price movements (jumps).[1] In this paper, we develop a model in which income and dividends have smooth Gaussian dynamics, however, asset prices are subject to large infrequent jumps. In our model, large moves in asset prices obtain from the actions of the representative agent to acquire more information about the unobserved state of the economy for a cost. We show that the optimal decision to incur a cost and learn the true economic state is directly related to the level of uncertainty in the economy. This implies that aggregate economic volatility, as well as market volatility, should predict jumps in returns. We show that indeed in the data, consistent with the model, return jumps are predicted by consumption volatility (market volatility). Further, the implied asset-price implications from our model are consistent with the key findings from parametric models about frequency and predictability of jumps as discussed in Singleton (2006) as well as nonparametric jump-detection analysis of Barndorff-Nielsen and Shephard (2006). Based on our evidence, we argue that our structural model provides an economic basis for realistic reduced-form models of stock price dynamics with time-varying volatility and jumps.

We rely on the long-run risks model of Bansal and Yaron (2004), which key ingredients are a small and persistent low-frequency expected growth component, time-varying income volatility, and recursive utility of Epstein and Zin (1989) and Weil (1989). The expected growth is unobserved and has to be estimated from the history of the data; in addition, the representative agent also has an option to incur a cost and learn the true economic state. This setup is designed to capture the intuition that some of the key aspects of the economy are not directly observable, but the agents can learn more about them through additional costly exploration. We show that the optimal decision to pay a cost and observe the true state endogenously depends on the aggregate volatility, the variance of the filtering error and agent's preferences. In particular, with preference for early resolution of uncertainty, the optimal frequency of learning about the true state after incurring a cost increases when the income volatility rises. On the other hand, with expected utility, the agent has no incentive to learn the true state even if costs are zero. Learning about the true state may lead to large revisions in expectations about future income, which translate into large moves in equilibrium asset prices. These large moves in asset prices obtain even though the underlying income and dividends in the economy are smooth and have no jumps. Such asset-price moves, we show, do not occur in economies where an option to learn about the true expected growth for a cost is absent.

[1] Jump-diffusion models are considered in Merton (1976), Naik and Lee (1990), Bates (1991), Bakshi, Cao, and Chen (1997), Pan (2002), Eraker, Johannes, and Polson (2003), Eraker (2004), Liu, Pan, and Wang (2005), Broadie, Chernov, and Johannes (2007). For a high-frequency analysis of intra-day data, refer also to Barndorff-Nielsen and Shephard (2006) and Andersen, Bollerslev, Diebold, and Vega (2003).

The learning mechanisms in our paper complement Van Nieuwerburgh and Veldkamp (2006) and Veldkamp (2006b). In these models, the information of the agent is endogenous and varies with the underlying state in the economy. In particular, the impact of bad news can be endogenously very large in good times when the information is abundant, so that asset prices fall precipitously and a sudden crash occurs. In a similar vein, in our model the endogenous actions of investors to obtain additional information about the underlying state lead to discrete changes in their expectations of future growth and consequently, large moves in the financial markets.

We solve the model from the perspective of the social planner, who optimally allocates social resources for an acquisition of costly information. As argued in Grossman and Stiglitz (1980), this setup may be difficult to decentralize in the presence of costly information acquisition. However, Veldkamp (2006a) presents a model which motivates the market price implications of a decentralized model with costly information acquisition. Veldkamp (2006a) allows the investors to hire a third party to acquire information on their behalf. The third party pays a fixed cost, determined endogenously in equilibrium, and shares the information to its clients. She shows that as more investors decide to purchase information, the per-investor cost of being informed declines in equilibrium, therefore, in many cases, most investors would get informed. The implications of this equilibrium are very similar to those of the representative agent setup with costly information acquisition featured in our paper.

One of the key implications of our model is that the income volatility predicts future large moves in returns. We provide empirical support that large moves in the stock market can be predicted by the volatility measures in the economy. Specifically, we document a positive correlation of the return jump indicator with lags of conditional variance of consumption. On an annual frequency, the volatility of annual consumption significantly predicts large moves in next-year market returns with an R^2 of 9%, which we show using two alternative measures of consumption volatility, including the usual GARCH model. Further, in the data there is no evidence for predictability of large moves in returns by the levels of the real aggregate variables. We show that the model can match both of these novel and important data features. Earlier evidence in Bates (2000), Pan (2002) and Eraker (2004) documents that market volatility also predicts jumps. In our structural model, the market variance is related to aggregate income volatility, which consequently enables us to match this data feature as well and provide an economic motivation for this empirical finding.

Our target is to match the key evidence on frequency, magnitude and predictability of jumps in the data. In the data we identify 25 years with at least one significant price move (i.e. jump) in daily returns for the 80 year period from 1926 to 2008; hence, the frequency of jump-years is once every 3.3 years.[2] In our sample, we find that the

[2]This provides a conservative estimate for the frequency of return jumps in the data, as there can be more than 1 jump in daily returns in a given year.

relative contribution of jumps to the total return variance is 7.5%, which is consistent with the evidence in Huang and Tauchen (2005) and other studies. We calibrate the model to match these dimensions along with other key asset-market facts. We use standard calibrations of income and preference parameters, while our calibration of learning costs is similar to observation and transactions costs in Abel, Eberly, and Panageas (2007).[3] We show that at the calibrated value of learning cost parameter, investors optimally choose to observe the true state about once every one and a half years. The expenditure on costly learning is 8.5% of the daily income; hence, the per annum expenditure on costly learning is about 0.03% of the aggregate income. Our model generates the mean market return of 6.4%, the volatility of returns of 15.5%, and the model-implied risk-free rate is 1%. Hence, our model can account for the usual equity premium and risk-free rate puzzles in the data. Further, the model with constant aggregate volatility delivers the average frequency of jump-years once every 4.8 years, and the contribution of jumps to return variance of 7.2%. When we allow for time-varying aggregate volatility, the average frequency of jump-years increases to once every 3.4 years, while the relative contribution of jumps to 12%. In standard models with no option to learn the true state for a cost, asset prices do no exhibit jumps. Further, we show that the model with costly learning delivers positive and significant correlation of the large return move indicator with endowment and return variances and zero correlation with endowment growth. The magnitudes of the correlation coefficients are comparable to the data.

A contribution of our paper is to develop an equilibrium asset-pricing model where financial markets display jumps even though the underlying economic input (endowment growth) is smooth. In the standard full-information long-run risks model of Bansal and Yaron (2004), there are no discrete changes in the economy and asset prices do not exhibit jumps. Croce, Lettau, and Ludvigson (2010) consider a similar long-run risks setup where the investors can learn from the history of the data only, i.e. no costly learning. David (1997), Veronesi (1999), Hansen and Sargent (2010), and Ai (2010) consider learning models in which the agents learn about unobserved state variables. It is worth noting that learning considered in these models does not generate jumps in returns. An alternative approach for motivating asset-price jumps is entertained in Eraker and Shaliastovich (2008), Drechsler and Yaron (2010), Liu et al. (2005), Barro (2006), Drechsler (2010) via exogenous jumps in the underlying income process. In these models, asset-price jumps are due to large shifts or disasters in macroeconomic fundamentals. However, our empirical evidence suggests that many asset-price jumps do not coincide with any tangible economic disasters. Consistent with this evidence, in our model asset price jumps are not linked to economic

[3]In the context of rational inattention literature, Sims (2003) feature similar adjustments costs related to information-processing constraint. Costs of acquiring, absorbing and processing information are also used to explain infrequent adjustments of stock portfolio (Duffie and Sun, 1990) or the consumption and saving plans of investors (Reis, 2006).

disasters. We view our approach as complementing the literature which motivates asset-price jumps by macroeconomic jumps/disasters.

The rest of the paper is organized as follows. In the next section we review the empirical evidence on large moves in asset prices in the data. In Section 3 we set up a model and describe preferences, information structure and income dynamics in the economy. In Section 4 we characterize solutions to the optimal learning policy and equilibrium asset valuations. Finally, in Section 5 we use numerical calibrations to quantify model implications for asset-price jumps. Conclusion and Appendix follow.

2 Evidence on Asset Price Jumps

Empirical evidence suggests that asset prices display infrequent large movements which are too big to be Gaussian shocks. In the first panel of Figure 1 we plot the time series of daily inflation-adjusted returns on a broad market index for the period of 1926-2008.[4] Occasional large spikes in the series suggest the presence of large moves (jumps). Consistent with this evidence, the kurtosis of market returns is 21, relative to 3 for Normal distribution, as shown in the first panel of Table 1.

For further evidence on large movements in asset prices, we apply non-parametric jump-detection methods (see Barndorff-Nielsen and Shephard, 2006), used in a stream of papers in financial econometrics. This approach allows us to identify years with one or more large price moves in daily returns.

Let R_T stand for a total return from time $T-1$ to T, and denote $R_{T,j}$ the jth intra-period return from $T-1+(j-1)/M$ to $T-1+j/M$, for $j=1,2,\ldots,M$. The two common measures which capture the variation in returns over the period are the Realized Variation, given by the sum of squared intra-period returns,

$$RV_T = \sum_{j=1}^{M} R_{T,j}^2 \tag{2.1}$$

and the Bipower Variation, which is defined as the sum of the cross-products of the current absolute return and its lag:

$$BV_T = \frac{\pi}{2}\left(\frac{M}{M-1}\right)\sum_{j=2}^{M} |R_{T,j-1}||R_{T,j}|. \tag{2.2}$$

[4]We prorate monthly inflation rate to daily frequency to obtain inflation-adjusted returns from nominal ones. The results for the nominal returns are very similar.

When the underlying asset-price dynamics is a general jump-diffusion process, for finely sampled intra-period returns the Realized Variation RV_T measures the total variation coming both from Gaussian and jump components of the price, while the Bipower Variation BV_T captures the contribution of a smooth Gaussian component only (see, e.g. Barndorff-Nielsen and Shephard, 2006).[5] Hence, these two measures reveal the magnitudes of smooth and jump components in the total variation of returns. A scaled difference between these two measures (Relative Jump statistics) provides a direct estimate of the percentage contribution of jumps to the total price variance:

$$RJ_T = \frac{RV_T - BV_T}{RV_T}. \tag{2.3}$$

Under the assumption of no jump and some regularity conditions, Barndorff-Nielsen and Shephard (2006) show that the joint asymptotic distribution of the two variation measures is conditionally Normal. This allows us compute a t-type statistic to test for abnormally large price movements, which are indicative of jumps. A popular version of this statistic is

$$z_T = \frac{RV_T - BV_T}{\sqrt{\left(\left(\frac{\pi}{2}\right)^2 + \pi - 5\right)\frac{1}{M}TP_t}}, \tag{2.4}$$

where the jump-robust Tri-Power Quarticity measure TP_t estimates the scale of the variation measures and is defined as

$$TP_T = \left(\frac{M^2}{M-2}\right)\left(E(|N(0,1)|^{4/3})\right)^{-3}\sum_{j=3}^{M}|R_{T,j-2}|^{4/3}|R_{T,j-1}|^{4/3}|R_{T,j}|^{4/3}. \tag{2.5}$$

Under the null hypothesis of no jumps and conditional on the sample path, the jump-detection statistic z_T is asymptotically standard Normal. Thus, if the value of z_T is higher than the cut-off corresponding to the chosen significance level, then the test detects at least one abnormal large price move during the period T.

[5]More precisely, under some technical conditions,

$$\lim_{M\to\infty} RV_T = \int_{T-1}^{T}\sigma_p^2(s)ds + \sum_{j=1}^{N_T}k_{T,j}^2, \quad \lim_{M\to\infty} BV_T = \int_{T-1}^{T}\sigma_p^2(s)ds,$$

where $\sigma_p(s)$ is the instantaneous volatility of the Brownian motion component of the price, $k_{T,j}$ is the jump size and N_T is the number of jumps within the period T.

To calculate the jump-detection statistics over a year, we use the data on 266 daily returns, on average.[6] We focus on the annual jump detection frequency, as it is well-recognized that one requires large number of intra-period observations to compute jump statistics. $M = 266$ is a typical number in high-frequency studies, where it roughly corresponds to using 5-minutes returns to compute daily (24 hours) statistics. Huang and Tauchen (2005) discuss the performance of the tests in finite samples.

On Figure 1 we plot daily inflation-adjusted market returns and the corresponding years detected by jump-detection statistic for the period of 1927 - 2008, while Figure 2 depicts the corresponding jump statistics z_T. Notably, high values of z_T above the corresponding cut-off point indicate the presence of large moves in daily returns. At the 1% significance level, we identify 25 years with at least one significant move in daily asset prices. 8 of those jump-years occur before 1945, that is, 8 out 18 years from 1927 to 1945 contain one or more large moves in daily returns. The remaining 17 of the jump-years occur in the post-war period of 63 years. Some of the salient jump dates include 1982, 1987, 1991, and 2003. The relative contribution of large movements to the total return variation, as measured by the average relative jump measure RJ, is 7.5%. This estimate is consistent with other studies.

2.1 Predictability of Large Price Moves

In this section, we provide empirical evidence that macroeconomic volatility and the market return variance can predict large asset-price moves in the data. On the other hand, in the data there is no persuasive evidence for the link between large moves in returns and the growth rates of aggregate macroeconomic variables at all leads and lags. That is, at the considered frequencies of large moves in returns, jumps in asset prices neither coincide with significant changes in the real economy, nor can they be predicted by them. This empirical evidence has important implications for identifying the sources of jump risk in financial markets, which motivate our model setup. The inputs in our model (i.e. endowment) are Gaussian. While there are no jumps in the real side of the economy, learning and costly information acquisition will trigger endogenous jumps in financial markets. In contrast, works by Eraker and Shaliastovich (2008), Drechsler and Yaron (2010), Pan (2002), Barro (2006) incorporate jumps into the exogenous inputs in the model, namely the consumption (i.e. endowment) process. Our evidence, of course, suggests that on average there is close to zero correlation between jumps in growth rates and asset prices.

On top panel of Figure 3 we plot correlation coefficients of jump-year indicators with annual consumption growth rate, its conditional variance and the conditional

[6]For predictability regressions, we also construct jump statistics on monthly and quarterly frequencies.

variance of market returns, up to 5 year leads and lags.[7] We further provide the correlation estimates and the standard errors in the top panel of Table 2. The correlations of large move indicators with lagged aggregate volatility are consistently positive and reach the $20-30\%$ range. Similarly, high market variance predicts an increase in future jump probability, and the correlations of market variance with contemporaneous and future jump-year indicators are about 10%. The jump-year indicator correlations with future variance measures decrease to zero after 2-3 years. Further, we do not find any strong evidence for the link between the asset-price jumps and contemporaneous or past consumption growth rates. The correlation coefficients for the jump-year indicator with consumption growth rate are negative at one and two year lags and are around -0.1. They are essentially zero at three year lags and beyond.

The above predictability patterns are even stronger at quarterly and monthly frequencies, as the persistence of the variance measure and the frequency of identified jump periods increase. As consumption data is not available at such frequencies for a long historical sample, we use the industrial production index growth, whose monthly and quarterly observations are available from 1930s.[8] On the bottom panels of Figure 3 we plot the lead-lag correlations of the jump indicator with levels and conditional volatilities of industrial production growth rate and variance of the market return at quarterly and monthly frequencies. The results present robust evidence for predictability of asset-price jump periods by the variance measures and absence of persuasive link between the asset-price jumps and contemporaneous or past levels of the real economic growth. We are going to match these jump predictability patterns in the model, alongside other key macroeconomic and financial data features.

To sharpen quantitative results, we construct a measure of macroeconomic volatility based on the financial markets data; similar volatility measures are entertained by Bansal, Kiku, and Yaron (2007) and Bansal, Khatchatrian, and Yaron (2005). We regress annual consumption growth on its own lag, the lags of market price-dividend ratio and junk bond spread and extract consumption innovation. The square of this innovation is further projected on the price-dividend ratio and junk bond spread, so that the fitted value $\hat{\sigma}_T^2$ captures the level of ex-ante aggregate volatility in the economy. The results of the two projections are summarized in the top panel Table 3. The R^2s are in excess of 20%, and the signs of the slope coefficients are economically intuitive: low asset valuations and high bond spreads predict low expected growth and high aggregate volatility.

We use the extracted factor $\hat{\sigma}_T^2$ to forecast next year jump indicator statistic. The probit regression of the next-period jump indicator on current measure of macroeco-

[7] Conditional variance computations are based on AR(1)-GARCH(1) fit.

[8] On annual frequency, the correlation of growth rates in consumption and industrial production is 0.55, while the correlation of their conditional variances is 0.84.

nomic volatility yields a statistically significant coefficient on $\hat{\sigma}_T^2$ with a t-statistics in excess of 2.5, and R^2 of 9%. Specifically,

$$\hat{Pr}(JumpIndicator_{T+1}) = \Phi\left(\underset{(0.21)}{-0.84} + \underset{(468.00)}{1186.23}\hat{\sigma}_T^2\right),$$

where $JumpIndicator_T$ is equal to 1 if year T is flagged as a jump-year and 0 otherwise. On Figure 2 we plot the jump-detection statistics z_T itself and the fitted probability of contemporaneous jump. The spikes in fitted probabilities broadly agree with large values of the jump statistics, even for the 1955-1980 period when no significant price moves were detected.

For robustness, we also check the results using GARCH measure of annual consumption volatility in the data. The bottom panel of Table 3 shows that the estimated aggregate consumption volatility is very persistent in the data. The probit estimation of predictability of the future jump-year indicator is given by,

$$\hat{Pr}(JumpIndicator_{T+1}) = \Phi\left(\underset{(0.20)}{-0.76} + \underset{(406.43)}{875.84}\hat{\sigma}_T^2\right),$$

so that the consumption volatility is a statistically significant predictor of future jump years with $t-$statistics of 2.16, and the R^2 of 6.3%. [9]

While consumption volatility forecasts jump periods, the level of consumption growth rate does not seem to predict future jump years in the data. In Table 4 we report the R^2 in the probit regression of next-period jump-year indicator on the realized consumption growth. The slope coefficient is insignificant from 0, and the R^2 is below 1%. We show that our calibrated model can match well this quantitative evidence on the predictability of jumps in returns.

Predictability of future jumps by the consumption variance is a novel dimension of this paper. Predictability of future return jumps by market variance is consistent with the evidence in earlier studies which estimate parametric models of asset-price dynamics, see Bakshi et al. (1997), Bates (2000), Pan (2002), Eraker (2004) and Singleton (2006). We provide further discussion of these model specifications in Section 5.5.

3 Model Setup

Our model builds on the long-run risks framework developed in Bansal and Yaron (2004), where the investor has full information about the economy. In contrast, we

[9] For robustness, we checked the regressions using post-war sample, and found very similar results.

assume that investors do not observe all the relevant state variables, and hence there is an important role for learning about the true underlying state of the economy. The exogenous endowment process is Gaussian and does not contain any exogenous jumps. However, we show that the optimal actions of the agents to learn the unobserved states for a cost can lead to asset-price dynamics which exhibits jumps.

3.1 Preferences and Information

Denote $\mathcal{I}_t$ the beginning-of-period information set of the agent, which includes current and past observed variables. The information set by the end of the period is endogenous and depends on the decision of investors to learn about the true state. Let us introduce a binary choice indicator $s_t \in \{0, 1\}$, which is equal to one if the agent learns about the true state for a cost in period t, and zero otherwise. Let $\mathcal{I}_t(s_t)$ be the time-t (end-of-period) information set following a choice s_t. With no learning about the true state ($s_t = 0$), the end-of-period information set coincides with that in the beginning of the period: $\mathcal{I}_t(0) \equiv \mathcal{I}_t$. On the other hand, when $s_t = 1$, investors acquire new information during the day which enriches their information set: $\mathcal{I}_t(1) \supset \mathcal{I}_t$. Further, let E_t denote the conditional expectation with respect to the information set $\mathcal{I}_t$, while denote $E_t^{s_t}$ the conditional expectation based on the information following a binary choice $s_t : E_t^{s_t}(.) \equiv E[.|\mathcal{I}_t(s_t)]$.

We consider recursive preferences of Epstein and Zin (1989) over the uncertain consumption stream, with the intertemporal elasticity of substitution parameter set to one:

$$\begin{aligned} U_t &= C_t^{1-\beta} \left(\mathbf{J}_t^{s_t}(U_{t+1})\right)^{\beta}, && (3.1) \\ \mathbf{J}_t^{s_t}(U_{t+1}) &= \left(E_t^{s_t} U_{t+1}^{1-\gamma}\right)^{\frac{1}{1-\gamma}}. && (3.2) \end{aligned}$$

C_t denotes consumption of the agent and $\mathbf{J}^{s_t}(U_{t+1})$ is the certainty equivalent function which formalizes how the agent evaluates uncertainty across the states. Parameter β is the subjective discount factor and γ is the risk-aversion coefficient of the agent. Note that certainty equivalent function depends on the choice indicator $s_t \in \{0, 1\}$, as the information set of the agent is different whether the investors learn about the true state ($s_t = 1$) or not ($s_t = 0$).

To derive our model implications, we solve the model from the perspective of the social planner, who optimally allocates social resources for acquisition of costly information. This setup, as is highlighted in Grossman and Stiglitz (1980), may be difficult to decentralize in the presence of costly information acquisition. However, Veldkamp (2006a) presents a mechanism which motivates the decentralized market price implications of model with costly information acquisition. Veldkamp (2006a) allows the investors to hire a third party to acquire information on their behalf. The third party pays a fixed cost, determined endogenously in equilibrium, and shares the information to its clients. She shows that as more investors decide to purchase information, the per investor cost of being informed declines in equilibrium. Therefore, in many cases most investors would get informed. While we do not explicitly introduce such information channels to keep the model tractable and maintain the focus on large moves in returns, the implications of this equilibrium are similar to those of the representative agent setup with costly information acquisition featured in our paper.

3.2 Social Planner Problem

Consider the life-time utility of the agent $U_t(s_t)$ for a given learning choice of the social planner $s_t \in \{0, 1\}$:

$$U_t(s_t) = C_t(s_t)^{1-\beta} \left(\mathbf{J}_t^{s_t}(U_{t+1})\right)^{\beta}, \tag{3.3}$$

where U_{t+1} is the optimal utility tomorrow, and $C_t(s_t)$ denotes a choice specific consumption of the agent. The risk-sensitive certainty equivalent operator $\mathbf{J}_t^{s_t}(U_{t+1})$ is specified in equation (3.2).

The objective of the social planner is to maximize the certainty equivalent of the life-time utility of the agent $U_t(s_t)$ with respect to the beginning-of-period information set $\mathcal{I}_t$ by choosing whether or not to learn about the true state for a cost:

$$s_t^* = arg \max_{s_t} \left\{\mathbf{J}_t(U_t(s_t))\right\}. \tag{3.4}$$

The true value of the state is not known to the planner in the beginning of the period. As the agents are risk-sensitive to the new information about the state, the planner chooses to learn about the state for a cost if the certainty equivalent of the agent's life-time utility with learning is bigger than the life-time utility without learning. Following a decision to learn, the social planner then uses part of the endowment to pay the learning cost.

Denote Y_t the aggregate income process. Then, the budget constraint of the social planner states that the aggregate income is equal to consumption and learning cost expenditures:

$$Y_t = C_t(s_t) + s_t \xi_t. \tag{3.5}$$

The learning cost ξ_t represents the resources required to acquire and process the new information about the underlying economic state. We interpret this cost to be a social cost in terms of research costs borne by social institutions (e.g Treasury and the Federal Reserve Bank) to gather information about the underlying state of the economy. An alternative interpretation is presented in Veldkamp (2006) who argues that these costs may be related to media costs (such as hiring a journalist) to gather information about economy. For analytical tractability, we make ξ_t proportional to the aggregate income:

$$\xi_t = \chi Y_t, \tag{3.6}$$

for $0 \leq \chi < 1$. This specification preserves the homogeneity of the problem and simplifies the solution of the model.

In Appendix A.1 we show that in equilibrium, the life-time utility of investors following learning choice s_t is proportional to the level of income,

$$U_t(s_t) = \phi_t(s_t) Y_t, \quad \text{for } s_t \in \{0, 1\}. \tag{3.7}$$

where the utility per income ratio $\phi_t(s_t)$ satisfies the following recursive equation:

$$\phi_t(s_t) = (1 - s_t\chi)^{1-\beta} \left(E_t^{s_t} \left[\phi_{t+1} \frac{Y_{t+1}}{Y_t} \right]^{1-\gamma} \right)^{\frac{\beta}{1-\gamma}}. \tag{3.8}$$

Learning about the true state has two effects on the utility of investors. First, the agent's consumption drops as part of the aggregate endowment is sacrificed to cover the learning costs. This decreases the agent's utility, as evident from examining the first bracket in the expression above. On the other hand, learning enriches the information set of investors, and the ensuing reduction in the uncertainty about future economy may increase their utility (second part of the expression (3.8)). The net effect depends on the attitude of investors to the timing of resolution of uncertainty and the magnitude of learning costs, as we discuss in detail in Section 4.

In Appendix A.2 we show that the equilibrium discount factor M_{t+1} depends on the income growth, future life-time utility and the endogenous information set of the agent:

$$M_{t+1} = \beta \left(\frac{Y_{t+1}}{Y_t} \right)^{-1} \frac{U_{t+1}^{1-\gamma}}{E_t^{s_t^*}(U_{t+1}^{1-\gamma})}. \tag{3.9}$$

Hence, we can solve for the price of any asset traded in the economy using a usual equilibrium Euler equation

$$E_t^{s_t^*} \left[M_{t+1} R_{i,t+1} \right] = 1, \tag{3.10}$$

where $R_{i,t+1}$ is the return on the asset, and s_t^* is an equilibrium costly learning choice.

3.3 Income Dynamics

The log income growth rate process incorporates a time-varying mean x_t and stochastic volatility σ_t^2:

$$\Delta y_{t+1} = \mu + x_t + \sigma_t \eta_{t+1}, \tag{3.11}$$
$$x_{t+1} = \rho x_t + \varphi_e \sigma_t \epsilon_{t+1}, \tag{3.12}$$
$$\sigma_{t+1}^2 = \sigma_0^2 + \nu(\sigma_t^2 - \sigma_0^2) + \sigma_w \sigma_t w_{t+1}. \tag{3.13}$$

where η_t, ϵ_t and w_t are independent standard Normal innovations. Parameters ρ and ν determine the persistence of the mean and variance of the income growth rate, respectively, while φ_e and σ_w govern their scale. The empirical motivation for the time-variation in the conditional moments of the income process comes from the long-run risks literature, see e.g. Bansal and Yaron (2004), Hansen, Heaton, and Li (2008) and Bansal and Shaliastovich (2007).

We assume that the volatility σ_t^2 is known to the agent at time t, which can be justified as the availability of high-frequency data allows for an accurate estimation of the conditional volatility in the economy. On the other hand, the true expected income state x_t is not directly observable to the investors. The investors can learn about the state from the observed data using standard filtering techniques, and they also have an additional option to pay a cost to learn its true value. This setup is designed to capture the intuition that some of the key aspects of the economy are not directly observable, but the agents can learn more about them through additional costly exploration.

To solve the learning problem of the agents, we follow a standard Kalman Filter approach.[10] Given the setup of the economy, the beginning-of-period information set of the agent consists of the history of income growth, income volatility and observed true states up to time $t : \mathcal{I}_t = \{y_\tau, \sigma_\tau^2, s_{\tau-1} x_{\tau-1}\}_{\tau=1}^t$. If the agent does not learn the true state in period t, the end-of-period information set is the same as in the beginning of the period: $\mathcal{I}_t(0) = \mathcal{I}_t$. On the other hand, if the agent learns the true value of the expected income state, the information set immediately adjusts to include x_t : $\mathcal{I}_t(1) = \mathcal{I}_t \cup x_t$. Define a filtered state $\hat{x}_t(s_t)$, which gives the expectation of the true state x_t given the information set of the agent and the costly learning decision s_t:

$$\hat{x}_t(s_t) = E_t^{s_t}(x_t), \tag{3.14}$$

and denote $\omega_t^2(s_t)$ the variance of the filtering error which corresponds to the estimate $\hat{x}_t(s_t)$:

$$\omega_t^2(s_t) = E_t^{s_t}(x_t - \hat{x}_t(s_t))^2. \tag{3.15}$$

[10] For Kalman Filter reference and applications, see Lipster and Shiryaev (2001), David (1997), and Veronesi (1999).

If the agent chooses to learn about the true state, we obtain, naturally, that $\hat{x}_t(1) = x_t$ and $\omega_t^2(1) = 0$.

Given the history of income, income volatility and past observed expected growth states, the agent updates the beliefs about the unobserved expected income state in a Kalman filter manner. Indeed, as the income volatility is observable, the evolution of the system is conditionally Gaussian, so that the expected mean and variance of the filtering error are the sufficient statistics to track the beliefs of the agent about the economy. Specifically, for a given choice indicator s_t today, the evolution of the states in the beginning of the next period follows from the one-step-ahead innovation representation of the system (3.11)-(3.13):

$$\begin{aligned}
\Delta y_{t+1} &= \mu + \hat{x}_t(s_t) + u_{t+1}(s_t), && (3.16)\\
\hat{x}_{t+1}(0) &= \rho \hat{x}_t(s_t) + K_t(s_t) u_{t+1}(s_t), && (3.17)\\
\omega_{t+1}^2(0) &= \sigma_t^2 \left(\varphi_e^2 + \rho^2 \frac{\omega_t^2(s_t)}{\omega_t^2(s_t) + \sigma_t^2} \right), && (3.18)
\end{aligned}$$

where the gain of the filter is equal to

$$K_t(s_t) = \frac{\rho \omega_t(s_t)^2}{\omega_t(s_t)^2 + \sigma_t^2}. \tag{3.19}$$

The filtered consumption innovation $u_{t+1}(s_t) = \sigma_t \eta_{t+1} + x_t - \hat{x}_t(s_t)$ is learning choice specific, and contains short-run consumption shock and filtering error. The two cannot be separately identified unless the agent learns the true x_t, in which case the filtered consumption innovation is equal to the short-run consumption shock, $u_{t+1}(1) = \sigma_t \eta_{t+1}$. Recall that the variance shocks w_{t+1} are assumed to be independent from the income innovations at all leads and lags. That is, future volatility shocks do not help predict tomorrow's expected income, and neither can learning about x_t affect the agent's beliefs about future income volatility. Therefore the dynamics of the income volatility is independent of the learning choice of the agent and follows (3.13). In particular, if income volatility is constant, we obtain a standard result that the variance of the filtering error $\omega_t^2(0)$ increases in a deterministic fashion since the last costly learning. On the other hand, when income volatility is stochastic, the variance of the filtering error fluctuates over time and is high at times of heightened aggregate volatility.

The key novel economic channel in our model is a discrete adjustment in agent's expectation about future growth, $\hat{x}_t$, at times when agent decides to learn the true state, that is, when $s_t = 1$. Indeed, if investors decide to pay a cost to learn the true state, the expected income growth and variance of the filtering error are immediately adjusted to reflect the new information. We can then express the values of the states in the following way:

$$\hat{x}_{t+1}(s_{t+1}) = s_{t+1}x_{t+1} + (1 - s_{t+1})\hat{x}_{t+1}(0), \tag{3.20}$$

$$\omega^2_{t+1}(s_{t+1}) = (1 - s_{t+1})\omega^2_{t+1}(0). \tag{3.21}$$

In equilibrium, such revisions in expected growth state endogenously trigger large moves in asset-prices which look like jumps, even though the fundamental income process is smooth Gaussian. We characterize the optimal decision to learn for a cost and the asset-pricing implications in the next section.

4 Model Solution

4.1 Optimal Costly Learning

We solve for the equilibrium life-time utility of the agent and characterize the optimal decision to learn about the expected growth for a cost. In Appendix A.4 we show that the life-time utility of the agent depends on the beginning-of-period information and, at times when the agent chooses to learn about the true state for a cost, on the true value of the expected income growth. In particular, as the volatility and income growth shocks are assumed to be independent, we can separate the expected growth and volatility components, so that the solution to the life-time utility per income ratio can be written in the following way:

$$\phi_t(s_t) = e^{B\hat{x}_t(s_t)+f(s_t,\sigma^2_t,\omega^2_t(0))}. \tag{4.1}$$

The sensitivity of the utility to expected income growth B is independent of the decision to learn for a cost and is given by

$$B = \frac{\beta}{1 - \beta\rho}. \tag{4.2}$$

The volatility function $f(s_t, \sigma^2_t, \omega^2_t(0))$ depends on the learning choice s_t, the exogenous income volatility σ^2_t and the beginning-of-period filtering variance $\omega^2_t(0)$, as well as the risk aversion of the agent γ, learning cost χ and other model and preference parameters. The recursive solution to this volatility component is provided in the Appendix equation (A.27).

Let us characterize the solution to the optimal costly learning decision of the agent and use it to illustrate some of the important features of the model. The agent chooses to observe the true state if the ex-ante life-time utility with learning exceeds the utility with no learning about the true state. Given the equilibrium solution to the life-time utility per income ratio in (4.1), the investor's life-time utility with no learning is equal to

$$\phi_t(0) = e^{B\hat{x}_t(0)+f(0,\sigma_t^2,\omega_t^2(0))}, \tag{4.3}$$

while the ex-ante life-time utility with costly learning is given by

$$\mathbf{J}_t(\phi_t(1)) = e^{B\hat{x}_t(0)+\frac{1}{2}(1-\gamma)B^2\omega_t^2(0)+f_t(1,\sigma_t^2,\omega_t^2(0))}. \tag{4.4}$$

The life-time utility of the agent depends on the estimate of expected growth $\hat{x}_t$ and the volatility factors σ_t^2 and $\omega_t^2(0)$, as is evident from the above equations. Further, the expected growth enters symmetrically across the ex-ante life-time utilities with and without costly learning (see equations (4.3) and (4.4)). The optimal decision to pay a cost to learn is optimally determined by evaluating the life-time utility across the two decisions, that is, comparing (4.3) and (4.4). From this comparison we find that the optimal costly learning choice s_t^* is given by,

$$\begin{aligned} s_t^* &= 1[\mathbf{J}_t^0(\phi_t(1)) > \phi_t(0)] \\ &= 1\left[\frac{1}{2}(1-\gamma)B^2\omega_t^2(0) + f_t(1,\sigma_t^2,\omega_t^2(0)) > f_t(0,\sigma_t^2,\omega_t^2(0))\right]. \end{aligned} \tag{4.5}$$

This decision implies the following important result:

Result 1: *The optimal costly learning rule depends only on the volatility states $w_t^2(0)$ and σ_t^2, and it does not depend on the expected growth $\hat{x}_t$.*

Indeed, in our model learning for a cost gives an agent a real option to reduce the uncertainty about the estimate of expected growth. Because of this option feature and a Gaussian dynamics of the economy, we obtain that the optimal decision only depends on the volatility states, as is evident from equation (4.5).

In general, the timing of costly learning is stochastic and determined by the income volatility σ_t^2 and variance of filtering error ω_t^2, as well as the model and preference parameters. In a particular case when income volatility is constant, the optimal costly learning rule considerably simplifies, as we demonstrate in the next result:

Result 2: *When income volatility is constant, investors optimally learn about the true state for a cost at constant time intervals.*

Indeed, when income shocks are homoscedastic, the optimal learning rule is driven only by the variance of the filtering error, and the agent chooses to exercise costly learning option and reduce the uncertainty about expected growth when this variance is high enough. However, since income volatility is constant, the variance of the filtering error is a deterministic function of time since the last costly learning.

Hence, investors optimally choose to learn the true state for a cost at determined in equilibrium constant time intervals. This result is similar to Abel et al. (2007), who show in a partial equilibrium setting that investors optimally choose to update their information in the presence of observation costs at equally spaced points in time.

The frequency of costly learning in our model depends on the model and preference parameters. In particular, we can show that the optimal costly learning policy depends on the risk-aversion and learning cost parameters in an intuitive way:

Result 3: *When income volatility is constant and agents prefer early resolution of uncertainty, the frequency with which agents learn for a cost increases when risk-aversion parameter increases, or when learning becomes less costly.*

The formal proof for these comparative statics results are shown in the Appendix A.6, and the importance of the preference for early resolution of uncertainty is discussed in detail in the next Section.

In the time-varying volatility model, the optimal learning policy depends both on the variance of the filtering error and the stochastic volatility of income growth, which complicates the formal comparative statics analysis of the model. In particular, the frequency of costly learning is no longer constant and depends on the conditional volatility of income growth. Using numerical solution to the model, we document the following result:

Result 4: *When income volatility is time-varying, agents exercise costly learning option more frequently when income volatility is high.*

While we do not provide a formal proof for this finding, it appears to be quite intuitive and follow from our previous discussion of the homoscedastic model. Indeed, as filtering uncertainty accumulates very quickly at times of heightened aggregate volatility, the incentives to learn and reduce the uncertainty are thus bigger in high relative to low income volatility periods. Hence, the frequency of costly learning is increasing in income volatility. This result implies that costly learning times are predictable by the income volatility in economy, which, we show, provides an economic basis for the predictability of asset-price jumps in financial markets.

4.2 Preferences and Information Acquisition

One of the key ingredients of the model which determines the optimal learning choice is the preferences of the agent. In our setup, the agent has recursive preferences when the risk-aversion coefficient γ is different from 1; when $\gamma = 1$ preferences collapse to a standard expected log utility case. The incentive to learn the unobserved state for a cost critically requires the recursive preferences of the agent, and in particular, the preference for early resolution of uncertainty ($\gamma > 1$). We establish the following important result:

Result 5: *With standard expected utility preferences, the agent is indifferent to the timing of the resolution of uncertainty, and as a consequence, has no incentive to learn for a cost.*

Indeed, consider a case when learning costs are zero, that is $\chi = 0$, so that and the consumption of the agent is equal to the income. Then, the utility of the agent corresponding to the indicator variable $s_t \in \{0, 1\}$ satisfies

$$U_t(s_t) = E_t^{s_t} \sum_{j=0}^{\infty} \beta^j u(Y_{t+j}). \tag{4.6}$$

The optimal learning policy in the expected utility case is based on the ex-ante expected utility given the beginning of period information. Applying the law of iterated expectations, we find that the ex-ante utility of the agent with the new information is equal to the life-time utility without the new information:

$$E_t^0 U_t(1) = E_t^0 \left(E_t^1 \sum_{j=0}^{\infty} \beta^j u(Y_{t+j}) \right) = E_t^0 \sum_{j=0}^{\infty} \beta^j u(Y_{t+j}) \equiv U_t(0). \tag{4.7}$$

In expectation, new information does not increase the utility of the agent. Therefore, with power utility investors have no incentive to gather new information about the economy, even if this information is costless. On the other hand, in Appendix A.3 we show that with recursive utility investors have incentives to learn the new information as long as they have a preference for early resolution of uncertainty ($\gamma > 1$). In this case, the value of learning the new information can exceed the immediate learning costs, so the agents optimally choose to pay a cost and acquire the information. This underscores the economic importance of the preference for early resolution of uncertainty in learning models.

4.3 Risk Compensation and Asset Prices

Using the solution to the agent's learning model, we can express the equilibrium discount factor (3.9) in terms of the underlying variables in the economy; see equation

(A.30) in Appendix for details. In particular, the innovation into the log discount factor satisfies,

$$
\begin{aligned}
m_{t+1}(s_t) - E_t^{s_t} m_{t+1}(s_t) = & - \left(1 + (\gamma - 1)(1 + BK_t(s_t))\right) u_{t+1}(s_t) \\
& - (\gamma - 1)(f_{t+1} - E_t^{s_t} f_{t+1}) - (\gamma - 1) B s_{t+1}^{*}(x_{t+1} - \hat{x}_{t+1}(0)).
\end{aligned}
\tag{4.8}
$$

In our economy the agent is exposed to three sources of risk: Gaussian consumption shocks $u_{t+1}(s_t)$, volatility shocks $(f_{t+1} - E_t^{s_t} f_{t+1})$ and discrete revisions in the true state $s_{t+1}^{*}(x_{t+1} - \hat{x}_{t+1}(0))$. The key novel dimension of the paper is the endogenous costly learning by the agent s_{t+1}^{*}, which triggers large adjustments, i.e. jumps, in the discount factor and equilibrium asset prices when investors decide to pay a cost and learn the true state. The non-trivial costly learning policy is essential to generate endogenous jumps in our model in the absence of corresponding jumps in economic fundamentals. The special cases of our model when agents always know the true state ($s_t \equiv 1$) or never know the true value of the state ($s_t \equiv 0$) do not lead to asset-price jumps when model inputs are smooth.

Indeed, when the agent knows the true expected growth state at all times ($s_t \equiv 1$), our model collapses to a standard long-run risks setup of Bansal and Yaron (2004). In this case, the price of short-run consumption risk is γ, the price of long-run risk is $(\gamma - 1)B$ and the price of volatility risk is constant and provided in the above study. In a standard long-run risks model, asset prices do not exhibit jumps as economic inputs are smooth and shocks are Normal. With costly learning, our model features an additional source of risk due to the discrete revisions of the expected growth state, $s_{t+1}^{*}(x_{t+1} - \hat{x}_{t+1}(0))$. A discrete revision in the expected growth can be quite large, in absolute value, as agent's estimate of expected growth moves away from the true underlying state between the relatively infrequent times of costly learning. Notably, the revision of the expected growth introduces an endogenous jump risk in the economy, as both the timing and the magnitude of the discount factor jump is determined in equilibrium by underlying states and model and preference parameters. Eraker and Shaliastovich (2008), Drechsler and Yaron (2010), Barro (2006) introduce exogenous jumps in the underlying endowment process, which trigger corresponding jumps in the equilibrium asset prices. However, as discussed earlier in the empirical section, an important feature of the data is that many of the jumps in asset markets do not seem to coincide with jumps in the real economy. Motivated by this empirical evidence, our model specification features instead smooth income dynamics and endogenous jumps in asset prices due to optimal actions of investors to pay a cost and learn the true state.

An alternative case considered in the literature is when investors never know the true value of the state ($s_t \equiv 0$), and they optimally estimate the unknown expected growth using the history of the data. Such an approach, within the long-run risks setup, is pursued in Croce et al. (2010), while Ai (2010), David (1997) and Veronesi (1999) develop a model where the agents learn about the regime shifts. It is worth emphasizing that learning considered in these models does not generate jumps in the asset prices. Hansen and Sargent (2010) consider an alternative approach and introduce preference for robustness in agent's learning. This, they show, magnifies the level and variation in risk premia relative to the standard models. However, to deliver jumps in the asset-prices, such specifications require exogenous jumps in the inputs of the economy (endowment process), as shown in Liu et al. (2005) and Drechsler (2010). On the other hand, our approach complements the approach taken in Veldkamp (2006b) and Van Nieuwerburgh and Veldkamp (2006), where information is endogenous and varies with the state in the economy. For example, in Veldkamp (2006b) the impact of bad news can be endogenously very large in good times when the information is abundant, so that asset prices move sharply. In a similar vein, in our model, the endogenous actions of investors to obtain additional information leads to discrete changes in expectations about future growth and therefore large asset-price movements.

To bring our model implications closer to the data, we calibrate a dividend asset, which is a levered claim with a dividend stream proportional to income growth:

$$\Delta d_t = \mu + \varphi_d(\Delta y_t - \mu). \tag{4.9}$$

Bansal and Yaron (2004) specify dividend dynamics which includes idiosyncratic dividend shock. The specification above is simpler as it does not require extension of the model to multivariate Kalman filter, but preserves model results and intuition.

Using the equilibrium solution to the discount factor in (4.8) and the Euler condition (3.10), we can solve for the equilibrium log price-dividend ratio,

$$v_t(s_t) = H\hat{x}_t(s_t) + h(s_t, \sigma_t^2, \omega_t^2(0)), \tag{4.10}$$

where the solutions for a constant H and the volatility component $h(s_t, \sigma_t^2, \omega_t^2(0))$ are given in Appendix A.5.

The asset valuations depend on filtered or, if $s_t = 1$, true expected income growth, and the volatility factors in the economy. When investors pay a cost and learn the true state, their estimate of expected growth can change substantially from what it was a period ago. The equilibrium price-dividend ratio responds to this change in the expected growth state magnified by H. For example, as H is positive in the model, when the true x_t is much lower than what the agent expected, asset prices can fall sharply. This discrete decline in asset prices, triggered by optimal decision of investors to learn the true state, is detected as a jump in the financial markets.

Hence, the distribution of asset prices in our economy is heavy-tailed, even though the underlying macroeconomic inputs are smooth Gaussian.

The probability of costly learning, and consequently, large asset-price moves depends on the volatility states in the economy. In particular, when aggregate volatility is high, investors learn for a cost more often, which triggers more frequent large moves in returns. Hence, asset-price jump times are predictable by the aggregate macroeconomic volatility. Further, as the volatility of equilibrium returns is positively related to aggregate volatility, the model can also explain the predictability of future asset-price jumps by the variance of the market return. Notably, in the model the probability of asset-price jumps is not related to the level of real economy, so the financial jumps are not predicted by the endowment growth rate.

In the next Section we calibrate the economy and show that the model-implied jump implications are quantitatively consistent with the data.

5 Model Output

5.1 Model Calibration

The model is calibrated on a daily frequency. The baseline calibration parameter values, which are reported, annualized, in Table 5, are similar to the ones used in standard long-run risks literature (see e.g. Bansal and Yaron, 2004). Specifically, we set the persistence in the expected income growth $\rho^{12\times 22}$ at 0.4 on annual frequency. The choice of φ_e and σ_0 ensures that the model matches the annualized aggregate volatility of about 1.4%, while the annualized volatility persistence is set to 0.77. To calibrate dividend dynamics, we set the leverage parameter of the corporate sector φ_d to 5. We calibrate the model on a daily frequency and then time-aggregate to annual horizon. Table 6 shows that we can successfully match the unconditional mean, volatility, and auto-correlations of the endowment dynamics in the data.

As for the preference parameters, we let the subjective discount factor equal 0.997 and set the risk aversion parameter at 10. The learning expenditure includes the resources that the investors spend to acquire and process the information about true value of the underlying economic state, which includes opportunity costs of time and effort. We calibrate the cost parameter similar to observation and information costs emphasized in Abel et al. (2007). They set the observation cost to a fraction of annual income, and show that investors choose to update their information once every 8 months. Motivated by the empirical evidence on the frequency of jumps, we calibrate expenditure on costly learning accounts to be 0.03% of annual aggregate income (8.5% of daily income). At this level of learning costs, investors are willing to optimally learn the true state about once every one and a half year. Even though the level of costly expenditure appears to be quite small, we show that it has important implications for the distributions of the asset prices and financial jumps in the economy, which are impossible to obtain when the costly learning option is absent. Naturally, the calibration of the learning cost parameter is sensitive to the assumed values of other model parameters, such as risk aversion and level of the volatility shocks.

As the model does not admit convenient closed-form solutions for the asset prices, we use calibrated parameter values to solve the model numerically. We first start with the model specification when income volatility is constant, that is, $\sigma_t = \sigma_0$. As shown in Result 2, the optimal learning policy in this setup is purely time-dependent, and the agents learn about the true state at constant, determined in equilibrium, time intervals. The details of the model solution for the optimal learning choice and the equilibrium asset prices are provided in Appendix A.6. In the general case when income volatility is time-varying, we first put income and filtering volatility states on a fine grid and search for a fixed-point solutions to the recursive volatility function equations (A.29) and (A.35). It turns out that the optimal solutions to the volatility components can be very accurately approximated by the linear functions of the volatility states, where the loading coefficients are learning-choice specific. We utilize these linear approximations to speed up and stabilize numerical computations.

5.2 Constant Volatility Case

Table 1 reports asset-pricing implications of the model with constant income volatility. Model-implied mean and volatility of market returns are 6.7% and 15.5%, respectively, and the risk-free rate is 1.1%, which match the empirical data. Hence, the model can account for the usual equity premium, the return volatility and the risk-free rate puzzlez. The model specification where agents have no option to pay a cost and learn the true state delivers comparable values for the first two moments of the return distributions; this is consistent with Bansal and Yaron (2004) who show that a standard long-run risks model without costly learning can explain the unconditional mean and volatility of returns.

The costly learning option, however, is central to account for large asset-price moves and the heavy tails of the distribution of returns. To illustrate the jump implications of our model, on Figure 4 we plot a typical simulation of the economy for 80 years. The log income growth is conditionally Normal, and the filtered expected income state closely tracks the true state with a correlation coefficient in excess of 70%. About every 2 years the agent pays the cost and learns the true state. The revision in expectations about future income growth triggers proportional adjustments to the equilibrium asset prices, as can be seen from equation (4.10). In presence of persistent expected growth shocks, asset prices are very sensitive to changes in expected income state. Therefore, even small deviations in the filtered state from the truth, when uncovered, can lead to large changes in asset valuations which are empirically detected as jumps.[11] As shown in Table 1, the frequency of detected jump-years is about once every 4.8 years, and the contribution of jumps to the total return variation is 7.2%, which is consistent with the data. Due to large moves in returns, the unconditional distribution of returns is heavy-tailed: the kurtosis of the return distribution in the model is equal to 18, which is close to the empirical estimate of 21 in the data. Notably, as the market volatility is constant, the heavy tails in return distribution obtain through the discrete adjustments to the asset prices due to costly learning.

Large asset-price moves cannot be obtained in the model where the agent had no option to learn for a cost and had to exclusively rely on standard Kalman filtering from the history of the data, as can be seen from the return simulations on the bottom panel of Figure 4 and the summary statistics in the second panel of Table 1. Indeed, with no costly learning, we do not find more than 1 or 2 instances of large price moves in 80 years of simulated daily data; the detected jumps represent pure-chance large random draws in the simulation. Consistent with the lack of large move in returns, the unconditional distribution of market returns does not possess heavy tails, as the kurtosis of return distribution of 3 is equal to that of Normal distribution.

Our constant volatility model can deliver the key result that the equilibrium asset prices can display infrequent large movements, while there are no corresponding jumps in the macroeconomic inputs (endowment growth). These asset-price jumps arise endogenously due to the optimal actions of investors pay a cost and learn the true expected growth state. However, since income shocks are homoscedastic, the agents exercise a costly learning option at constant time intervals, and steady-state volatilities of macroeconomic and financial variables are constant, which cannot account for the predictability of asset-price jumps in the data. We can address these issues by opening up stochastic volatility channel, which we discuss in the next section.

5.3 Time-Varying Volatility Model

The bottom panel of Table 1 depicts summary statistics for model-implied return distribution in the time-varying volatility specification. When agents have an option to learn for a cost, the model generates the mean market return of 6.4% and the volatility of returns of 15.5%. The model-implied risk-free rate is 1%. Hence, as in the constant volatility case, the model accounts for the usual equity premium and risk-free rate puzzles. Notably, while the model specification without costly learning

[11] Although costly learning occurs at constant time intervals, the years with flagged jumps do not always occur at regular intervals, as can be seen on Figure 4. Indeed, the jump-detection statistics is designed to pick out only large jumps, hence the significance level of 1%, so that some of the smaller price adjustments remain undetected.

can generate similar level of the equity premium, an option to learn for a cost leads to a considerably higher variation in the conditional equity premium in the time-series. For the consumption asset, the annualized volatility of the equity premium is 0.44% with costly learning versus 0.16% without costly learning (the total volatility of return on consumption asset is about 2%). For the dividend asset, the variation in annualized equity premium is 2.4% versus 0.32% without costly learning. This underscores the economic importance of the costly learning channel for the risk premium fluctuations.

Time-variation in income volatility implies that the optimal costly learning is stochastic and no longer occurs at constant time intervals. In particular, an important prediction of the model, stated in Result 4, is that the frequency of costly learning and consequently, asset-price jumps, increases at times of heightened income volatility. For a graphical illustration of these model implications, we show a typical simulation of income growth, income volatility and the variance of filtering error on Figure 5, and the equilibrium returns with and without costly learning option on Figure 6. The income growth process is conditionally Gaussian and hence does not exhibit large moves. Further, unlike the constant volatility model, the variance of the filtering error now fluctuates over time, and one can observe the occasional sharp reductions of the filtering uncertainty to zero at times when agents choose to learn the true state for a cost. These times of costly learning correspond to the periods of high income volatility and high variance of the filtering error. We highlight the dependence of the costly learning rule on the volatility states on Figure 7, which depicts the expected number of periods till the next costly learning given current filtering variance for high, medium and low values of aggregate volatility. Consistent with our earlier discussion, investors choose to learn for a cost if the variance of the filtering error grows too high in the economy, and the frequency of costly learning increases at times of heightened income volatility.

The actions of investors to learn about the underlying state can lead to large adjustments in daily asset prices, detected as jumps by annual jump-detection statistics, as shown in return simulation on Figure 6. To illustrate the response of the asset valuations to the revisions of the expected growth, we show a scatter plot of the change in price-dividend ratio Δv_t versus the revision in expectations $x_t - \hat{x}_t(0)$ on Figure 8. When the agents do not exercise the costly learning option, the unobserved gap between the true and filtered state has no information about asset prices, hence, the horizontal line in the middle of the graph. On the other hand, when agents pay a cost and observe the true state, the price-dividend ratio changes to reflect the adjustment in the agent's expectations about future growth. Positive revisions lead to upward moves in the asset prices, while lower than expected growth implies large negative jumps in returns.

Relative to constant volatility case, the detected jump-years are more frequent, averaging once every 3.4 years, and contribute more to the total variation in returns, 12% versus 7% in a constant-volatility case and in the data (see Table 1). These large moves in returns cannot be obtained in the economy without costly learning, as can be visually seen on the time-series plot of returns on Figure 6. Without costly learning, the average frequency of detected jump-years is less than 2 in 80 years, and the detected "jumps" are merely pure-chance large random draws. The comparison of the higher-order moments of model-implied return distribution is revealing: without an option to learn, the kurtosis of market returns is 3, and it reaches 36 when the agent can learn the true state for a cost.

Naturally, the frequency of detected jump-years depends on the significance of the jump detection test, which we set to 1%.. For robustness, on Figure 9 we show the jump-year frequency for a range of significance levels from 0.5% to 10%. As the significance level increases, the null of no jumps is rejected more often, so that the frequency of detected jump-years increases. As the Figure shows, the model can match very well the evidence on the average frequency of jump-years in the data, as the model-implied jump-year frequency is nearly on top of the empirical one and is well within the $5\% - 95\%$ confidence band.

5.4 Predictability of Jumps

An important prediction of our model is that asset-price jumps are predictable by the persistent variance measures, such as endowment and market volatility. To put our results in perspective, note that if the jump arrival intensity is constant, as is sometimes assumed in reduced-form asset-pricing models, the number of periods between successive jumps follows an exponential distribution. On Figure 10 we plot the unconditional distribution of the number of periods between the detected jump-years from the long simulation of the time-varying volatility model, alongside with the an exponential fit to this distribution. The mean of the fitted exponential distribution is 3.6 years, which agrees with the estimate of the jump-year frequency reported in Table 1. While the exponential distribution generally fits the distribution of jump duration, there is evidence for clustering of jumps – the unconditional distribution has heavier left tail than an exponential, so a jump-year is likely to follow another.

The persistence of asset-price jumps is a natural outcome of the model result that the frequency of learning and consequently, the likelihood of price jumps, is increasing with aggregate volatility. As discussed before in Section 2, the predictability of return jumps by the aggregate volatility is an important feature of the data, and our model can capture this effect. Furthermore, as the aggregate volatility also drives the variation in equilibrium market returns, our model can provide an economic explanation for the predictability of large asset-price moves by the variance of returns in the data. Finally, as in the data, the level of endowment growth does not predict future return jumps, as the optimal learning choice depends only on the income volatility and variance of filtering error. This highlights an important aspect of the model and the data that the second moments are critical to forecast future jumps, while the movements in the level are not informative about future jumps in returns.

The model can quantitatively reproduce the key features of predictability of return jumps by consumption and market variance, and absence of predictability of future jumps by the level of consumption growth. In the lower panels of Table 2 we show model-implied population values for the lead-lag correlations of jump indicator with endowment growth and conditional variance of endowment growth and returns at an annual frequency, constructed in the same way as the empirical counterparts in the data. As shown in Table, return jump indicator and the consumption growth rate have zero correlation, and return jumps and macroeconomic volatility have positive correlations of about the same magnitude as in the data. Most of the model correlations are within one standard deviation from their estimates in the data. For robustness, we checked the results using monthly frequency, and obtained similar results. In sharp contrast, in the model without costly learning all correlations are zeros (shown in lower panel of Table). That is, in the absence of endogenous jumps due to costly learning, the model is unable to capture the correlation between return volatility and the jump indicator, as well as the correlation between macro-volatility and the jump indicator. That is, it misses on key economically significant jump dimensions of the data. This underscores the importance of the costly learning channel, developed in the paper, to explain asset-price jumps.

In Table 4 we show the probit regression results for the predictability of jump-years by consumption variance and consumption growth. Using the consumption variance, we obtain the R^2 of 6% and 5% in the data and in the model, respectively. The R^2 drops to less than 1% in the data and 2% in the model when we use the lags of consumption growth rate to predict future jumps. Without the costly learning option, the R^2s are zero. Hence, our costly learning model can capture the predictability evidence of jumps in the data.In the next section we show that the jump predictability implications from the model are also consistent with the evidence from the parametric asset-pricing models.

5.5 Parametric Jump Model

The predictability of large moves in returns that our model is able to capture is consistent with the evidence from parametric models for asset prices, which feature stochastic volatility and jumps in returns whose arrival intensity is increasing in market variance; see examples in Bates (2000), Pan (2002), Eraker (2004) and Singleton (2006). To further compare our model implications to the results from the parametric studies of return dynamics, we fit a discrete-time GARCH-jump specification for returns, which feature autoregressive stochastic volatility and time-varying arrival intensity of jumps in returns.[12] Specifically, the return dynamics is given by

$$r_t = \mu_r + a_{1,t} + a_{2,t}. \tag{5.1}$$

The first component $a_{1,t}$ represents a smooth Gaussian component of returns, whose conditional volatility is time-varying and follows GARCH(1,1) process:

$$a_{1,t} = \sigma_{r,t-1}\tilde{a}_t, \quad \tilde{a}_t \sim N(0,1), \tag{5.2}$$
$$\sigma^2_{r,t} = \sigma^2_v + \beta_v \sigma^2_{r,t-1} + \alpha_v (r_t - \mu_r)^2. \tag{5.3}$$

The second shock $a_{2,t}$ is driven by Poisson jumps:

$$a_{2,t} = \sum_{k=1}^{n_t} \xi_{t,k} - \mu_j \lambda_{t-1}. \tag{5.4}$$

The jump size distribution is assumed to be Normal:

$$\xi_{t,k} \sim N(\mu_j, \sigma^2_j), \tag{5.5}$$

and the arrival of number of jumps $n_t = 0, 1, 2, \ldots$ is described by a conditional Poisson distribution with intensity λ_t, so that

$$Pr_{t-1}(n_t = j) = \frac{exp(-\lambda_{t-1})\lambda^j_{t-1}}{j!}. \tag{5.6}$$

As we are interested in the predictability of jumps by market variance, we follow the literature and model the jump intensity to be linear in the variance of returns,

$$\lambda_t = \lambda_0 + \lambda_l \sigma^2_{r,t}. \tag{5.7}$$

The above specification of return dynamics can be readily estimated by MLE using the sample and simulated data. In particular, we use monthly returns from 1926 to 2008 in the data, and perform a Monte-Carlo study with 100 simulations of 85 years of monthly returns from a time-varying volatility model.[13] The estimation results in the data and the model are shown in Table 7. As can be seen in Table 7, the model matches quite well the dynamics of the time-varying volatility of smooth component of the returns, as the volatility persistence parameters α_v and β_v are

[12] Similar specification is considered in Bates and Craine (1999). See Maheu and McCurdy (2004) for extensions and estimation details.

[13] We chose to focus on monthly rather than daily data as estimation using monthly data is more stable and shows less evidence of model mis-specification.

quite close in the data and in the model. The model can also capture the key findings regarding the frequency and predictability of jumps. The estimated probability of jumps loads positively and significantly on the market variance in the data, and the slope coefficient in the model matches the data estimate very well. The estimated mean jump size is -7% in the data, and the estimate is not statistically significant. In the model, as investors can learn that the true state is bigger or smaller than their estimate, the jumps are symmetric, so that the mean jump size is zero. The standard deviations of the jump distribution in the data and model are very close and equal to 6% and 7%, respectively.

Thus, the model can account for the key features of the conditional distribution of returns in the data, so it can serve as an economic basis for realistic reduced-form models of asset prices which incorporate time-varying volatility and jump components.

6 Conclusion

We present a general equilibrium model which features smooth Gaussian dynamics of income and dividends and large infrequent movements in asset prices (jumps). The large moves in asset prices are triggered by the optimal actions of investors to learn the unobserved expected growth. We show that the optimal decision to learn the true state is stochastic and depends on the time-varying volatility of income growth and the variance of the filtering error, as well as the preference parameters. The revisions in the expected income due to costly learning lead to large moves in asset valuations which look like jumps. These large price moves cannot be obtained in the economy without costly learning of the true state, or in the economy with standard expected utility.

A prominent feature of the model is that the frequency of costly learning, and consequently, the likelihood of asset price jumps, increases in the income volatility, so that returns jumps are more frequent at times of high aggregate volatility. We show that predictability of returns jumps by consumption variance is an important and novel aspect of the data. Furthermore, the model can provide an economic explanation for the predictability of large asset-price moves by the variance of returns, and lack of return jump predictability by the levels of consumption growth in the data. This highlights an important aspect of the model and the data that the second moments are critical to forecast future jumps, while the movements in the level are not informative about future jumps in returns.

Using calibrations, we find that the model can quantitatively reproduce the key features of predictability of return jumps by consumption and market variance, and absence of predictability of future jumps by the level of consumption. In addition, the model can account for the frequency and magnitude of price jumps in the data, fat-tail distribution of market returns, equity premium, and other asset-pricing features. We argue that our structural model can serve as an economic basis for realistic reduced-form models of asset prices which incorporate time-varying volatility and jump components.

A Model Solution

A.1 Social Planner's Problem

The learning decision of the social planner maximizes the ex-ante utility of the agent:

$$s_t^* = arg \max_s \{\mathbf{J}_t(U_t(s))\}, \tag{A.1}$$

subject to the resource constraint (3.5):

$$Y_t = C_t(s_t) + s_t \xi_t, \tag{A.2}$$

where the learning cost ξ_t is proportional to the aggregate income Y_t :

$$\xi_t = \chi Y_t, \tag{A.3}$$

for $0 \leq \chi < 1$. From the resource constraint, it immediately follows that

$$C_t(s_t) = Y_t(1 - \chi s_t). \tag{A.4}$$

Therefore, when the planner does not learn about the true state ($s_t = 0$), the agent's consumption is equal to the aggregate income. On the other hand, when the planer learns about the true state, ($s_t = 1$), part of the endowment is sacrificed to cover the learning cost.

Conjecture that the life-time utility functions are proportional to income:

$$U_t(s_t) = \phi_t(s_t) Y_t, \tag{A.5}$$

for $s_t \in \{0, 1\}$. The optimal utility of the agent then is given by the learning choice specific counterpart evaluated at the optimal indicator $s_t^* : U_t = U_t(s_t^*)$. The optimal utility next period takes into account the optimal learning choice tomorrow and can be written as $U_{t+1} = \phi_{t+1} Y_{t+1}$, where to simplify the notations, we denote $\phi_{t+1} \equiv \phi_{t+1}(s_{t+1}^*)$.

Substitute the conjecture for $U_t(s_t)$ and U_{t+1}, and the consumption rule (A.4) into the definition of the life-time utility of the agent in (3.3) to obtain the following recursive formula for the utility per income $\phi_t(s_t)$:

$$\phi_t(s_t) = (1 - s_t \chi)^{1-\beta} \left(E_t^{s_t} \left[\phi_{t+1} \frac{Y_{t+1}}{Y_t} \right]^{1-\gamma} \right)^{\frac{\beta}{1-\gamma}}. \tag{A.6}$$

As aggregate income Y_t is known in the beginning-of-the period, it can be factored out from the optimal condition for learning (A.1). We can then rewrite it in the following way:

$$\begin{aligned} s_t^* &= 1 \quad \text{if } \mathbf{J}_t(\phi_t(1)) > \mathbf{J}_t(\phi_t(0)) \\ &= 0 \quad \text{if } \mathbf{J}_t(\phi_t(1)) \leq \mathbf{J}_t(\phi_t(0)). \end{aligned} \tag{A.7}$$

A.2 Representative Agent Problem

For completeness, we also show the solution to the representative agent problem.

Denote W_t the agent's wealth in period t. Denote ω_t the fraction of agent's wealth invested into each asset (i.e., the portfolio weights), so that $\omega_t' 1 = 1$. Then, for a vector of asset returns R_{t+1}, we can write down the return on the aggregate wealth,

$$R_{c,t+1} = \omega_t' R_{t+1}. \tag{A.8}$$

Hence, we can write down the budget constraint of the agent in the following way:

$$W_{t+1} = (W_t - C_t - s_t \xi_t) \, R_{c,t+1}. \tag{A.9}$$

For convenience, we make the learning costs ξ_t be a fixed proportion of the consumption level of the agent,

$$\xi_t = \frac{\chi}{1-\chi} C_t, \tag{A.10}$$

It is easy to see that with this parametrization, in equilibrium the learning costs are going to be proportional to the aggregate income of the agent with a coefficient of proportionality χ, as stated in (3.6).

The budget constraint can then be rewritten in the following way:

$$W_{t+1} = \left(W_t - \frac{C_t}{1 - \chi s_t} \right) R_{c,t+1}. \tag{A.11}$$

The optimization problem of the agent is given by:

$$U_t(s_t) = \max_{C_t, \ w_t} \left\{ \mathbf{J}_t(U_t(s)) \right\} \tag{A.12}$$

subject to the budget constraint above.

Taking a first-order condition with respect to consumption, we obtain, after some algebra, the optimal consumption rule

$$\frac{C_t}{W_t} = (1-\beta)(1 - s_t \chi). \tag{A.13}$$

Now let us take a first-order condition with respect to a portfolio weight i, subject to the restriction $\omega_t' 1 = 1$:

$$E_t^{s_t}\left(\left(U_{t+1}/W_{t+1}\right)^{1-\gamma} R_{c,t+1}^{-\gamma} R_{i,t+1}\right) = const, \quad \forall i. \tag{A.14}$$

Multiplying each of these conditions by a corresponding asset weight and summing up, we obtain that

$$E_t^{s_t}\left(\left(U_{t+1}/W_{t+1}\right)^{1-\gamma} R_{c,t+1}^{-\gamma} R_{i,t+1}\right) = E_t^{s_t}\left(\left(U_{t+1}/W_{t+1}\right)^{1-\gamma} R_{c,t+1}^{1-\gamma}\right),$$

from where we can read the discount factor,

$$M_{t+1} = \frac{\left(U_{t+1}/W_{t+1}\right)^{1-\gamma} R_{c,t+1}^{-\gamma}}{E_t^{s_t}\left(\left(U_{t+1}/W_{t+1}\right)^{1-\gamma} R_{c,t+1}^{1-\gamma}\right)}. \tag{A.15}$$

Let us impose equilibrium restrictions that income is equal to consumption plus learning costs, $Y_t = C_t + s_t \xi_t$. For our parametrization of the learning cost in (A.10), this implies that in equilibrium, the learning costs are proportional to the aggregate income, $\xi_t = \chi Y_t$.

Note that while consumption per wealth ratio is not constant (see (A.13)), the total income per wealth ratio is constant:

$$\frac{Y_t}{W_t} = 1 - \beta. \tag{A.16}$$

Using the budget constraint in (A.11), we obtain that the equilibrium aggregate wealth return is proportional to the income growth:

$$R_{c,t+1} = \frac{1}{\beta}\frac{Y_{t+1}}{Y_t}. \tag{A.17}$$

From here, we obtain that in equilibrium,

$$\begin{aligned}\left(U_{t+1}/W_{t+1}\right) R_{c,t+1} &= \left(\frac{U_{t+1}}{W_{t+1}}\right)\left(\frac{1}{\beta}\frac{Y_{t+1}}{Y_t}\right) \\ &= \frac{1-\beta}{\beta}\frac{U_{t+1}}{Y_t}.\end{aligned}$$

Substitute this into the expression for the discount factor (A.15) to obtain

$$M_{t+1} = \beta\left(\frac{Y_{t+1}}{Y_t}\right)^{-1} \frac{U_{t+1}^{1-\gamma}}{E_t^{s_t} U_{t+1}^{1-\gamma}}. \tag{A.18}$$

This is the equilibrium discount factor for a given model choice of the agent s_t.

We can also plug in the optimal consumption policy into the utility equation (A.12) to verify the solution to the utility per wealth in (A.6) obtained using the social planner problem.

A.3 Timing of Resolution of Uncertainty

The key aspect of our model is that the agent has a preference for a timing of the resolution of uncertainty. In Section 4 we showed that with standard utility preferences, the agents have no incentive to learn for a cost, even if the cost is zero. Let us now formally prove that when investors have preference for early resolution of uncertainty ($\gamma > 1$), they have incentives to learn the new information.

Using the recursive solution for the utility per income ratio in (A.6), the solution to the optimal choice indicator above can be expanded in the following way:

$$s_t^* = 1\left[(1-\chi)^{1-\beta}\left(E_t^0\left[E_t^1\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma}\right]^{\beta}\right)^{\frac{1}{1-\gamma}} > \left[E_t^0\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma}\right]^{\frac{\beta}{1-\gamma}}\right]. \tag{A.19}$$

Consider a case when the learning cost is zero, $\chi = 0$. As $\beta < 1$, using Jensen's inequality argument it follows that $E(U^{\beta}) < (E(U))^{\beta}$ for any random variable U with positive support. Now apply this inequality for $U = E_t^1\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma}$ to obtain

$$E_t^0\left[E_t^1\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma}\right]^{\beta} < \left[E_t^0\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma}\right]^{\beta}.$$

To plug this inequality into optimal policy equation (A.19), we need to take both sides of it to the $1/(1-\gamma)$ power. With preference for early resolution of uncertainty, $\gamma > 1$ and the exponent $1/(1-\gamma)$ is negative, so the inequality reverses. Hence, with preference for early resolution of uncertainty, agents can always increase their life-time utility by learning additional information at zero cost. On the other hand, when $\gamma \le 1$, agents have no preference for early resolution of uncertainty, and hence they have no incentive to acquire additional information even if it is costless. Hence, preference for early resolution of uncertainty plays a key role in investors decision to pay a cost and learn the new information.

A.4 Utility and Learning Choice

As the volatility and consumption shocks are uncorrelated, we can separate the expected growth and volatility components in the equilibrium utility per income ratio, which simplifies the solution to the fixed-point recursion in (3.8). In this section we consider a general case with time-varying volatility, while in Appendix A.6 we show that the solution can be simplified even further when the volatility is constant.

Conjecture that for each choice indicator s_t and corresponding states $\hat{x}_t(s_t), \omega_t^2(0)$ and σ_t^2 today, the life-time utility per income ratio satisfies,

$$\phi(s_t, \hat{x}_t(s_t), \omega_t^2(0), \sigma_t^2) = e^{B\hat{x}_t(s_t)+f(s_t,\sigma_t^2,\omega_t^2(0))}, \tag{A.20}$$

for some utility loading B and volatility function $f(s_t, \sigma_t^2, \omega_t^2(0))$. Note that the variance of the filtering error used in the value function is based on the beginning of period information; the actual value $\omega_t^2(s_t)$ depends deterministically on the beginning-of-period estimate $\omega_t^2(0)$ and learning choice s_t, see equation (3.21).

Conjecture that the optimal learning choice tomorrow s_{t+1}^* depends only on the income volatility and beginning-of-period variance of the filtering error, and not on the expected income and dividend factors, i.e. $s_{t+1}^* = s^*(\sigma_{t+1}^2, \omega_{t+1}^2(0))$. Consider the equilibrium life-time utility from next period onward:

$$\begin{aligned} \phi_{t+1} &= \phi(s_{t+1}^*, \hat{x}_{t+1}(s_{t+1}^*), \omega_{t+1}^2(0), \sigma_{t+1}^2) \\ &= e^{B\hat{x}_{t+1}(s_{t+1}^*)+f_{t+1}}, \end{aligned} \tag{A.21}$$

where for notational simplicity, we define $f_{t+1} = f(s_{t+1}^*, \sigma_{t+1}^2, \omega_{t+1}^2(0))$. Now, using (3.20),

$$\log\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right) = B\left(\hat{x}_{t+1}(0) + s_{t+1}^*(x_{t+1} - \hat{x}_{t+1}(0))\right) + f_{t+1} + \Delta y_{t+1}. \tag{A.22}$$

Consider a recursive equation for the optimal utility per income ratio (A.6) for a given choice indicator s_t today. To evaluate $E_t^{s_t}\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma}$, we use the law of iterated expectations where we first condition on $\mathcal{I}_{t+1}$. Then, $\Delta y_{t+1}, \sigma_{t+1}^2$ and therefore $\hat{x}_{t+1}(0), s_{t+1}^*$ and f_{t+1} are known, while the only random component is the true state x_{t+1}. Due to the Kalman filter procedure,

$$x_{t+1}|\mathcal{I}_{t+1} \sim N(\hat{x}_{t+1}(0), \omega_{t+1}^2(0)), \tag{A.23}$$

where $\hat{x}_{t+1}(0)$ and $\omega_{t+1}^2(0)$ satisfy (3.20) and (3.21). Therefore the right-hand side expectation in the utility recursion (3.8) is equal to,

$$\begin{aligned} E_t^{s_t}\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma} &= E_t^{s_t} e^{(1-\gamma)\left[B\hat{x}_{t+1}(0)+f_{t+1}+\Delta y_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega_{t+1}^2(0)s_{t+1}^*\right]} \\ = e^{(1-\gamma)(\mu+(B\rho+1)\hat{x}_t(s_t))} E_t^{s_t} e^{(1-\gamma)\left[(BK_t(s_t)+1)u_{t+1}(s_t)+f_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega_{t+1}^2(0)s_{t+1}^*\right]}&. \end{aligned} \tag{A.24}$$

Now by conjecture, s^*_{t+1} and thus f_{t+1} and $\omega_{t+1}(s^*_{t+1})$ are driven by income volatility shocks, which are independent from income innovations and therefore, from the filtered shock $u_{t+1}(s_t)$. Thus,

$$\begin{aligned} E_t^{s_t}\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma} &= e^{(1-\gamma)\left(\mu+(B\rho+1)\hat{x}_t(s_t)+\frac{1}{2}(1-\gamma)(BK_t(s_t)+1)^2(\omega_t^2(s_t)+\sigma_t^2)\right)} \\ &\times E_t^{s_t} e^{(1-\gamma)\left[f_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega_{t+1}^2(0)s^*_{t+1}\right]} \end{aligned} \tag{A.25}$$

Therefore, using the equilibrium utility recursion (A.6) and the conjectured solution for the life-time utility of the agent (A.20) and matching the coefficients, we obtain that loading on expected growth is equal to

$$B = \frac{\beta}{1-\beta\rho}, \tag{A.26}$$

while the volatility function satisfies

$$\begin{aligned} &f(s_t,\sigma_t^2,\omega_t^2(0)) = (1-\beta)\ln(1-s_t\chi) + \beta\mu \\ &+ \beta\frac{1}{2}(1-\gamma)(BK_t(s_t)+1)^2(\omega_t^2(s_t)+\sigma_t^2) + \frac{\beta}{1-\gamma}\ln E_t^{s_t} e^{(1-\gamma)\left[f_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega_{t+1}^2(0)s^*_{t+1}\right]}. \end{aligned} \tag{A.27}$$

Solution to B and f verifies the conjecture for the life-time utility of the agent.

Now, given the utility equation (A.20) and the dynamics of the factors, we can rewrite the optimal condition for a learning choice (3.4). Notably, the expected growth component drops out, so that the optimal choice indicator depends only on the learning and aggregate variance:

$$s^*_t = 1\left[\frac{1}{2}(1-\gamma)B^2\omega_t^2(0) + f_t(1,\sigma_t^2,\omega_t^2(0)) > f_t(0,\sigma_t^2,\omega_t^2(0))\right]. \tag{A.28}$$

Using the optimal condition for s^*_{t+1} tomorrow to rewrite the recursive equation of the volatility function (A.27) in the following way:

$$\begin{aligned} &f(s_t,\sigma_t^2,\omega_t^2(0)) = (1-\beta)\ln(1-s_t\chi) + \beta\mu + \beta\frac{1}{2}(1-\gamma)(BK_t(s_t)+1)^2(\omega_t^2(s_t)+\sigma_t^2) \\ &+ \frac{\beta}{1-\gamma}\ln E_t^{s_t} e^{(1-\gamma)\max\left[\frac{1}{2}(1-\gamma)B^2\omega_{t+1}^2(0)+f_{t+1}(1,\sigma_{t+1}^2,\omega_{t+1}^2(0)),f_{t+1}(0,\sigma_{t+1}^2,\omega_{t+1}^2(0))\right]}. \end{aligned} \tag{A.29}$$

That is, the volatility function f can be obtained as fixed-point solution to the equation above, given the evolution of the variance of the filtering error in (3.18) and (3.21).

Using the solution to the equilibrium discount factor in (A.18), we can express the equilibrium discount factor in terms of the underlying variables in the economy:

$$
\begin{aligned}
&m_{t+1}(s_t) = \log\beta - \mu - \hat{x}_t(s_t) - \frac{1}{2}(1-\gamma)^2(BK_t(s_t)+1)^2(\omega_t^2(s_t)+\sigma_t^2) \\
&- \ln E_t^{s_t} e^{(1-\gamma)\left[f_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega_{t+1}^2(0)s_{t+1}^*\right]} \\
&- (1+(\gamma-1)(1+BK_t(s_t)))\,u_{t+1}(s_t) - (\gamma-1)Bs_{t+1}^*(x_{t+1}-\hat{x}_{t+1}(0)) - (\gamma-1)f_{t+1}.
\end{aligned} \tag{A.30}
$$

Using the equilibrium solution to the discount factor, we obtain that the risk-free rate satisfies,

$$
r_{ft} = -\log\beta + \mu + \hat{x}(s_t) - \frac{1}{2}(2\gamma-1)(BK_t+1)^2(\omega_t^2+\sigma_t^2) \tag{A.31}
$$

A.5 Dividend Asset

We follow a standard approach of Bansal and Yaron (2004) to solve for the equilibrium price-dividend ratio in the economy.

Conjecture that the equilibrium price-dividend ratio satisfies

$$
v_t(s_t) = H\hat{x}_t(s_t) + h(s_t, \sigma_t^2, \omega_t^2(0)), \tag{A.32}
$$

We log-linearize the market return, which, using the dividend specification in (4.9), the conjectured solution for the price-dividend ratio and the dynamics of the state, can be expressed in the following way:

$$
\begin{aligned}
r_{d,t+1} &= \kappa_0 + \varphi_d\mu + (H(\kappa_1\rho-1)+\varphi_d)\hat{x}_t(s_t) - h(s_t,\sigma_t^2,\omega_t^2(0)) \\
&\quad + (\kappa_1 HK_t(s_t)+\varphi_d)u_{t+1}(s_t) + \kappa_1 h(s_{t+1}^*, \sigma_{t+1}^2, \omega_{t+1}^2(0)) + \kappa_1 H s_{t+1}^*(x_{t+1}-\hat{x}_{t+1}(0)).
\end{aligned} \tag{A.33}
$$

for endogenous log-linearization coefficients κ_0 and κ_1.

Using Euler conditions and the equilibrium solution for discount factor for the log-linearized dividend return, we obtain that the loading H is given by

$$
H = \frac{\varphi_d - 1}{1-\kappa_1\rho}, \tag{A.34}
$$

while the price-dividend volatility component satisfies a recursive equation

$$
\begin{aligned}
&h_t(s_t, \sigma_t^2, \omega_t^2(0)) = \ln\beta + \kappa_0 \\
&+ \frac{1}{2}(\varphi_d - 1 + \kappa_1 H K_t(s_t))(\varphi_d - 1 + \kappa_1 H K_t(s_t) - 2(\gamma-1)(1 + BK_t(s_t)))(\sigma_t^2 + \omega_t^2(s_t)) \\
&+ \ln E_t^{s_t} e^{\kappa_1 h_{t+1} + \frac{1}{2}(\kappa_1 H - (\gamma-1)B)^2 s^*_{t+1}\omega^2_{t+1}(0) - (\gamma-1) f_{t+1}} - \ln E_t^{s_t} e^{(1-\gamma)\left(f_{t+1} + \frac{1}{2}(1-\gamma)B^2 s^*_{t+1} w^2_{t+1}(0)\right)}.
\end{aligned}
\tag{A.35}
$$

To solve for the approximating constants κ_0 and κ_1, we use the numerical procedure discussed in Bansal et al. (2007), who develop a method to solve for the endogenous constants associated with each return and document that the numerical solution to the model is accurate.

A.6 Constant Volatility Case

When the income volatility is constant $\sigma_t = \sigma_0^2$, the variance of the filtering error becomes a deterministic function of time since the last learning about the true state. In this case, the optimal learning decision is purely time-dependent, so that the investors choose to learn about the underlying state if the last time they did so was N or more periods ago.

Assume we know the optimal N, and consider the time interval from 1 to N. In equilibrium, the agent starts filtering in period 1 and learns about the true state for a cost in period N, afterwards the solution repeats itself.

The equilibrium volatility functions are non-random functions of time, so to simplify the notations, denote them f_i :

$$
\begin{aligned}
f_i &= f(0, \sigma_0^2, \omega_i^2(0)), \quad 1 \le i < N, \\
f_N &= f(1, \sigma_0^2, \omega_N^2(0)).
\end{aligned}
$$

Now we can rewrite the recursions in (A.27) as a system of linear equations (to simplify the exposition, we consider the case $N > 2$):

$$
\begin{aligned}
f_1 - \beta f_2 &= \beta\mu + \frac{1}{2}\beta(1-\gamma)(BK_1(0)+1)^2(\omega_1^2(0) + \sigma_0^2), \\
&\dots \\
f_i - \beta f_{i+1} &= \beta\mu + \frac{1}{2}\beta(1-\gamma)(BK_i(0)+1)^2(\omega_i^2(0) + \sigma_0^2), \quad 2 \le i < N-1 \\
f_{N-1} - \beta f_N &= \beta\mu + \frac{1}{2}\beta(1-\gamma)\left((BK_{N-1}(0)+1)^2(\omega_{N-1}^2(0) + \sigma_0^2) + B^2\omega_N^2(0)\right), \\
f_N - \beta f_1 &= (1-\beta)\ln(1-\chi) + \beta\mu + \frac{1}{2}\beta(1-\gamma)(BK_N(1)+1)^2(\omega_N^2(1) + \sigma_0^2).
\end{aligned}
\tag{A.36}
$$

This system can be easily solved for equilibrium volatility functions $f_i, i = 1, 2, ...N$.

Now we need to make sure that the chosen N is indeed optimal, that is, the agent is not better off deviating from the conjectured learning rule. If investors were to learn about the state earlier than in the Nth period, their utility would be f_N. To preclude this deviation, we need to have that (see condition (A.28))

$$\frac{1}{2}(1-\gamma)B^2\omega_i^2(0) + f_N < f_i, \tag{A.37}$$

for $1 \leq i < N$.

On the other hand, consider a scenario when investors fail to learn about the true state at time N. By conjecture, the optimal behavior in period $N+1$ is to learn, therefore, from the expression (A.27), the utility that the investors would get by deviating is given by,

$$\tilde{f}_N = \beta\mu + \frac{1}{2}\beta(1-\gamma)\left((BK_N(0)+1)^2(\omega_N^2(0)+\sigma_0^2) + B^2\omega_{N+1}^2(0)\right) + \beta f_N.$$

Following the optimality condition for the choice indicator (A.28), we then need to have that

$$\frac{1}{2}(1-\gamma)B^2\omega_N^2(0) + f_N > \tilde{f}_N. \tag{A.38}$$

In practice, we loop from a low value of N until we satisfy both optimality conditions (A.37)-(A.38), solving a linear system (A.36) for the volatility functions f_i. In numerical calibrations we verify that the optimal N is always unique: when N is lower than optimum, we violate the last condition (A.38), so that the agent can increase the utility by estimating, rather than learning about the state for a cost; for N higher than optimum, (A.37) is not satisfied, and investors would want to learn sooner.

We follow the same approach to find the volatility functions in the price-dividend ratio. As hs are no longer random, we can rewrite their recursion in (A.35) much in the same way as (A.36). The solution for the price-dividend volatility components then follows directly as we already know the optimal choice indicator and utility functions f_i. To solve for the approximating constants κ_0 and κ_1, we use the numerical procedure discussed in Bansal et al. (2007).

Now let us prove formally that if investors have preference for early resolution of uncertainty, the frequency of costly learning increases when the agents get more risk-averse or learning becomes less costly. Indeed, when $\gamma > 1$, we can express the constraint (A.37) using the system of equations (A.36) as

$$0 > \frac{f_i - f_N}{1-\gamma} - \frac{1}{2}B^2\omega_i^2(0) = -\frac{(1-\beta)^2}{1-\beta^N}\ln(1-\chi)\frac{1}{1-\gamma} + q_i, \tag{A.39}$$

where the term q_i does not depend on the risk-aversion γ or learning cost χ parameters. Notably, the sign of the inequality critically depends on $\gamma > 1$, that is, preference for early resolution of uncertainty. When $\gamma < 1$, the inequality would reverse, and in this case, as we discussed in Section 4.2, it is never optimal to learn for a cost, and N is infinity.

It is easy to see that the right-hand side of the inequality (A.39) is unambiguously increasing in the risk-aversion, and decreasing in the learning cost parameter. Hence, if we start with an equilibrium solution to the model and increase the risk-aversion (decrease learning cost), the optimality constraint on number of period N (A.37) gets monotonically more restrictive. When the risk aversion rises high enough (learning cost drops low enough), the constraint gets violated and the agent optimally chooses to decrease the number of learning periods N. That is, the frequency of costly learning increases with the risk-aversion coefficient and decreases with the costly learning parameter.

Tables and Figures

Table 1: **Summary Statistics: Data and Model**

	Mean	Std Dev	Kurt	Jump-year Freq	Jump Contribution
Data:					
Return	7.36	17.09	21.03	3.32	7.48
Model:					
Constant Volatility:					
Return with costly learning	6.70	15.49	17.69	4.84	7.16
Return, no costly learning	6.95	15.55	3.01	41.42	1.95
Time-Varying Volatility:					
Return with costly learning	6.35	15.52	35.51	3.35	11.93
Return, no costly learning	6.22	13.94	3.17	44.78	1.99

Mean, standard deviation and kurtosis of returns, and frequency and variance contribution of jumps. The first panel presents statistics in the data, while the second one – for the model specifications with constant and time-varying volatility, respectively. No costly learning refers to the case when the agent has no option to learn the true state for a cost. Jump-year frequency is the average frequency of years with jumps detected by the jump statistics, in years. Jump Contribution measures the average percent contribution of large price moves to the total return variance. Data are daily inflation-adjusted market returns from 1926 to 2008. Model statistics are based on the average across 100 simulations of 85 years of data. Jump-detection statistics are based on 1% significance level.

Table 2: **Jump Correlations: Data and Model**

	-2y	-1y	0y	1y	2y
Data:					
Growth Rate	-0.15	-0.08	-0.08	-0.07	0.10
	(0.16)	(0.14)	(0.10)	(0.12)	(0.15)
Macro vol	0.18	0.23	0.26	0.26	0.20
	(0.11)	(0.10)	(0.09)	(0.11)	(0.10)
Return Vol	0.10	0.11	0.12	0.03	0.01
	(0.14)	(0.15)	(0.15)	(0.13)	(0.13)
Costly Learning Model:					
Growth Rate	0.00	0.00	0.00	0.00	0.00
Macro Vol	0.07	0.09	0.16	0.13	0.12
Return Vol	0.04	0.07	0.11	0.09	0.06
Model Without Costly Learning:					
Growth Rate	0.00	0.00	0.00	0.00	0.00
Macro Vol	0.00	0.00	0.00	0.00	0.00
Return Vol	0.00	0.00	0.00	0.00	0.00

Correlation of return jump indicator with past and future economic growth rate, aggregate economic volatility and conditional variance of returns, at 1 and 2 year leads and lags. The data are annual observations of real consumption growth and returns from 1930 to 2008. Model statistics are population values at annual frequency.

Table 3: **Estimation of Consumption Volatility**

	Δc	pd	$spread$	R^2
Projection:				
$\widehat{\Delta c}$	0.351	0.006	-0.004	0.268
	(0.063)	(0.006)	(0.003)	
$\widehat{\sigma}^2 \times 10^4$	22.358	-0.571	5.046	0.232
	(33.241)	(1.057)	(2.112)	

GARCH Model:

	ρ_0	$\bar{\sigma}$	α_c	β_c	R^2
$\widehat{\Delta c}$	0.30	1.81e-05	0.72	0.18	0.10
	(0.11)	(1.5e-05)	(0.14)	(0.09)	

Estimation of the conditional consumption volatility. Top panel presents slope coefficients and R^2 in the projections of consumption growth and squared consumption residual on price-dividend ratio and junk bond spread. Bottom panel presents estimation results of the AR(1)-GARCH(1,1) specification $\Delta c_{t+1} = \mu_0 + \rho_0 \Delta c_t + \widehat{\sigma}_t \epsilon_{c,t+1}$, $\widehat{\sigma}^2_{t+1} = \bar{\sigma} + \alpha_c \widehat{\sigma}^2_t + \beta_c(\widehat{\sigma}_t \epsilon_{c,t+1})^2$. Annual observations of real consumption growth, price-dividend ratio and AAA-BAA junk bond spread from 1930 to 2008. Standard errors are in parentheses.

Table 4: **Jump Predictability: Data and Model**

	Data $R^2, \%$	Model $R^2, \%$
Consumption Variance	6.28	4.92
Consumption Growth	0.60	1.94

Predictability of jump-years by the consumption volatility and realized consumption growth. The table reports R^2 in probit regressions of jump-year indicator on the lags of the level or variance of consumption growth. Data is based on annual observations of consumption and returns from 1930 to 2008. Model output is based on 100 simulations of 85 years of daily data aggregated to annual horizon.

Table 5: **Configuration of Model Parameters**

Preferences and Learning	β	γ	χ		
	0.997*	10	8.5%		
Consumption and Dividend	μ	ρ	σ	φ_e	φ_d
	1.92*	0.38*	1.39*	0.15*	5
Volatility	σ_w	ν			
	1.37e-02*	0.77*			

The model is calibrated on daily frequency. Stars indicate annualized parameter values. On average, there are 22×12 trading days a year, so the annualized values are, $12 \times 22\mu, \rho^{12\times 22}, \sqrt{12 \times 22}\sigma, \sqrt{12 \times 22}\sigma_w, \nu^{12\times 22}, \sqrt{12 \times 22}\varphi_e, \beta^{12\times 22}$. Mean and volatility parameters are in percent.

Table 6: **Consumption Dynamics: Data and Model**

	Data		Model		
	Estimate	S.E.	Median	5%	95%
Mean	1.92	(0.29)	1.88	1.24	2.53
Vol	2.13	(0.59)	2.18	1.81	2.62
AR(1)	0.45	(0.11)	0.51	0.36	0.65
AR(2)	0.16	(0.14)	0.14	-0.06	0.36
AR(5)	-0.01	(0.09)	-0.01	-0.25	0.18

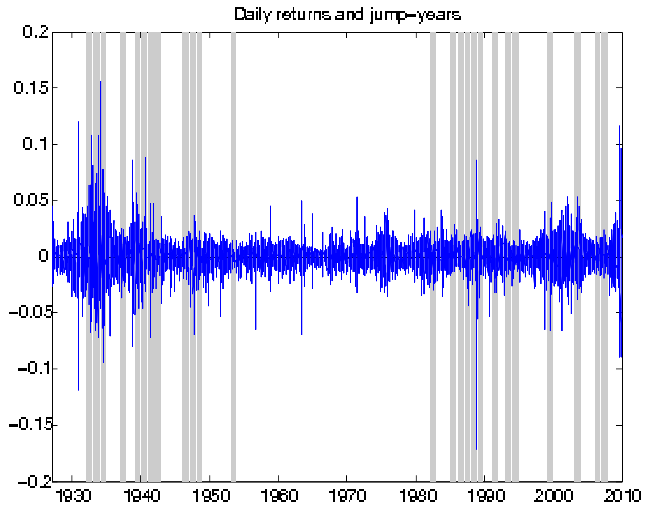
Calibration of consumption dynamics. Data is annual real consumption growth for 1930-2008. Model is based on 100 daily simulation of 85 years of consumption growth aggregated to annual horizon. Standard errors are Newey-West with 10 lags.

Table 7: **Estimation of GARCH-Jump Model**

	Data		Model	
	Estimate	S.E.	Mean	S.E.
σ_v	0.01	(0.01)	0.01	(0.01)
β_v	0.85	(0.03)	0.88	(0.13)
α_v	0.07	(0.02)	0.01	(0.01)
λ_0	-0.06	(0.03)	-0.03	(0.10)
λ_l	73.33	(32.46)	62.62	(64.19)
μ_j	-0.07	(0.04)	0.00	(0.02)
σ_j	0.06	(0.02)	0.07	(0.02)

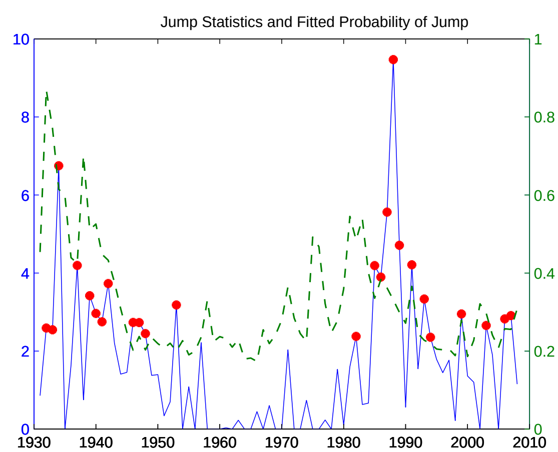
MLE estimation of parametric GARCH-jump model. Data is monthly real returns from 1926 to 2008. Model is based on 100 simulation of 85 years of daily returns aggregated to monthly horizon.

Figure 1: **Time-series of Returns**



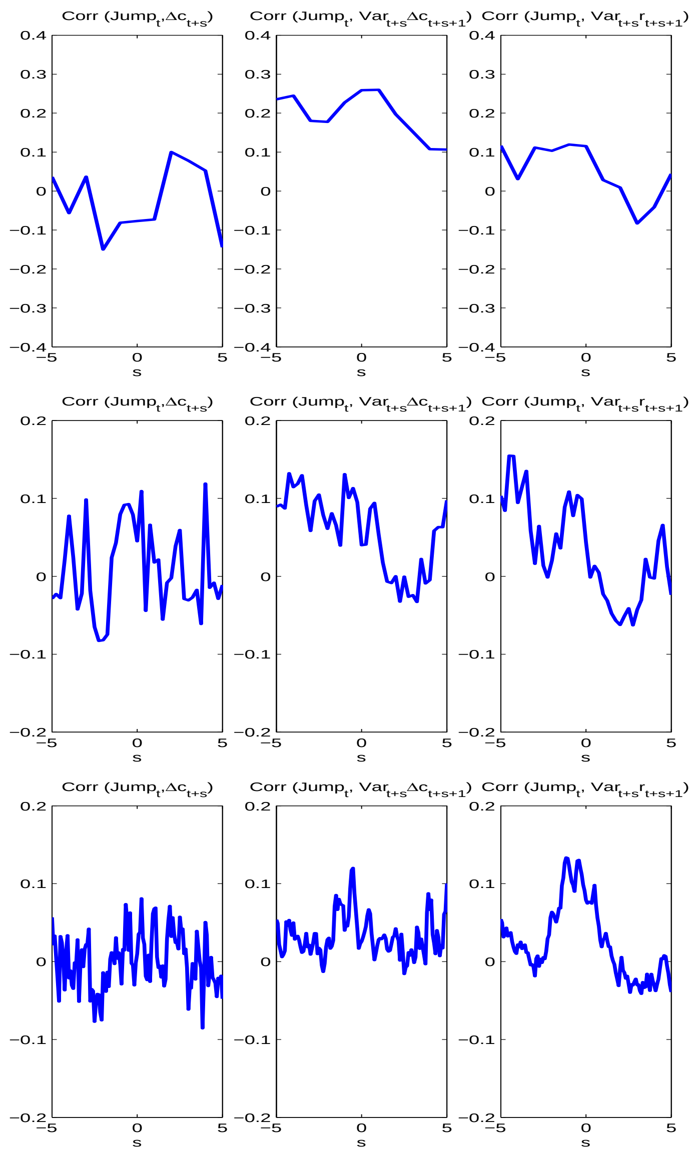
Daily observations on real market returns from 1926 to 2008. Grey regions correspond to periods with at least one significant large price move, at 1% significance level.

Figure 2: **Predicted Probability of Large Price Moves**



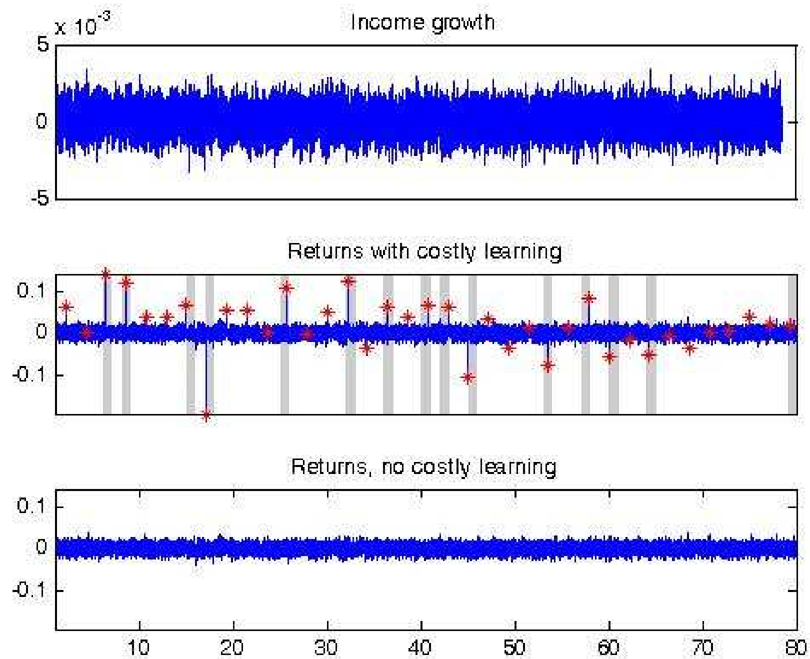
Annual jump statistics (solid line with values on the left Y-axis) and the predicted probability of large price moves (dashed line with values on the right Y-axis), based on the ex-ante consumption variance. Stars indicate years with at least one large price move.

Figure 3: **Jump Correlations in the Data**



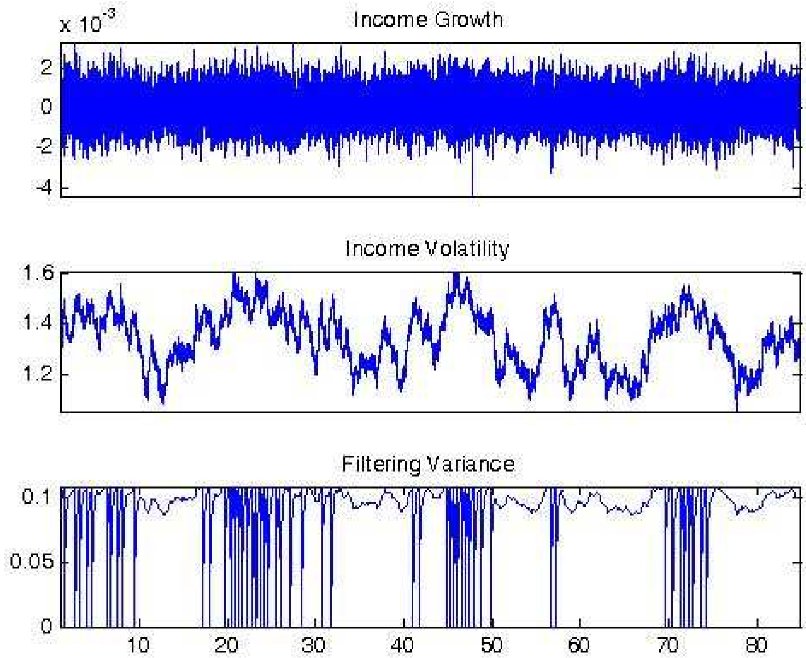
Correlation of return jump indicator with the level of economic growth rate (left panel), aggregate economic volatility (middle panel) and conditional variance of returns, at up to 5 year leads and lags. Top panel is based on annual observations of real consumption growth and returns from 1930 to 2008; middle and bottom panels are based on industrial production and return data from 1926 to 2008 at quarterly and monthly frequency, respectively.

Figure 4: **Income and Return Simulation in Constant Volatility Model**



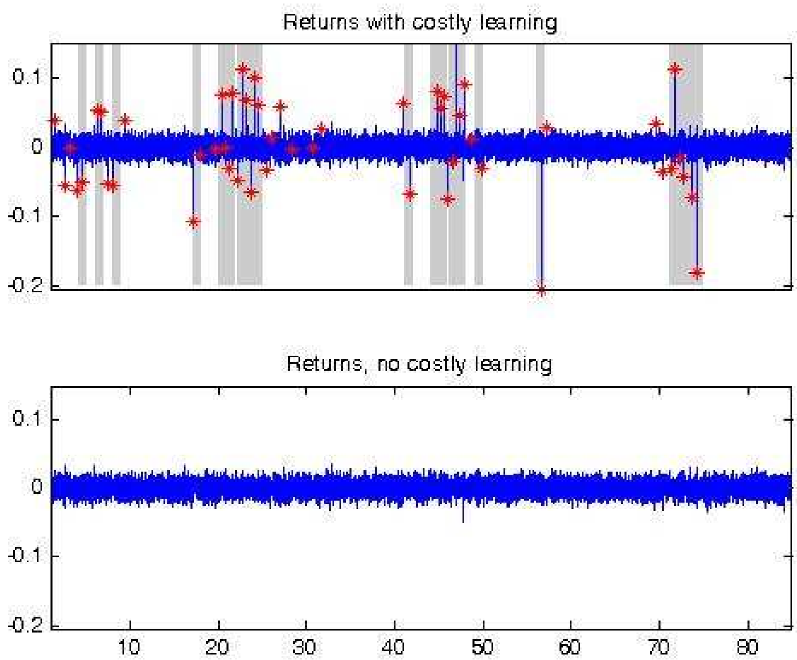
Simulation of the economy for 85 years in constant volatility model. Top panel depicts daily income growth. The next two panels show daily market returns in the models when the agent has an option to learn the true state for a cost and with no option to learn, respectively. Red stars indicate days of costly learning, while grey regions correspond to the years with at least one significant jump, detected by the jump statistics.

Figure 5: **Income Simulation in Time-Varying Volatility Model**



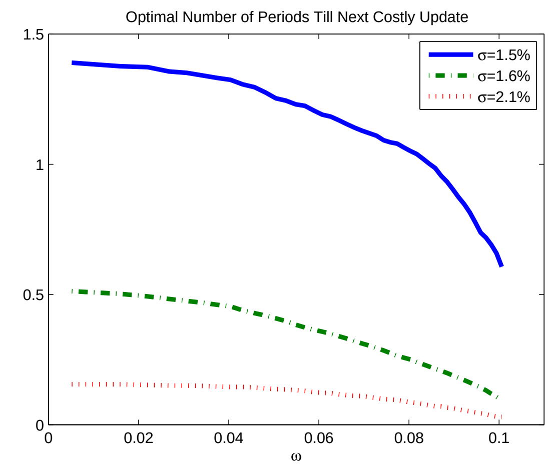
Simulation of the economy for 85 years in time-varying volatility model. Top panel depicts daily income growth. The next two panels show conditional volatility of income growth and the volatility of filtering error, annualized in percent.

Figure 6: **Return Simulation in Time-Varying Volatility Model**



Simulation of the economy for 85 years in time-varying volatility model. The two panels show daily market returns in the models when the agent has an option to learn the true state for a cost and with no option to learn, respectively. Red stars indicate days of learning, while grey regions correspond to the years with at least one significant jump, detected by the jump statistics.

Figure 7: **Costly Learning Frequency in Time-Varying Volatility Model**

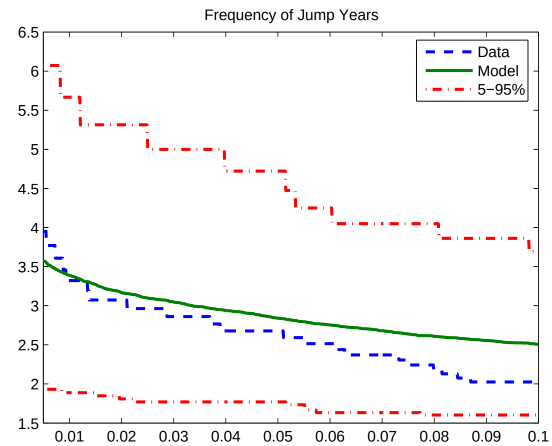


Average number of periods until next costly learning update, in years, for a given filtering variance and 3 levels of income volatility. Based on a long simulation of the time-varying volatility model. Volatilities are annualized, in percent.

Figure 8: **Change in PD Ratio Due to Revision in Expected Growth**

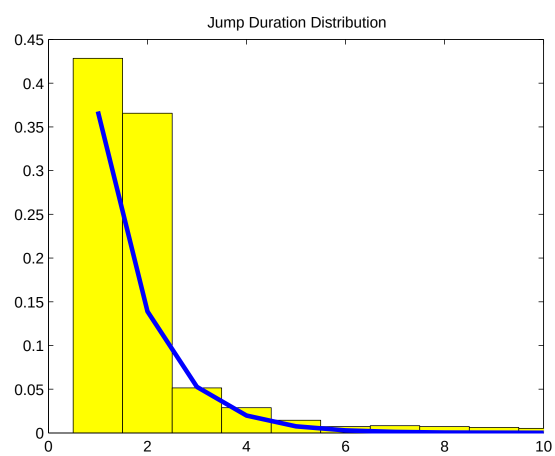
Scatter plot of the change in price-dividend ratio, Δv_t, versus the revision in expected growth $x_t - \hat{x}_t(0)$. Based on a long simulation of the time-varying volatility model.

Figure 9: **Frequency of Jump-Years**



Average number of years between detected jump periods for a range of significance levels of the jump-detection test. Data (solid line) is based on daily observations on real market returns from 1926 to 2008, while model average (dashed line) and $5\% - 95\%$ confidence band are based on 100 simulations of the time-varying volatility model.

Figure 10: **Model Frequency of Large Moves**



Model-implied distribution of time between years with detected large asset-price moves and an exponential distribution fit. Based on a long simulation of the time-varying volatility model.

References

Abel, Andrew, Janice Eberly, and Stavros Panageas, 2007, Optimal inattention to the stock market, *American Economic Review* 97, 244–249.

Ai, Hengjie, 2010, Information about long-run risks: Asset pricing implications, forthcoming in Journal of Finance.

Andersen, T., T. Bollerslev, F.X. Diebold, and C Vega, 2003, Micro effects of macro announcements: Real-time price discovery in foreign exchange, *American Economic Review* 93, 38–62.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52, 2003 – 2049.

Bansal, Ravi, Varoujan Khatchatrian, and Amir Yaron, 2005, Interpretable asset markets, *European Economic Review* 49, 531–560.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2007, Risks for the long run: Estimation and inference, working paper.

Bansal, Ravi, and Ivan Shaliastovich, 2007, Risk and return in bond, currency and equity markets, working paper.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Barndorff-Nielsen, Ole E., and Neil Shephard, 2006, Econometrics of testing for jumps in financial economics using bipower variation, *Journal of Financial Econometrics* 4, 1–30.

Barro, Robert, 2006, Rare disasters and asset markets in the Twentieth century, *Quarterly Journal of Economics* 121.

Bates, D., 1991, The crash of 87: Was it expected? The evidence from options markets, *Journal of Finance* 46, 1009 – 1044.

Bates, D., and R. Craine, 1999, Valuing the futures market clearinghouses default exposure during the 1987 crash, *Journal of Money, Credit, and Banking* 31.

Bates, David, 2000, Post-'87 crash fears in S&P 500 futures options, *Journal of Econometrics* 94, 181–238.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model specification and risk premia: Evidence from futures options, *Journal of Finance* 62, 1453–1490.

Croce, Mariano, Martin Lettau, and Sydney Ludvigson, 2010, Investor information, long-run risk, and the duration of risky cash flows, working paper.

David, Alexander, 1997, Fluctuating confidence in stock markets: Implications for returns and volatility, *Journal of Financial and Quantitative Analysis* 32, 427–462.

Drechsler, Itamar, 2010, Uncertainty, time-varying fear, and asset prices, working paper.

Drechsler, Itamar, and Amir Yaron, 2010, What's vol got to do with it, forthcoming in Review of Financial Studies.

Duffie, D., and T. Sun, 1990, Transactions costs and portfolio choice in a discrete-continuous-time setting, *Journal of Economic Dynamics and Control* 14, 35–51.

Epstein, Larry G., and Stanley Zin, 1989, Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Eraker, Bjorn, 2004, Do stock prices and volatility jump? Reconciling evidence from spot and option prices, *Journal of Finance* 59, 1367–1403.

Eraker, Bjorn, M. Johannes, and N.G. Polson, 2003, The impact of jumps in returns and volatility, *Journal of Finance* 53, 1269–1300.

Eraker, Bjørn, and Ivan Shaliastovich, 2008, An equilibrium guide to designing affine pricing models, *Mathematical Finance* 18, 519–543.

Grossman, Sanford, and Joeseph Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Hansen, Lars, and Thomas Sargent, 2010, Fragile beliefs and the price of uncertainty, *Quantitative Economics* 1, 129–62.

Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long-run risk, *Journal of Political Economy* 116, 260–302.

Huang, Xin, and George Tauchen, 2005, The relative contribution of jumps to total price variance, *Journal of Financial Econometrics* 3, 256–499.

Lipster, R.S., and A.N. Shiryaev, 2001, *Statistics of Random Processes* (Springer).

Liu, Jun, Jun Pan, and Tan Wang, 2005, An equilibrium model of rare-event premia and its implication for option smirks, *Review of Financial Studies* 18, 131–164.

Maheu, John M., and Thomas H. McCurdy, 2004, News arrival, jump dynamics, and volatility components for individual stock returns, *Journal of Finance* 59, 755–793.

Merton, R., 1976, Option pricing when underlying returns are discontinuous, *Journal of Financial Economics* 3, 125–144.

Naik, V., and M. Lee, 1990, General equilibrium pricing of options on the market portfolio with discontinuous returns, *Review of Financial Studies* 3, 493–522.

Pan, Jun, 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.

Reis, Ricardo, 2006, Inattentive consumers, *Journal of Monetary Economics* 53, 1761–1800.

Sims, C., 2003, Implications of rational inattention, *Journal of Monetary Economics* 50, 665–690.

Singleton, Kenneth J., 2006, *Empirical Dynamic Asset Pricing - Model Specification and Econometric Assessment* (Princeton University Press).

Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2006, Learning asymmetries in real business cycles, *Journal of Monetary Economics* 53, 753–772.

Veldkamp, Laura, 2006a, Media frenzies in markets for financial information, *American Economic Review* 96, 577–601.

Veldkamp, Laura, 2006b, Slow boom, sudden crash, *Journal of Economic Theory* 124, 230–257.

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

Weil, Philippe, 1989, The equity premium puzzle and the risk free rate puzzle, *Journal of Monetary Economics* 24, 401–421.