NBER WORKING PAPER SERIES

LEARNING AND ASSET-PRICE JUMPS

Ravi Bansal
Ivan Shaliastovich

Working Paper 14814
http://www.nber.org/papers/w14814

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2009

We would like to thank Hengjie Ai, Tim Bollerslev, Mikhail Chernov, Riccardo Colacito, Janice Eberly, Neil Shephard, George Tauchen and the participants of the 2008 AEA meeting and CCAPMWorkshop in Aarhus, Denmark for their comments. The usual disclaimer applies. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2009 by Ravi Bansal and Ivan Shaliastovich. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Learning and Asset-Price Jumps
Ravi Bansal and Ivan Shaliastovich
NBER Working Paper No. 14814
March 2009
JEL No. E0,E4,E44,G0,G1,G12

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 income volatility predicts future jumps, while the variation in the level of income does not. We find that indeed in the data large moves in returns are predicted by consumption volatility, but not by the changes in the consumption level. We show that the model can quantitatively capture these novel features of the data.

Ravi Bansal
Fuqua School of Business
Duke University
1 Towerview Drive
Durham, NC 27708
and NBER
ravi.bansal@duke.edu

Ivan Shaliastovich
Duke University
Durham, NC 27708
ivan.Shaliastovich@duke.edu

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 consumption 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 asset prices. These large moves in asset prices obtain even though the underlying income in the economy is smooth and has 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).

Earlier studies, such as Gennotte (1986), David (1997), Veronesi (1999) and Ai (2007) feature Kalman filter learning about the unobserved states to derive implications about the asset valuations in the economy. Veldkamp (2006) and Van Nieuwerburgh and Veldkamp (2006) specify a learning model where the endogenous information flow varies with the level of economic activity. Hansen and Sargent (2006) consider alternative approach to learning, which captures investors' concerns about robustness and potential model misspecification. In our approach, we modify standard Kalman filtering problem to account for endogenous learning about the true state after paying a cost. The endogeneity of agent's information set changes the sources and prices of risks relative to standard models; in particular, the actions of investors to learn about the true state alter asset valuations and can lead to asset-price jumps.

One of the key implications of the 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 volatility measures in the economy. Specifically, we document a positive correlation of return jump-indicator with lags of conditional variance of consumption. On annual frequency, the volatility of annual consumption significantly predicts next year large moves in 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, 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 24 years with at least one significant price move (i.e. jump) in daily return for the 80 year period from 1926 to 2006; hence, the frequency of jump-years is one every 3.4 years[2]. In our sample, we find that the 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 so we can 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 (2007a, b).[3] We show that at the calibrated value of learning cost parameters, investors optimally choose to observe the true state about once every one

[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 on a given year.

[3]Rational inattention channel is also used to explain infrequent adjustments of stock portfolio (Duffie and Sun 1990) or the consumption and saving plans of investors (Reis, 2006) .

and a half year, and the per annum expenditure on costly learning is merely 0.02% of the aggregate income. The model with constant aggregate volatility delivers the average frequency of jump-years once every 4.5 years, and the contribution of jumps to return variance of 7.7%. When we allow for time-varying aggregate volatility, the average frequency of jump-years increases to once every 4 years, while the relative contribution of jumps — to 9.5%. 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 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.

The rest of the paper is organized as follows. In the next section we review the empirical evidence on large infrequent movements in asset valuations in the data. In Section 3 we set up a model and describe preference, 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 follows.

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 1926-2006[4]. Occasional large spikes in the series suggest presence of large moves (jumps). Indeed, the empirical quantiles plot on Figure 2 indicates that there are substantial deviations in the distribution of market returns from Normality, with a number of observations falling far in the tails. 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 $j=1,2,\dots,M$. The

[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.

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 current absolute return and its lag:

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

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 statistics to test for abnormally large price movements, which are indicative of jumps. A popular version of this statistics is

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

[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.

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

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

Under the null hypothesis of no jumps and conditional on the sample path, the jump-detection statistics 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 large price move during the period T.

To calculate the jump-detection statistics over a year, we use the data on 266 daily returns, on average[6]. $M = 266$ 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 - 2006. At 1% significance level, we flag 24 years with at least one significant move in daily asset prices. 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.

Naturally, the detection of jump-years depends on the chosen significance level of the test. When the significance level drops, the cut-off value for the jump statistics in equation (2.4) increases, so that only larger jumps get flagged. On Figure 3 we plot the average frequency of detected jump-years in the data for a range of significance levels from 0.5% to 5%. As the significance level increases, the detected frequency of jump periods increases from one every 4 years to one every 2.5 years. In subsequent analysis, we fix the significance level to 1% – the results for other values are qualitatively very similar.

2.1 Predictability of Large Price Moves

In this section, we provide empirical evidence that fundamental macroeconomic volatility and variance of market returns can predict future large moves in asset prices in the data. On the other hand, there is no consistent evidence in the data for the link between large moves in returns and the levels of aggregate macro variables. That is,

[6]For predictability regressions, we construct similar measures on monthly and quarterly frequencies.

at the considered frequencies of large moves in returns, jumps in asset prices are not predicted by the variations in the level of real economy.

On the top panel of Figure 5 we plot the correlations of jump year indicator with annual consumption growth rate, its conditional variance and the conditional variance of market returns up to 5 year leads and lags (conditional variance calculations are based on AR(1)-GARCH(1) fit). The correlations of large move indicator with lagged aggregate volatility are all positive and are within 0.2-0.3 range. Similarly, positive market variance positively predicts future jump years 1 to 4 years ahead, with the correlation coefficient of about 0.1. As for the level of consumption growth, while the correlation coefficients are negative for 1 and 2 year lags and around -0.1, they turn positive at 3 year lag.

The predictability patterns are stronger at quarterly and monthly frequencies, as the persistence of variance measure and the frequency of detected 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.[7] On the bottom panels of Figure 5 we plot the lead-lag correlations of the jump indicator with levels and conditional volatilities of industrial production growth rate and variance of market return at quarterly and monthly frequency. The results present a robust evidence for positive correlations of future large move indicators with variance measures and no consistent link with the level of the real economy.

To sharpen quantitative results, we construct a measure of macroeconomic volatility based on the financial markets data. 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 2. 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 uncertainty.

We use the extracted factor $\hat{\sigma}^2$ to forecast next year jump indicator statistic. The probit regression of the next-period jump indicator on current measure of macroeconomic volatility yields a statistically significant coefficient on $\hat{\sigma}_T^2$ with a t-statistics in excess of 3, and R^2 of 9%. Specifically,

$$\hat{Pr}(JumpIndicator_{T+1}) = \Phi\left(\underset{(0.22)}{-0.95} + \underset{(469.35)}{1411.35}\hat{\sigma}_T^2\right),$$

[7] 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.

where $JumpIndicator_T$ is equal to 1 if year T is flagged as a jump-year and 0 otherwise. On Figure 4 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 2 shows that the estimated aggregate consumption volatility is very persistent in the data. The probit estimation of predictability of future jump-year indicator is given by,

$$\hat{Pr}(JumpIndicator_{T+1}) = \Phi\left(\underset{(0.21)}{-0.80} + \underset{(395.68)}{899.23}\hat{\sigma}_T^2\right),$$

so that the consumption volatility is statistically significant predictor of future jump years with $t-$statistics of 2.3, and the R^2 of 6%.

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).

3 Model Setup

Our model builds on the long-run risks framework developed in Bansal and Yaron (2004), where the investor has a full information about the economy. In contrast, we 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. In the model we show that the actions of the agents to learn the unobserved states 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 solve the model, we consider an optimization problem of the social planner, who wants to optimally allocate the exogenous output stream between consumption and learning expenditures. The solution to the social planner problem can then be decentralized in the competitive markets, which we verify by solving the representative agent problem directly.

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 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.

Define 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. It is similar to costs of observing the value of wealth and costs of transferring assets and rebalancing the portfolio featured in rational inattention literature, see Abel, Eberly, and Panageas (2007a, b). 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 are 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 opposite 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. For example, in Appendix A.2 we show that with expected utility, the agent never learns about the true state even at zero costs, as there is no preference for early resolution of uncertainty. On the other hand, if investors have preference for early resolution of uncertainty, they will choose to learn for a cost if the the cost is small enough.

The decentralization of the social planner problem leads to the usual equilibrium Euler equation

$$E_t^{s_t^*}\left[M_{t+1}R_{i,t+1}\right] = 1, \tag{3.9}$$

where $R_{i,t+1}$ is the return on any asset traded in the economy. The expression for the discount factor M_{t+1} is also standard, safe for an endogenous information set which depends on the optimal choice indicator s_t^* :

$$M_{t+1} = \beta\left(\frac{Y_{t+1}}{Y_t}\right)^{-1}\frac{U_{t+1}^{1-\gamma}}{E_t^{s_t^*}(U_{t+1}^{1-\gamma})}. \tag{3.10}$$

3.3 Income Dynamics

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

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

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. The learning costs in this paper are similar to the observation and information costs in rational inattention literature; see Abel, Eberly, and Panageas (2007a), Reis (2006) and Duffie and Sun (1990).

We apply a separation principle and solve a filtering problem before the optimization problem of the agent. 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 $\hat{x}_t(s_t)$:

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

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 unobserved expected income state in a Kalman filter manner. Indeed, as the income volatility is observable, 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 next period. 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 true consumption shock, $u_{t+1}(1) = \sigma_t\eta_{t+1}$. The filtered consumption innovation is used to update the estimate of the expected growth $\hat{x}_{t+1}(0)$, as shown in (3.17). This Kalman filter estimate is known to the agent at the beginning of next period. If investors decide to pay the cost to learn the true x_{t+1}, the expected income and variance of the filtering error are immediately adjusted to reflect the new information. We can then express the values of the states tomorrow in the following way:

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

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

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 volatility. Therefore the dynamics of the income volatility is independent of the learning choice of the agent and follows (3.13). If income volatility is constant, we obtain a standard Kalman Filter result that the variance of the filtering error $\omega^2_t(0)$ 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 and typically increases faster at times of heightened aggregate volatility. Learning models considered by David (1997) and Veronesi (1999) use regime-shift specification for expected growth component and feature alternative time-varying dynamics of the filtering uncertainty.

We specified the evolution of the economy in the beginning of the next period in (3.16)-(3.18), and instantaneous adjustments of the expected income and variance of the filtering error when the agent chooses to learn the true state for a cost in (3.20)-(3.21). The decision to learn is endogenous and is determined as a part of the equilibrium solution of the model, which we discuss in the next section.

4 Model Solution

4.1 Optimal Costly Learning

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. 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 model. In Appendix A.3 we show that the life-time utility per income ratio can be written in the following way:

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

where the sensitivity of the utility to expected income growth is independent of the costly learning choice and is given by

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

The volatility function $f(s_t, \sigma_t^2, \omega_t^2(0))$ depends on the learning choice s_t, the volatility states σ_t^2 and $\omega_t^2(0)$, as well as risk aversion of the agent γ and learning cost parameter χ. It satisfies the recursive equation given in the Appendix (A.18).

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), investor's life-time utility with no learning is

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

while the certainty equivalent of the life-time utility (per income) with costly learning is

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

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 depends on the recursive preferences of the agent. Indeed, when $\gamma = 1$, the volatility functions $f(s_t, \sigma_t^2, \omega_t^2(0))$ are constant and depend only on the learning choice s_t; moreover, the level of life-time utility following a learning choice is smaller than the level of utility without learning for any positive learning cost parameter χ. Hence, in expected utility case, the agent never learns for a cost. On the other hand, with recursive utility, the volatility functions $f(s_t, \sigma_t^2, \omega_t^2(0))$ are time-varying, so that the ex-ante life-time utilities of the agent with and without learning depend on the income volatility and variance of the filtering error. Learning choice is optimally determined by the relative difference between these utilities. Given our distributional assumptions on the economy, this difference depends only on volatility states $w_t^2(0)$ and σ_t^2, but not on the expected growth:

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

Hence, the optimal learning choice of the agent is governed by the income volatility, variance of the filtering error and preference and learning cost parameters. In general, solutions for the volatility function f and the optimal learning choice s_t^* are not available in a closed form, so we have to solve for them numerically. In the case when income shocks are homoscedastic, the variance of the filtering error is a deterministic function of time since the last costly learning, so investors optimally learn about the true state at constant time intervals. On the other hand, when income volatility is time-varying, we show that the agent chooses to learn the true state when the filtering variance gets too high relative to the underlying volatility of the income growth, so that the utility benefits from reducing the uncertainty outweigh the learning costs. Further, the frequency of costly learning increases at times of heightened income volatility, as the filtering uncertainty accumulates faster when income volatility rises.

4.2 Risk Compensation and Asset Prices

Using the solution to the equilibrium discount factor in (3.10), we can express the equilibrium discount factor in terms of the underlying variables in the economy. The conditional mean of the discount factor is equal to the negative of the expected income growth plus the contribution of the income and filtering variance:

$$\begin{aligned} E_t^{s_t} m_{t+1}(s_t) = \log\beta - \mu - \hat{x}_t(s_t) - \frac{1}{2}(1-\gamma)^2(BK_t(s_t)+1)^2(\omega_t^2(s_t)+\sigma_t^2) \\ - (\gamma-1)E_t^{s_t} f_{t+1} - \ln E_t^{s_t} e^{(1-\gamma)\left[f_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega_{t+1}^2(0)s_{t+1}^*\right]}. \end{aligned} \tag{4.6}$$

The innovation into the log discount factor satisfies,

$$\begin{aligned} m_{t+1}(s_t) - E_t^{s_t} m_{t+1}(s_t) = &- \left(1+(\gamma-1)(1+BK_t(s_t))\right) u_{t+1}(s_t) \\ &- (\gamma-1)Bs_{t+1}^*(x_{t+1}-\hat{x}_{t+1}(0)) - (\gamma-1)(f_{t+1}-E_t^{s_t} f_{t+1}). \end{aligned} \tag{4.7}$$

If the expected endowment growth factor is observable by the agent, (e.g., the information costs are zero), we obtain a standard long-run risks setup. The price of a short-run consumption risk is then γ, and prices of long-run and volatility risks are constant and provided in the above studies.

The option to learning about the true state for a cost changes the sources and prices of risks relative to standard models. In our model, the price of immediate consumption risk u_{t+1} is time-varying because the expected growth state is unobservable and has to be learned from the data. As the decision to learn about the true state depends on income volatility, the pricing of income volatility shocks is also more complicated, which is reflected in the non-linear volatility function f_{t+1}. Finally, while in the standard long-run risks model the agent fears the true innovations into the expected growth, $x_{t+1} - E_t x_{t+1}$, in our model its counterpart is the revision of the state, $s^*_{t+1}(x_{t+1} - \hat{x}_{t+1}(0))$. Although both risks have the same market price $(\gamma - 1)B$, the amount of risk can be much higher in our case as the variance of the filtering error can significantly exceed the conditional volatility of long-run risks shocks. Therefore, the option to learn about the true state can have a considerable effect on the asset valuations in the economy. In an alternative specification with robust control, Hansen and Sargent (2006) show that the concerns of the agents about model misspecification can also generate substantial magnification of the risk premia.

Consider an asset with a dividend stream proportional to income growth,

$$\Delta d_t = \mu + \varphi_d(\Delta y_t - \mu). \tag{4.8}$$

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 (4.6)-(4.7) and the Euler condition (3.9), we can solve for the equilibrium log price-dividend ratio,

$$v_t(s_t) = H\hat{x}_t(s_t) + h(s_t, \sigma_t^2, \omega_t^2(0)), \tag{4.9}$$

where the solutions for H and $h(s_t, \sigma_t^2, \omega_t^2(0))$ are given in Appendix A.4.

The asset valuations depend on filtered or, if $s_t = 1$, true expected income growth. When investors learn about the true state, the equilibrium price-dividend ratio responds to the the revision in the expected income state magnified by H. For example, as H is positive, when the true x_t is lower than what the agent expected, asset prices can fall sharply. 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. As the volatility of equilibrium returns also increases in aggregate volatility, the model can thus explain the predictability of future asset-price jumps by macroeconomic and return variance. In Section 5 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 daily frequency. The baseline calibration parameter values, which are reported in Table 3, are similar to the ones used in standard long-run risks literature (see e.g. Bansal and Yaron, 2004), safe for an adjustment to daily frequency. Specifically, we set the persistence in the expected income growth ρ at 0.9963. The choice of φ_e and σ_0 ensures that the model matches the annualized aggregate volatility of about 2%, while the volatility persistence is set to $\nu = 0.99$. 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 4 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 δ equal 0.9999 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 in rational inattention literature, see Abel, Eberly, and Panageas (2007a, b). At this level of learning costs, investors are willing to optimally learn the true state about once every one and a half year, so that expenditure on costly learning accounts merely for 0.02% of annual aggregate income. 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. We discuss these issues later in the section.

5.2 Constant Volatility Case

First we consider a special case when income volatility is constant, that is, $\sigma_t = \sigma_0$. This provides a useful benchmark case as the numerical solution to the model is greatly simplified. Indeed, in a standard homoscedastic Kalman Filter setup the variance of the filtering error ω_t^2 is a deterministic function of time since the last learning about the true state. Therefore, the optimal learning policy is purely time-dependent, so that the agent chooses to learn every Nth period, where N depends on the preference and income dynamics parameters of the model. We provide the details of the solution to the model in Appendix A.5.

On Figure 6 we show the optimal length of the filtering period N as a function of the cost parameter χ for risk aversion levels of 5, 10 and 15. When cost of learning increases, the agent chooses to learn the true expected income state less frequently.

This decision is also very sensitive to the risk attitude of the investors, so that more risk-averse investors learn about the true state more often for any value of the learning cost parameter. On the other hand, when the risk-aversion coefficient is less than or equal to 1, the agent will never choose to learn the true state for a cost, as there is no preference for early resolution of uncertainty.

Table 1 reports asset-pricing implications of the option to learn for a cost in the model with constant income volatility. Simulated mean and volatility of returns are 7.9% and 15.3%, respectively, which match statistics in the data. On Figure 8 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 0.7. 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 (4.9). In presence of highly persistent long-run risks shocks, asset prices are very sensitive changes in expected income state. Therefore, even small deviations in the filtered state from the truth, when uncovered, can lead to large changes in valuations that look like large price moves. Notably, although the number of periods between successive days with costly learning is constant, the years with flagged jumps do not have to occur at regular intervals, as shown on Figure 8. 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.

For comparison, the last graph on Figure 8 depicts the equilibrium market returns which would obtain in this economy if agent could not learn the true state for a cost and have to exclusively rely on standard Kalman filtering. As can be seen in the second panel of Table 1, the specification with no learning cannot deliver large price movements observable before. The jump-detection statistics typically do not find more than 2 or 3 instances of large price moves in 80 years of simulated daily data; the detected jumps represent pure-chance large random draws in the simulation. On the other hand, when agent can learn about the true state for a cost, the detected jump frequency is about 4.5 years, and the contribution of jumps to the total return variation is 7.7%. These numbers are consistent with the data (see first panel of Table 1).

Similar conclusions obtain from comparison of the unconditional distributions of returns. When the agent has no option to learn the true state, the kurtosis of return distribution is equal to 3, relative to 18 for the returns in the economy with learning, and 21 for the data. As income volatility is kept constant, the heavy tails in the return distribution are driven solely by the discrete adjustments to the asset valuations.

The constant-volatility case can deliver the key result that the equilibrium asset prices can display infrequent large movements which cannot be explained by standard

Gaussian shocks. However, when income shocks are homoscedastic, the decision to learn is purely time-dependent, and volatilities of macroeconomic and financial variables are constant. We can address these issues by opening up stochastic volatility channel, which we discuss in the next section.

5.3 Time-Varying Volatility Case

The asset-pricing implications of costly learning in a time-varying volatility setup take into account state-dependence and time-variation of the optimal costly learning rule. Indeed, as both aggregate volatility and filtering variance are now time-varying, the optimal decision to learn about the true state is stochastic, as shown in income dynamics simulation on Figure 9. We further characterize the dependence of optimal costly learning on filtering and income volatilities on Figure 7, which depicts the expected number of periods till next costly learning given current filtering variance for high, medium and low values of aggregate volatility. Investors choose to learn for a cost if the variance of the filtering error grows too high in the economy, so that it is optimal to sacrifice part of the current endowment, pay learning cost and reduce the amount of uncertainty. Costly learning is more frequent in high income volatility states, as in those states the uncertainty about the filtered estimate accumulates faster and is expected to reach the costly learning cut-off point sooner. These 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 10. Relative to constant volatility case, the detected jump-years are more frequent, averaging one every 4 years, and contribute more to the total variation in returns, 9.5% versus 7.5% in a constant-volatility case and in the data (see Table 1). For robustness, on Figure 11 we also show the model-implied jump-year frequency for a range of significance levels for the jump-detection test. 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 broadly match the evidence on the average frequency of jump-years in the data, as all the values are well within the $5\% - 95\%$ confidence band. These large moves in returns cannot be obtained in economy without costly learning, as can be visually seen on the time-series plot of returns on Figure 10. Without costly learning, the average frequency of detected jump-years is less than 1 in 80 years, and the detected "jumps" are merely pure-chance large random draws. The comparison of the fourth moments of return distribution is revealing: without an option to learn, the kurtosis of market returns is 3, and it reaches 35 when the agent can learn the true state for a cost.

On Figure 11 we plot the unconditional distribution of the number of periods between the detected jump-years based on long simulation from the full model with time-varying consumption volatility. In jump-diffusion models of asset prices with constant arrival intensity of jumps, the number of periods between successive jumps is exponential, so for comparison, we also provide an exponential fit to the jump-year duration distribution in the model. 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 tale than exponential, so a jump-year is likely to follow another. Clustering and predictability of jumps is an important aspect of our model, which we discuss in the next section.

5.4 Predictability of Jumps

In the model with time-varying consumption volatility, the frequency of learning and consequently, the likelihood of price jumps, is increasing with aggregate volatility, so that returns jumps are more frequent at times of high 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 levels of income 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. On Figure 12 we show model-implied lead-lag correlations of jump indicator with endowment growth and conditional variance of endowment growth and returns at monthly frequency, constructed in the same way as the empirical counterparts on Figure 5. The model with costly learning delivers positive and significant correlation of 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. As shown on the bottom panel of Figure 12, the model with no costly learning cannot account for the predictability of return jumps in the data, as all the correlation coefficients are zero. Results on quarterly and annual frequencies are very similar, and are omitted in the interest of space.

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. In Appendix A.6 we discuss the estimation results in the data and full model. The model can match quite well the dynamics of the time-varying volatility of returns, as well as the key findings in the literature regarding the frequency and predictability of jumps. In particular, the frequency of jumps is positive and highly significant in the data and the model. The jumps explain 10% of the variance of returns in the model and in the data, and the estimated frequency of jumps is one every one and two years, 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 in the economy, 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 income 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 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 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. 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 $\chi = 0$. Then, the utility of the agent corresponding to the indicator variable $s_t \in \{0, 1\}$ satisfies

$$U_t(s_t) = E_t^{s_t} \sum_{j=0}^{\infty} \beta^j u(Y_{t+j}). \tag{A.8}$$

The optimal learning policy in the expected utility case is based on the ex-ante expected utility given the beginning of period information. However, applying the law of iterated expectations, we obtain that

$$E_t U_t(1) = E_t \left(E_t^1 \sum_{j=0}^{\infty} \beta^j u(Y_{t+j}) \right) = E_t^0 \sum_{j=0}^{\infty} \beta^j u(Y_{t+j}) = U_t(0). \tag{A.9}$$

In expectation, new information does not increase the utility of the agent. Therefore, even though the new information is costless, the agent has no incentive to learn it.

The results are very different when the agent has a preference for a timing of the resolution of uncertainty. 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:

$$s_t^* = 1 \left[(1-\chi)^{1-\beta} \left(E_t \left[E_t^1 \left(\phi_{t+1} \frac{Y_{t+1}}{Y_t} \right)^{1-\gamma} \right]^{\beta} \right)^{\frac{1}{1-\gamma}} > \left[E_t \left(\phi_{t+1} \frac{Y_{t+1}}{Y_t} \right)^{1-\gamma} \right]^{\frac{\beta}{1-\gamma}} \right]. \tag{A.10}$$

In general, the equilibrium utility per income ratio ϕ_t depends on the whole history of choice indicators $(s_t, s_{t-1}, \ldots)$. Let us make a simplifying assumption, which holds in the model setup of this paper, that if the agent learns the state today, then all the past learned states become irrelevant for the life-time utility of the agent. That is, $\phi_t(1)$ does not depend on the variables observed before time t. Now consider a case when the agent has a preference for early resolution of uncertainty, i.e. when the risk aversion parameter γ is greater than one. Conjecture that it is optimal to learn tomorrow, so that by assumption above, ϕ_{t+1} does not depend on the current choice indicator s_t. Then, using Jensen's inequality type argument, it is easy to show that the agent learns the true state for a cost today if cost parameter χ is 0. Therefore, for small information costs, it is always optimal to learn about the underlying states if the agent prefers early resolution of uncertainty. On the other hand, if information costs are large, ($\chi \to 1$), it is never optimal learn, so that s_t is fixed at 0 in all time periods. For medium costs, the optimal solution to the choice indicator depends on the underlying state variables in the economy.

A.3 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.5 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,

$$\phi(s_t, \hat{x}_t(s_t), \omega_t^2(0), \sigma_t^2) = e^{B\hat{x}_t(s_t)+f(s_t,\sigma_t^2,\omega_t^2(0))}, \tag{A.11}$$

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).

Let us fix the optimal choice s_{t+1}^* tomorrow. We conjecture that 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:

$$\begin{aligned}\phi_{t+1} &= \phi(s_{t+1}^*, \hat{x}_{t+1}(s_{t+1}^*), \omega_{t+1}^2(0), \sigma_{t+1}^2) \\ &= e^{B\hat{x}_{t+1}(s_{t+1}^*)+f_{t+1}},\end{aligned} \tag{A.12}$$

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),

$$\log\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right) = B\left(\hat{x}_{t+1}(0) + s_{t+1}^*(x_{t+1} - \hat{x}_{t+1}(0))\right) + f_{t+1} + \Delta y_{t+1}. \tag{A.13}$$

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,

$$x_{t+1}|\mathcal{I}_{t+1} \sim N(\hat{x}_{t+1}(0), \omega^2_{t+1}(0)), \tag{A.14}$$

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

$$\begin{aligned} E_t^{s_t}\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma} &= E_t^{s_t} e^{(1-\gamma)\left[B\hat{x}_{t+1}(0)+f_{t+1}+\Delta y_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega^2_{t+1}(0)s^*_{t+1}\right]} \\ &= e^{(1-\gamma)(\mu+(B\rho+1)\hat{x}_t(s_t))} E_t^{s_t} e^{(1-\gamma)\left[(BK_t(s_t)+1)u_{t+1}(s_t)+f_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega^2_{t+1}(0)s^*_{t+1}\right]}. \end{aligned} \tag{A.15}$$

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,

$$\begin{aligned} E_t^{s_t}\left(\phi_{t+1}\frac{Y_{t+1}}{Y_t}\right)^{1-\gamma} &= e^{(1-\gamma)\left(\mu+(B\rho+1)\hat{x}_t(s_t)+\frac{1}{2}(1-\gamma)(BK_t(s_t)+1)^2(\omega^2_t(s_t)+\sigma^2_t)\right)} \\ &\times E_t^{s_t} e^{(1-\gamma)\left[f_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega^2_{t+1}(0)s^*_{t+1}\right]} \end{aligned} \tag{A.16}$$

Therefore, using the equilibrium utility recursion (A.6) and the conjectured solution for the life-time utility of the agent (A.11) and matching the coefficients, we obtain that loading on expected growth is equal to

$$B = \frac{\beta}{1-\beta\rho}, \tag{A.17}$$

while the volatility function satisfies

$$\begin{aligned} &f(s_t, \sigma^2_t, \omega^2_t(0)) = (1-\beta)\ln(1-s_t\chi) + \beta\mu \\ &+ \beta\frac{1}{2}(1-\gamma)(BK_t(s_t)+1)^2(\omega^2_t(s_t)+\sigma^2_t) + \frac{\beta}{1-\gamma}\ln E_t^{s_t} e^{(1-\gamma)\left[f_{t+1}+\frac{1}{2}(1-\gamma)B^2\omega^2_{t+1}(0)s^*_{t+1}\right]}. \end{aligned} \tag{A.18}$$

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

Now, given the utility equation (A.11) 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:

$$s^*_t = 1\left[\frac{1}{2}(1-\gamma)B^2\omega^2_t(0) + f_t(1, \sigma^2_t, \omega^2_t(0)) > f_t(0, \sigma^2_t, \omega^2_t(0))\right]. \tag{A.19}$$

Using the optimal condition for s^*_{t+1} tomorrow to rewrite the recursive equation of the volatility function (A.18) in the following way:

$$f(s_t, \sigma_t^2, \omega_t^2(0)) = (1-\beta)\ln(1-s_t\chi) + \beta\mu + \beta\frac{1}{2}(1-\gamma)(BK_t(s_t)+1)^2(\omega_t^2(s_t)+\sigma_t^2)$$
$$+ \frac{\beta}{1-\gamma}\ln E_t^{s_t} e^{(1-\gamma)\max\left[\frac{1}{2}(1-\gamma)B^2\omega_{t+1}^2(0)+f_{t+1}(1,\sigma_{t+1}^2,\omega_{t+1}^2(0)),f_{t+1}(0,\sigma_{t+1}^2,\omega_{t+1}^2(0))\right]}. \tag{A.20}$$

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).

A.4 Dividend Asset

Consider an asset with dividend stream $\Delta d_t = \mu + \varphi_d(\Delta y_t - \mu)$.

The equilibrium price-dividend ratio solves,

$$v_t(s_t) = H\hat{x}_t(s_t) + h(s_t, \sigma_t^2, \omega_t^2(0)), \tag{A.21}$$

The log-linearized returns satisfy

$$r_{d,t+1} = \kappa_0 + \varphi_d\mu + (H(\kappa_1\rho - 1) + \varphi_d)\hat{x}_t(s_t) - h(s_t, \sigma_t^2, \omega_t^2(0))$$
$$+ (\kappa_1 HK_t(s_t) + \varphi_d)u_{t+1}(s_t) + \kappa_1 h(s^*_{t+1}, \sigma_{t+1}^2, \omega_{t+1}^2(0)) + \kappa_1 H s^*_{t+1}(x_{t+1} - \hat{x}_{t+1}(0)). \tag{A.22}$$

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

$$H = \frac{\varphi_d - 1}{1 - \kappa_1\rho}. \tag{A.23}$$

The price-dividend levels are given recursively by

$$h_t(s_t, \sigma_t^2, \omega_t^2(0)) = \ln\beta + \kappa_0$$
$$+ \frac{1}{2}(\varphi_d - 1 + \kappa_1 HK_t(s_t))(\varphi_d - 1 + \kappa_1 HK_t(s_t) - 2(\gamma-1)(1+BK_t(s_t)))(\sigma_t^2 + \omega_t^2(s_t))$$
$$+ \ln E_t^{s_t} e^{\kappa_1 h_{t+1} + \frac{1}{2}(\kappa_1 H - (\gamma-1)B)^2 s^*_{t+1}\omega_{t+1}^2(0) - (\gamma-1)f_{t+1}}$$
$$- \ln E_t^{s_t} e^{(1-\gamma)\left(f_{t+1} + \frac{1}{2}(1-\gamma)B^2 s^*_{t+1} w_{t+1}^2(0)\right)}. \tag{A.24}$$

To solve for the approximating constants κ_0 and κ_1, we use the numerical procedure discussed in Bansal, Kiku, and Yaron (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.5 Constant Volatility Case

In the general case with time-varying volatility, we first solve the model numerically by discretizing income and filtering volatility states and applying fixed-point iterations to the volatility functions $f(s_t, \sigma_t^2, \omega_t^2(0))$ and $h(s_t, \sigma_t^2, \omega_t^2(0))$. We verify that the numerical solutions to these functions are very close to being linear, so in calibrations we approximate these function to be affine in the two volatilities, and solve for the loadings numerically.

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 :

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

Now we can rewrite the recursions in (A.18) as a system of linear equations (to simplify the exposition, we consider the case $N > 2$):

$$\begin{aligned} f_1 - \beta f_2 &= \beta\mu + \frac{1}{2}\beta(1-\gamma)(BK_1(0)+1)^2(\omega_1^2(0)+\sigma_0^2), \\ &\dots \\ f_i - \beta f_{i+1} &= \beta\mu + \frac{1}{2}\beta(1-\gamma)(BK_i(0)+1)^2(\omega_i^2(0)+\sigma_0^2), \quad 2 \leq i < N-1 \\ f_{N-1} - \beta f_N &= \beta\mu + \frac{1}{2}\beta(1-\gamma)\left((BK_{N-1}(0)+1)^2(\omega_{N-1}^2(0)+\sigma_0^2) + B^2\omega_N^2(0)\right), \\ f_N - \beta f_1 &= (1-\beta)\ln(1-\chi) + \beta\mu + \frac{1}{2}\beta(1-\gamma)(BK_N(1)+1)^2(\omega_N^2(1)+\sigma_0^2). \end{aligned} \tag{A.25}$$

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.19))

$$\frac{1}{2}(1-\gamma)B^2\omega_i^2(0) + f_N < f_i, \tag{A.26}$$

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.18), 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 optimality condition for choice indicator (A.19), we then need to have that

$$\frac{1}{2}(1-\gamma)B^2\omega_N^2(0) + f_N > \tilde{f}_N. \tag{A.27}$$

In practice, we loop from a low value of N until we satisfy both optimality conditions (A.26)-(A.27), where the volatility functions f_i solve the linear system (A.25). In numerical calibrations, the optimal N is always unique: when N is lower than optimum, we violate the last condition (A.27), 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.26) 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.24) much in the same way as (A.25), 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).

A.6 Parametric Jump Model

To compare our model implications to the results from the above studies, we fit a discrete-time GARCH-jump specification for returns, which feature autoregressive stochastic volatility and time-varying arrival intensity of jumps in returns[8]. Specifically, the return dynamics is given by

$$r_t = \mu_r + a_{1,t} + a_{2,t}. \tag{A.28}$$

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

Estimation of GARCH-jump model				
	Data		Model	
	Estimate	S.E.	Mean	S.E.
σ_v	0.08	(0.004)	0.12	(0.021)
β_v	0.91	(0.004)	0.94	(0.018)
α_v	0.07	(0.003)	0.03	(0.005)
λ_0	0.00	(0.001)	0.00	(0.002)
λ_l	25.78	(16.21)	152.00	(60.06)
μ_j	-1.6	(0.46)	0.60	(0.78)
σ_j	2.9	(0.33)	4.30	(0.77)

Estimation of GARCH-jump model using the sample and simulated data.

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:

$$a_{1,t} = \sqrt{v_{t-1}} z_t, \quad z_t \sim N(0,1), \tag{A.29}$$

$$v_t = \sigma_v^2 + \beta_v v_{t-1} + \alpha_v (r_t - \mu_r)^2. \tag{A.30}$$

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

$$a_{2,t} = \sum_{k=1}^{n_t} \xi_{t,k} - \mu_j \lambda_{t-1}. \tag{A.31}$$

The jump size distribution is Normal:

$$\xi_{t,k} \sim N(\mu_j, \sigma_j^2), \tag{A.32}$$

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

$$Pr_{t-1}(n_t = j) = \frac{exp(-\lambda_{t-1}) \lambda_{t-1}^j}{j!}. \tag{A.33}$$

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,

$$\lambda_t = \lambda_0 + \lambda_l v_t. \tag{A.34}$$

The above specification of return dynamics can be readily estimated by MLE using the sample and simulated data. In estimations, as we want to capture large, infrequent moves in asset prices, we restrict the unconditional jump intensity not to exceed 1 jump per year.

We perform a Monte-Carlo study where we estimate specification (A.28)-(A.34) for 100 simulations of 80 years of daily returns from a time-varying volatility model, and we compare the results to the estimates based on the sample data. As can be seen in Table A.6, the model matches quite well the dynamics of the time-varying volatility of smooth component of the returns: the overall persistence is 0.96 in the model, compared to 0.98 in the data, and the intercept and ARCH and GARCH parameter coefficients are close as well. The model can also capture the key findings in the literature regarding the frequency and predictability of jumps. The estimated jump arrival intensity loading on market variance is positive and highly significant both in the model and in the data, though, it is estimated with a large standard error. The mean jump size is -1.6% in the data, and is slightly positive but insignificant in the model. The standard deviations of the jump distribution are 3% and 4% in the data and model, respectively. The jumps explain about 10% of the variation of the returns in the data and the model, and their average frequency is once in one and two years, respectively.

Tables and Figures

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

	Mean	Std Dev	Kurt	Jump-year Freq	Jump Contribution
Data:					
Return	7.98	16.61	21.19	3.38	7.46
Model Output:					
Constant Volatility:					
Return with costly learning	7.92	15.28	18.37	4.47	7.68
Return, no learning	8.10	14.10	3.01	47.93	1.94
Time-Varying Volatility:					
Return with costly learning	7.82	15.22	35.14	3.97	9.47
Return, no learning	7.92	14.12	3.21	48.82	2.00

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 aggregate volatility, respectively. No 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 large price movements as flagged by jump-detection statistics, in years. Jump Contribution measures the average percent contribution of large price moves to total return variance. Data are daily inflation-adjusted market returns for 1926 - 2006. Model statistics are based on the average across 100 simulations of 80 years of data. Jump-detection statistics are based on 1% significance level.

Table 2: **Estimation of Consumption Volatility**

Projection:	Δc	pd	$spread$		R^2
$\widehat{\Delta c}$	0.347	0.007	-0.004		0.28
	(0.065)	(0.006)	(0.004)		
$\widehat{\sigma}^2 \times 10^4$		-0.546	4.602		0.23
		(1.156)	(2.056)		

GARCH:	$AR(1)$	Vol	$GARCH$	$ARCH$	R^2
$\widehat{\Delta c}$	0.32	2.81e-06	0.82	0.14	0.10
	(0.11)	(7.5e-06)	(0.07)	(0.06)	

Estimation of conditional consumption volatility based on projection of consumption growth and extracted squared consumption residual on price-dividend ratio and junk bond spread (top panel) and on AR(1)-GARCH(1,1) specification. Data is annual real consumption growth for 1930-2006.

Table 3: **Model Calibration**

Parameter	Value
μ	7.27e-05
ρ	0.9963
σ	8.53e-04
ν	0.99
σ_w	5.07e-05
φ_e	8.95e-03
φ_d	5
β	0.9999
γ	10
χ	0.080087

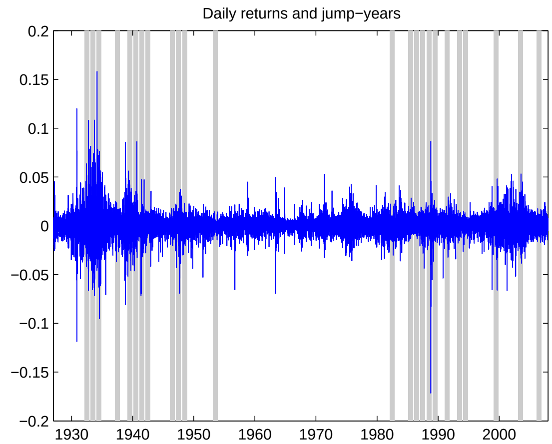
Calibrated parameter values, daily frequency.

Table 4: **Endowment Dynamics: Data and Model Calibration**

	Data		Model		
	Estimate	S.E.	Median	5%	95%
Mean	1.95	(0.32)	1.94	1.18	2.57
Vol	2.13	(0.52)	2.17	1.81	2.60
AR(1)	0.44	(0.13)	0.52	0.37	0.64
AR(2)	0.16	(0.18)	0.15	-0.06	0.36
AR(5)	-0.01	(0.10)	-0.02	-0.27	0.17

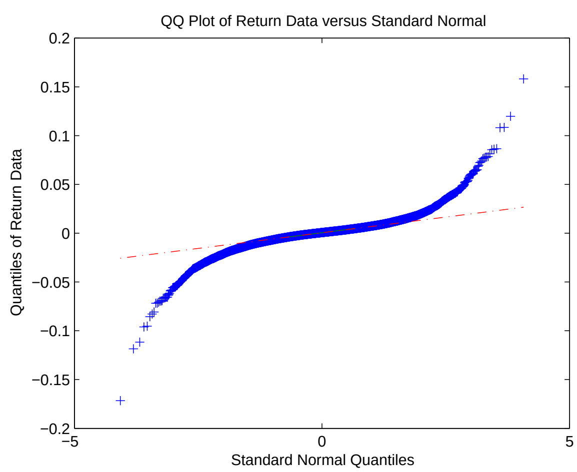
Calibration of income dynamics. Data is annual real consumption growth for 1930-2006. Model is based on 100 daily simulation of 80 years of income aggregated to annual horizon.

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



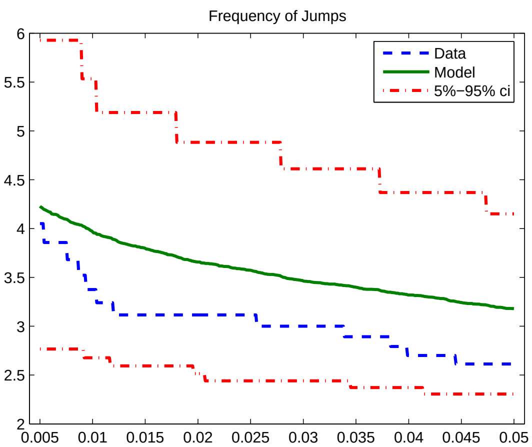
Daily observations on real market returns from 1926 to 2006. Grey regions correspond to periods with at least one significant large price move, at 1% significance level.

Figure 2: **Return Quantiles in the Data**



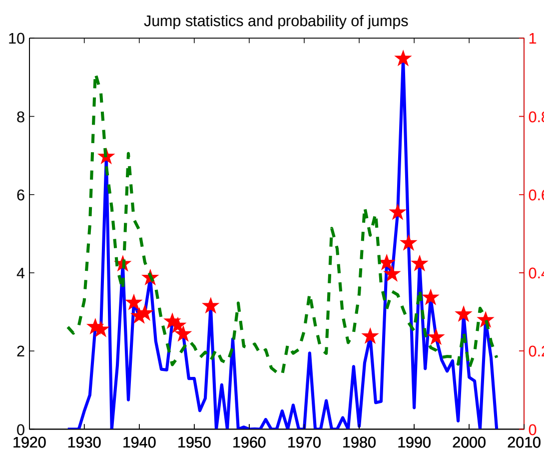
Quantiles of return data versus Normal distribution. Daily observations on real market returns from 1926 to 2006.

Figure 3: **Frequency of Jump-Years**



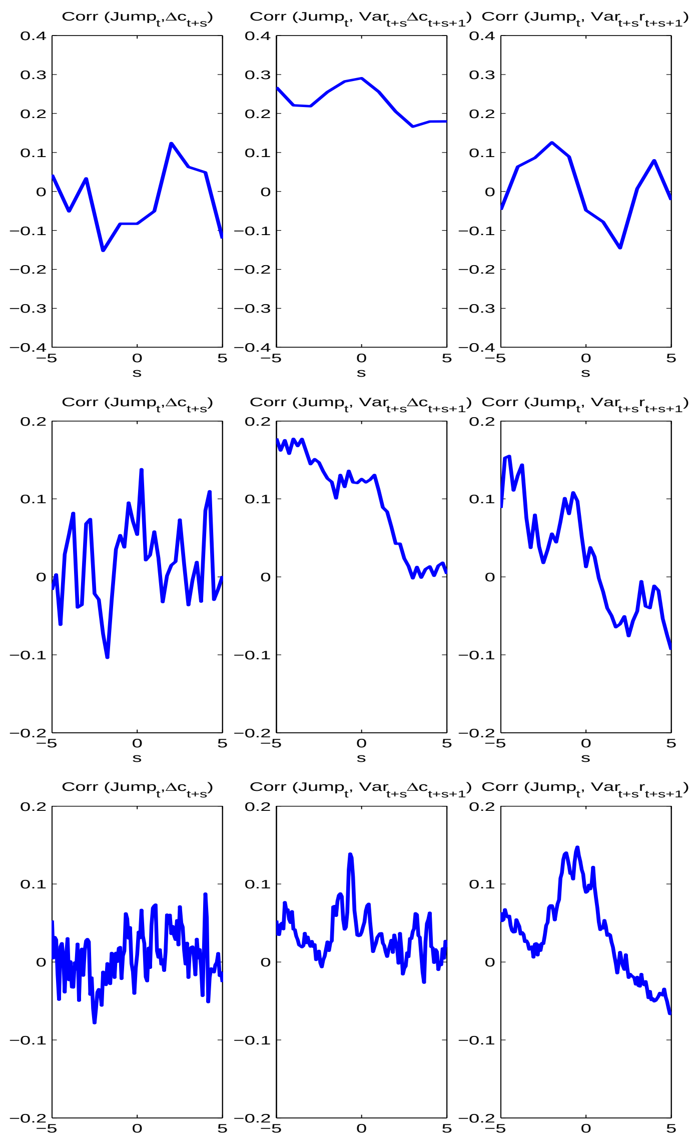
Average number of years between detected jump periods for a range of significance levels of jump-detection test. Data (solid line) is based on daily observations on real market returns from 1926 to 2006, while model average (dashed line) and 5% − 95% confidence band are based on 100 simulations of the full model.

Figure 4: **Predicted Probability of Large Price Moves**



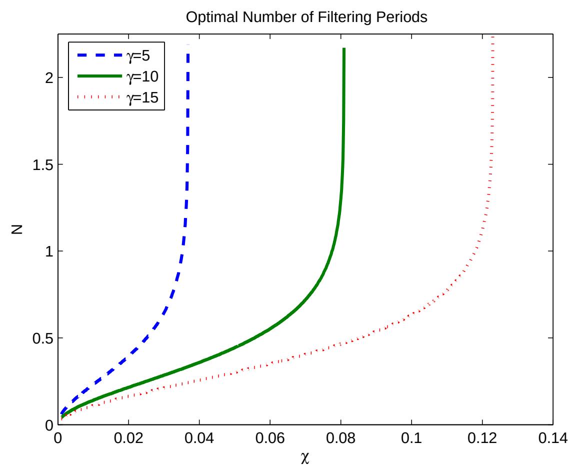
Annual jump statistics (solid line with values on the left Y-axis) and the predicted probability of large price moves (dashed line with values on the right Y-axis) based on the aggregate volatility. Stars indicate years with at least one large price move.

Figure 5: **Jump Correlations in the Data**



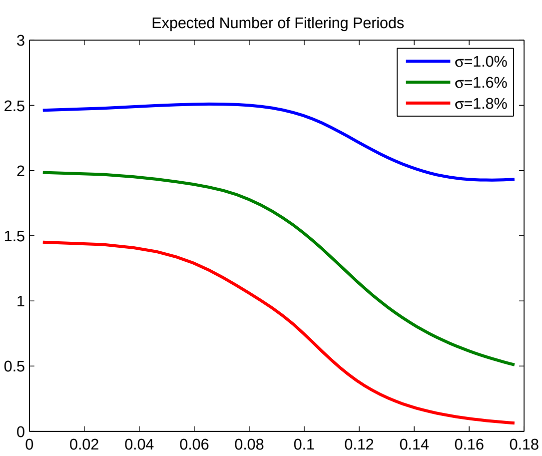
Correlation of return jump indicator with level of real economy (left panel), aggregate volatility (middle panel) and conditional variance of returns at up to 5 year leads and lags. Top panel is based on annual observations of real consumption growth from 1930 to 2006; middle and bottom panels are based on the series of industrial production from 1926 to 2006, quarterly and monthly, respectively. Jump indicators for the considered horizons are based on daily returns non-parametric jump detection at 1% significance level.

Figure 6: **Costly Learning in Constant Volatility Case**



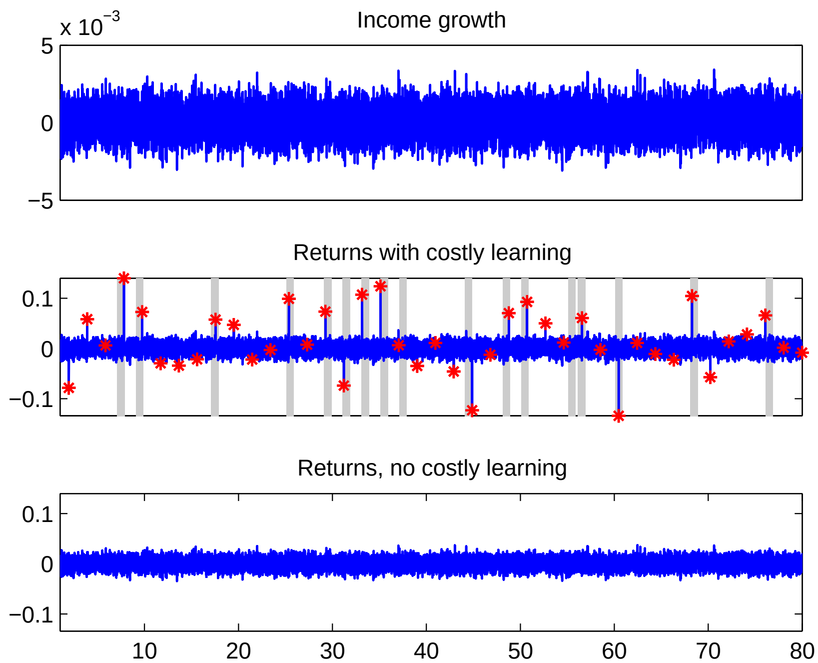
Optimal number of filtering periods, in years, in the constant volatility case for a range of information cost parameter χ. The risk aversion parameter γ is set at 15 (lower dotted line), 10 (middle solid line) and 5 (upper dashed line).

Figure 7: **Costly Learning in Time-Varying Volatility Case**



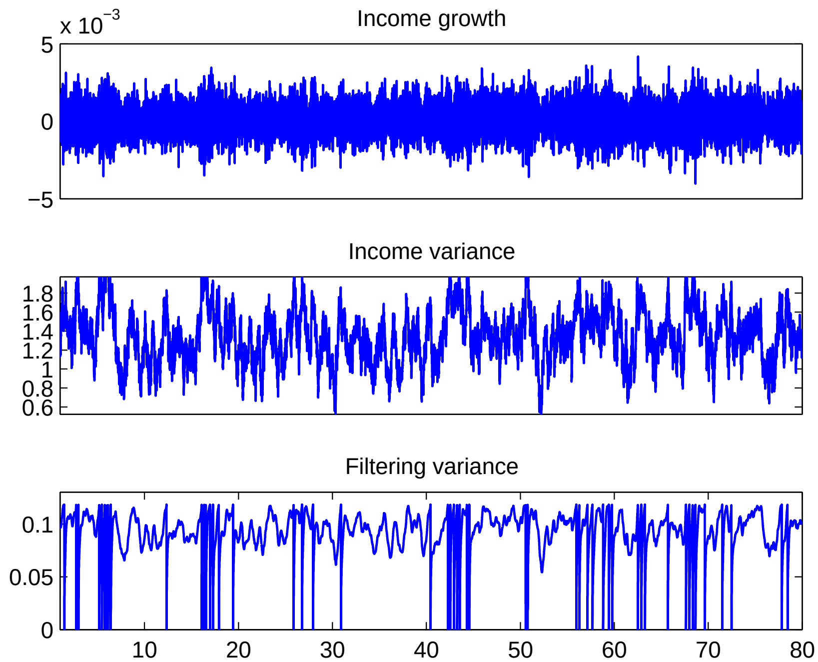
Expected number of periods till next costly learning, in years, in time-varying volatility case conditional on the level of filtering and income volatilities. Volatilities are annualized.

Figure 8: **Income and Return Simulation in Constant Volatility Case**



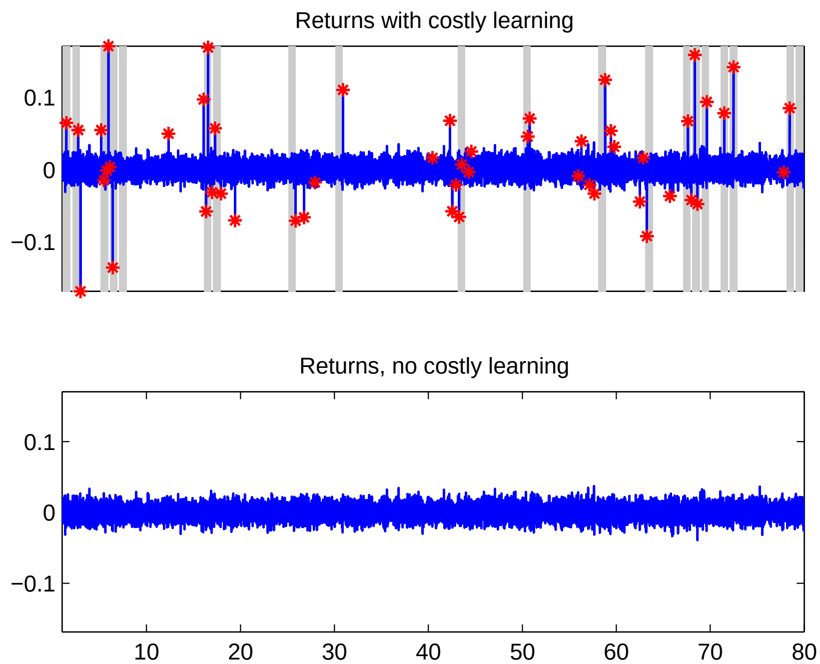
Simulation of the economy for 80 years in constant volatility case. Top panel depicts daily income growth. The next two panels show daily market returns when the agent learns about the true state for a cost, and with no option to learn, respectively. Red stars indicate days of learning, while grey regions correspond to flagged years with at least one significant jump using the jump-detection statistics.

Figure 9: **Income Simulation in Time-Varying Volatility Case**



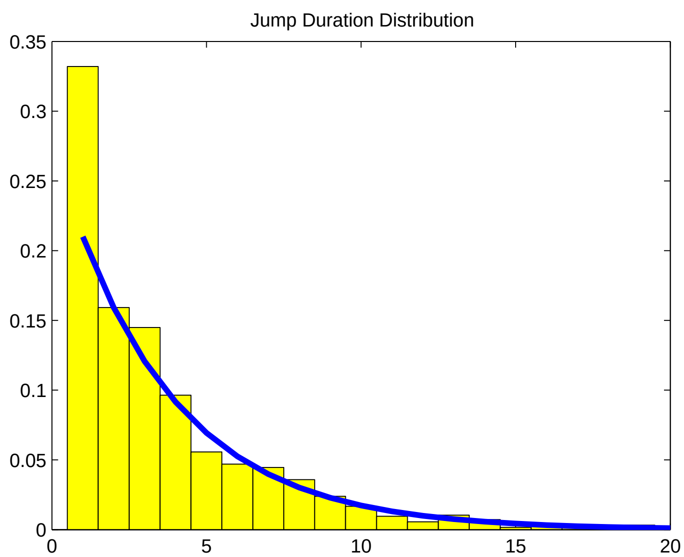
Simulation of the economy for 80 years in time-varying volatility case. Top panel depicts log income growth. The next two panels show income volatility and filtering volatility, annualized.

Figure 10: **Return Simulation in Time-Varying Volatility Case**



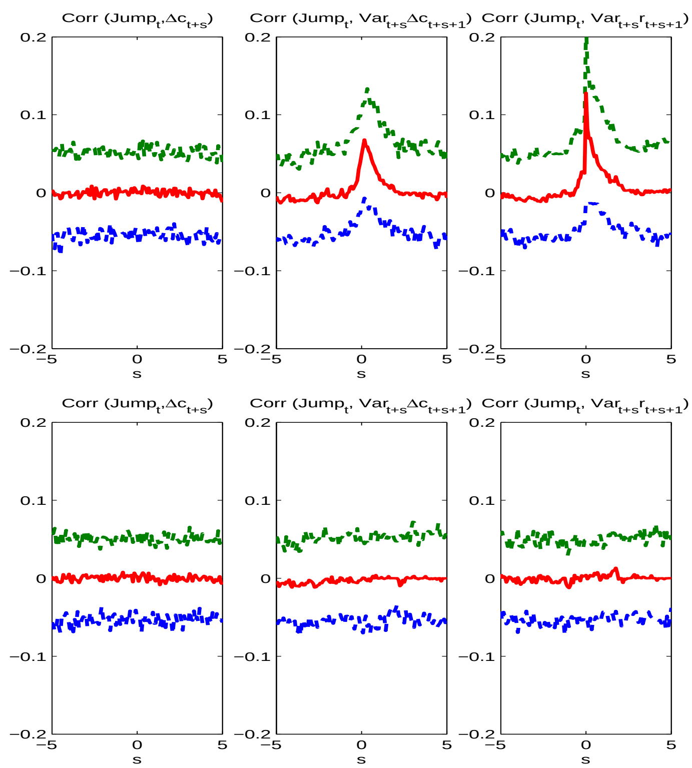
Simulation of the economy for 80 years in time-varying volatility case. The two panels show daily market returns when the agent learns about the true state for a cost, and with no option to learn, respectively. Red stars indicate days of learning, while grey regions correspond to flagged years with at least one significant jump using the jump-detection statistics.

Figure 11: **Model Frequency of Large Moves**



Model-implied distribution of time between years with large asset-price moves and an exponential distribution fit. Output is based on a long simulation from a time-varying volatility model.

Figure 12: **Model-Implied Jump Correlations**



Model-implied correlation of return jump indicator with endowment growth (left panel), conditional variance of endowment growth (middle panel) and conditional variance of returns (right panel), monthly frequency, at up to 5 year leads and lags based on the model with costly learning (top panel) and the model with no option to learn the true state for a cost (bottom panel). Model statistics are based on the average across 100 simulations of 80 years of data aggregated to monthly horizon; dashed lines show $5\% - 95\%$ confidence band. Jump indicator is based on daily returns non-parametric jump detection at 1% significance level.

References

Abel, A., J. Eberly, and S. Panageas, 2007a, Optimal inattention to the stock market in the presence of transaction costs, working paper.

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

Ai, Hengjie, 2007, Information quality and long-run risk: Asset pricing and welfare implications, working paper, Duke University.

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

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.

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.

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

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.

Gennotte, G., 1986, Optimal portfolio choice under incomplete information, *Journal of Finance* 41, 733–746.

Hansen, Lars, and Thomas Sargent, 2006, Fragile beliefs and the price of model uncertainty, working paper.

Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long-run risk, journal of Political Economy, forthcoming.

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

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.

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, 2006, 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.