Macro Announcement Premium and Risk Preferences

Hengjie Ai and Ravi Bansal *

May 4, 2016

Empirical evidence shows that a large fraction of equity premium is realized on a relatively small number of trading days with significant macroeconomic news announcements. In the 1961-2014 period, for example, about 55% of the entire equity premium is earned on about 30 trading days per year with significant macroeconomic announcements. In addition, the market equity premium typically rises prior to the announcement and falls immediately afterwards. In this paper, we develop an abstract theory and a quantitative model for the equity premium associated with macroeconomic news announcements. We demonstrate that the announcement premium identifies the compensation for investors' uncertainty aversion on capital markets. We present a dynamic model to account for the evolution of equity premium around macroeconomic announcements.

*Hengjie Ai (hengjie@umn.edu) is affiliated with the Carlson School of Management, University of Minnesota, and Ravi Bansal (ravi.bansal@duke.edu) is at the Fuqua School of Business, Duke University and NBER. We would like to thank seminar participants at Duke Finance workshop, HKUST, SAIF, and University of Texas, Austin for their helpful comments on the paper.

1 Introduction

Empirical evidence shows that a large fraction of equity premium is realized on a small number of trading days with significant macroeconomic news announcements. During the period of 1961-2014, for example, the cumulative excess return of the $SP500$ index on the thirty days per year with significant macroeconomic news announcements averages 3.36%, which accounts for 55% of the total annual equity premium during this period (7.50%). In this paper, we develop a theory for equity premium for macroeconomic news announcements.

We define pure news events as the arrival of news that affects asset prices but is not associated with an immediate impact on the real economy. The concept of pure news events is motivated by the fact that modern financial markets, especially with high-frequency trading, allow asset prices to react instantaneously to news before any real decisions (for example, consumption and investment) take place. Assuming rational expectation, surprises in news must be zero on average due to the law of iterated expectation. Does this imply that the response of asset prices to pure news events must average to zero? We study this question in a general equilibrium asset pricing model with pure news events.

Our main theorem is that the stochastic discount factor for pure news events can be interpreted as investors' probability distortion relative to rational expectation. This result has two implications. First, pure news events are not associated with the realization any equity premium if and only if investors have expected utility. Second, uncertainty aversion (in the sense of Gilboa and Schmeidler [19]) implies positive premia for pure news events on risky payoffs. Our framework assumes fairly weak regularity conditions on preferences and applies to most of the non-expected utility models proposed in the literature.

The equity premium in non-expected utility models typically contain two components, compensation for risk aversion (or risk premium) and compensation for uncertainty aversion (which can be labeled as uncertainty premium). Interpreted this way, our theoretical setup provides a thought experiment that separates risk aversion from uncertainty aversion using asset market data because it implies that any premium for pure news events must be compensation for uncertainty aversion.

We present a continuous-time model with learning to quantitatively account for the evolution of equity premium before, at, and after macroeconomic announcement. We model uncertainty aversion by the stochastic differential utility of Duffie and Epstein [12]. In our model, investors update their beliefs about the hidden state variables that govern the dynamics of aggregate consumption based both on their observations of the realizations of consumption and on pre-scheduled macroeconomic announcements. We establish several results in this environment. First, as in Breeden [7], because consumption follows a

continuous-time diffusion process, the equity premium investors receive in periods without news announcements is proportional to the length of the holding period of the asset. At the same time, a pure news event results in a non-trivial reduction of uncertainty, and is associated with the realization of a substantial amount of equity premium in an infinitesimally small window of time.

Second, equity premium typically increases before pure news events, peaks at the event, and drops sharply afterwards. The reduction of uncertainty right after pure news events implies a simultaneous decline in equity premium going forward. At the same time, after the current news announcement and before the next news announcement, because investors do not observe the movement in the hidden state variable, uncertainty slowly builds up over time, and so does the equity premium.

Third, we show that asset returns around pure news events impose a tight restriction on the preference parameters of the representative agent. As a result, the equity premium for macroeconomic announcements provides a market-based evidence on agent's preference for early resolution of uncertainty.

Our paper builds on the literature that studies decision making with non-expected utility and its asset pricing implications. We use the general representation of dynamic preferences of Strzalecki [41], and our framework includes most of the non-expected utility models appear in the literature as special cases. Some of the prominent examples are: the multiple prior model of Gilboa and Schmeidler [19] and Epstein and Schneider [16], the recursive preference of Kreps and Porteus [28] and Epstein and Zin [17], the robust control preference of Hansen and Sargent [21], the variational ambiguity averse preference of Maccheroni, Marinacci, and Rustichini [33, 34], and the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji [26, 27]. From the decision theoretical point of view, most of the above models satisfy Gilboa and Schmeidler [19]'s uncertainty aversion axiom and accounts for Ellsberg [14] type of behavior. We provide a necessary and sufficient condition for dynamic preferences to require an announcement premium. Our results imply that under mild technical condition, all uncertainty averse preferences are consistent with a positive premium for pure news events. For this perspective, our thought experiment can be interpreted as an asset market based axiom of uncertainty aversion.

Our empirical results are related to previous research on stock market returns on macroeconomic announcement days. Savor and Wilson [39] document that stock market returns and Sharpe ratios are significantly higher on days with macroeconomic news releases. Lucca and Moench [31] found similar patterns for equity returns on pre-scheduled FOMC announcement days. Cieslak, Morse, and Vissing-Jorgensen [11] provides an extensive study of the cyclical pattern of stock return around FOMC announcements. Rosoiu [37] found

similar patterns in high-frequency data. None of the above papers study a broad set of macroeconomic news and provide a theoretical foundation for equity premium for pure news events as we do.

Our framework also provides a simple litmus test for leading asset pricing models. The recursive utility based long-run risks model (Bansal and Yaron [4]), the robust control model (Hansen and Sargent [21]) and other non-expected utility based models are consistent with the empirical evidence of the positive average returns around news announcements, while the expected utility based habit model of Campbell and Cochrane [8] and the rare disaster model of Barro [5] are not.

The rest of the paper is organized as follows. We document some stylized facts for equity premium for macroeconomic news announcements in Section 2. In Section 3, we develop a theory for equity premium for pure news events in a simple two-period model. We present a continuous-time model in Section 4 to quantitatively account for the evolution of equity premium around macroeconomic news announcement days. Section 5 concludes.

2 Stylized facts

To demonstrate the significance of equity premium for macroeconomic announcements and to highlight the difference between news-announcement days and non-announcement days, we focus on a relatively small set of pre-scheduled macroeconomic news that are announced at the monthly or lower frequency. Within this category, we select the top five macroeconomic announcements ranked by investor attention by Bloomberg users. This procedure yields fifty announcement days per year in the 1997-2014 sample period where all news announcements are available, and on average thirty announcement days in the longer sample period of 1961-2014. We summarize our main findings below and provide the description of data construction and more details of the empirical evidence in Appendix A of the paper.

1. A large fraction of equity premium is realized on a relatively small number of trading days with pre-scheduled macroeconomic news announcements.

   As we show in Table 4, during the 1961-2014 period, there are on average thirty trading days per year with significant macroeconomic news announcements. The cumulative equity premium on the thirty news announcement days is 3.36%, which accounts for about 55% of the annual equity premium (6.19%) during this period. This pattern is even more pronounced if we focus on the period of 1997-2014, where more announcements are available. In this latter part of the sample, the market equity premium is 7.44% per year, while the cumulative excess return of the S&P500 index

on the fifty days with significant macroeconomic announcements is 8.24% on average. The equity premium on the rest of trading days is not statistically different from zero.

2. The equity premium increases before news announcements, peaks at the announcements, and drops immediately afterwards.

   The above pattern of the equity premium is shown in Figure 1, where we use the high-frequency data for S&P 500 SPDR during the 1997-2013 period from the TAQ database to construct thirty-minutes market returns around the pre-scheduled macroeconomic announcements.

3 Theoretical framework

In this section, we develop a theory for equity premium for pure news events. We first discuss several examples in a simple two-period model to illustrate conditions under which pure news events require a premium. We then provide a general theorem that characterizes the set of dynamic preferences that require a non-negative premium for pure news events.

3.1 A two-period model

We consider a representative agent economy with two periods (or dates), 0 and 1. There is no uncertainty in period 0, and the constant aggregate endowment in period 0 is denoted C_0. Aggregate endowment in period 1, denoted C_1 is a random variable.

Period 0 is further divided into two subperiods. In period 0^-, before any information about about C_1 is revealed, asset market opens, and a full set of Arrow-Debreu security is traded. The asset prices at this point will be called pre-news prices. For any asset with payoff X, we denote its pre-news price as $P^-(X)$. Note that $P^-(X)$ cannot depend on the realization of C_1, which is not known at this point.

In period 0^+, the true value C_1 is fully revealed to the agent. Immediately after the arrival of news, asset markets open again. Asset prices at this point are called post-news prices and are denoted as $P^+(X)$. In general, $P^+(X)$ depends on C_1, which is publicly known at 0^+. In period 1, the payoff of the Arrow-Debreu securities are realized and C_1 is consumed. We illustrate the timing of information and consumption in Figure 2 and that of the asset markets in Figure 3 assuming two states in period 1.

We define a pure news event to be the arrival of news that is not associated with an immediate impact on the allocation of real quantities and focus on the asset pricing implications of pure news events. The revelation of information about C_1 in our two-period

model captures two key properties of pure news events. First, it affects the conditional distribution of future consumption but rational expectation implies that surprises must average to zero in expectation. Second, news does not impact the current period consumption.

We believe that the concept of pure news events is the appropriate abstraction to study equity premium for macroeconomic announcements because it allows us to focus on the effect of macroeconomic news on asset prices without confounding this issue with the immediate impact of news on allocations. The latter is unlikely to be the key factor that determines equity premium realized upon news announcements for several reasons. First, modern financial markets allow asset prices to react immediately to news while the impact of news on consumption and investment takes time. As we show in Section 2, most of the equity premium associated with macroeconomic news is realized within the 30-minute interval of announcement. Real quantities react to news at much lower frequencies. Second, most of the equity premium are realized in the thirty or fifty trading days per year with significant macroeconomic announcements, while it is hard to imagine that most of the real economic activities are concentrated only on these small number of days as well. Third, by focusing on pure news events, we are effectively assuming that consumption does not respond to news in the current period. This becomes a very minimal restriction in our continuous-time model in Section 4, where the length of a period is infinitesimal.

The timing convention in the two-period model highlights the difference between standard consumption-based asset pricing models and our analysis. Standard consumption-based asset pricing models (for example, Lucas [30]) focus on the intertemporal substitution of consumption and its implications on returns received in period 1 after the action of consuming C_1 is completed. We study the reaction of asset prices to the pure news event at time 0^+ before any intertemporal consumption decisions takes place. Our purpose is to provide conditions on agents' intertemporal preferences under which pure news events require a premium.

The announcement return of an asset with payoff X is defined as the return of a strategy that long the asset right before the pre-scheduled pure news event and sell right afterwards. The expected announcement return of an asset with payoff X is therefore

$$ER(X) = \frac{E[P^+(X)]}{P^-(X)}. \tag{1}$$

We say that asset X requires a positive announcement premium if

$$ER(X) > 1. \tag{2}$$

3.2 Expected utility

We first consider the case in which the representative agent has expected utility:

$$u(C_0) + \beta E[u(C_1)],$$

where u is continuously differentiable. In this case, the pre-news price of an asset with payoff X is given by:

$$P^-(X) = E\left[\frac{\beta u'(C_1)}{u'(C_0)} X\right]. \tag{3}$$

In period 0^+, after the uncertainty is resolved, the agent's preference is represented by

$$u(C_0) + \beta u(C_1). \tag{4}$$

Therefore, for any C_1, the post-news price of the asset is

$$P^+(X) = \frac{\beta u'(C_1)}{u'(C_0)} X. \tag{5}$$

Clearly, the expected announcement return of any asset is

$$ER(X) = \frac{E[P^+(X)]}{P^-(X)} = 1. \tag{6}$$

There can be no announcement premium on any asset under expected utility.

3.3 An example with uncertainty aversion

Consider an agent with the constraint robust control preference of Hansen and Sargent [22]:

$$\begin{aligned} & u(C_0) + \beta \min_m E[mu(C_1)] \\ & subject\ to: \ E[m \ln m] \leq \eta \\ & E[m] = 1. \end{aligned} \tag{7}$$

The above expression can also be interpreted as the maxmin expected utility of Gilboa and Schmeidler [19]. The agent treats the reference probability measure, under which equity premium is evaluated (by econometricians) as an approximation. As a result, he takes into account of a class of alternative probability measures, represented by the density m, that are close to the reference probability measure. The inequality $E[m \ln m] \leq \eta$ requires that the relative entropy of the alternative probability models to be less than η.

In this case, the pre-news price of an asset with payoff X is:

$$P^{-}(X) = E\left[m^{*}\frac{\beta u'(C_1)}{u'(C_0)}X\right], \tag{8}$$

where m^{*} is the density of the minimizing probability for (7) and can be expressed as a function of $u(C_1)$, $m^{*} = m^{*} \circ u(C_1)$, where the m^{*} function is given by:

$$m^{*}(v) = \frac{e^{-\frac{1}{\theta}v}}{E\left[e^{-\frac{1}{\theta}v}\right]}. \tag{9}$$

The positive constant in the above expression, θ is determined by the binding relative entropy constraint $E[m^{*}\ln m^{*}] = \eta$.

In period 0^{+}, after the resolution of uncertainty, the agent's utility reduces to (4). As a result, the post-news price of the asset is the same as that in (5). Therefore, we can write the pre-news price as:

$$P^{-}(X) = E\left[m^{*}P^{+}(X)\right].^{1} \tag{10}$$

Because m^{*} is a decreasing function of date-1 utility $u(C_1)$, it is straightforward to prove the following claim:

Claim 1. *Suppose that the post-news price, $P^{+}(X)$ is a strictly increasing function of date-1 utility, $u(C_1)$, then $P^{-}(X) < E[P^{+}(X)]$. As a result, the announcement premium for the asset is strictly positive.*

The intuition of the above result is clear. After the pure news event, asset prices are discounted using marginal utilities. Under expected utility, the pre-news price is computed using probability weighted marginal utilities, therefore the pre-news price must equal expected post-news prices and there can be no announcement premium under rational expectation. Under the robust control preference, pre-news price is computed not using the reference probability, but with the pessimistic probability that over-weighs low utility states and under-weighs high utility states as shown in equation (9). As a result, uncertainty aversion applies an extra discounting for payoffs positively correlated with utility and therefore the asset market requires a premium for such payoffs relative to risk-free returns.

Under the robust control preference, the discounting is represented by the pessimistic probability density m^{*} that resembles the stochastic discount factor (SDF) in standard asset

[1] To minimize notation, we use m^{*} both as the probability density that prices post-news payoffs and as the real valued function of date-1 utility such that $m^{*} = m^{*} \circ u(C_1)$. Theorem 1 provides a sufficient condition under which the SDF for pure news events can always be represented as a measurable function of continuation utility.

pricing models but with a crucial difference. The SDF in standard asset pricing models reflects agents' intertemporal rate of substitution of consumption. In our model, there is no intertemporal consumption decision before and after the pure news event. The term m^* reflects investors' uncertainty aversion and identifies the probability distortion relative to rational expectation. As we show in the next section, the link between discounting with m^* and investors' uncertainty aversion holds under much more general conditions.

3.4 Uncertainty aversion and announcement premium

In this section, we provide a necessary and sufficient condition on preferences under which the announcement premium is positive. We consider a dynamic economy where the representative agent's preference has a recursive representation of the following form

$$V_t = u\left(C_t\right) + \beta \mathcal{I}\left[V_{t+1}\right], \tag{11}$$

where $\beta \in (0,1)$ and $\mathcal{I}$ is the certainty equivalence functional that maps the next period utility (which is a random variable) into its certainty equivalent (which is a real number). Our main result is that all risky assets require a non-negative announcement premium if and only if the certainty equivalence functional $\mathcal{I}[\cdot]$ is increasing in second-order stochastic dominance.

The representation (11) is very general. For example, the expected utility is the special case in which $\mathcal{I}[\cdot]$ is the expectation operator and the robust control preference corresponds to the specification where $\mathcal{I}[\cdot]$ is the robust control operator as in (7). More generally, Strzalecki [41] shows that (11) incorporates most of the dynamic uncertainty-averse preferences proposed in the literature.[2] Below we describe informally a general setup of a dynamic model with markets for pure news events and state the main theorems. We refer the reader to Appendix C for formal definitions and technical details.

A dynamic model with markets for pure news events Consider a representative agent, pure-exchange economy where time is finite and indexed by $t = 1, 2, \cdots, T$. The endowment process is denoted $\{Y_t, s_t\}_{t=1}^{T}$, where Y_t is aggregate endowment in period t and s_t is a signal publicly announced in period t after the realization of Y_t. The arrival of the

[2] As Strzalecki [41] shows, this representation includes the maxmin expected utility of Gilboa and Schmeidler [19], the second order expected utility of Ergin and Gul [18], the smooth ambiguity preferences of Klibanoff, Marinacci, and Mukerji [26], the variational preferences of Maccheroni, Marinacci, and Rustichini [33], the multiplier preferences of Hansen and Sargent [21] and Strzaleck [40], as well as the confidence preferences of Chateauneuf and Faro [9]. In addition, our setup is more general than that of Strzalecki [41], as we do not require the function $u(\cdot)$ to be affine. We provide expressions for the SDF for pure news events for all of the above mentioned decision models in Appendix C.

signal s_t is a pure news event because it carries information about future realizations of $Y_{t+1}, Y_{t+2}, \ldots Y_T$, but do not affect current-period endowment. In every period t, after the realization of Y_t but before the announcement of s_t, the time t^- asset market opens where a vector of $J+1$ returns are traded: $\left\{R^+_{j,t}(s_t)\right\}_{j=0,1,\cdots,J}$. Here $R^+_{j,t}$ are announcement returns that require one unit of investment at time t^- and provide a signal-contingent payoff at time t^+ after s_t is announced. At time t^+, after s_t is revealed, the intertemporal asset market opens and the agents trade a vector of $J+1$ returns with payoff contingent on the realization of next period endowment: $\left\{R^-_{j,t+1}(Y_{t+1})\right\}_{j=0,1,\cdots,J}$. To save notation, we suppress the dependence of $R^+_{j,t}(s_t)$ on s_t and the dependence of $R^-_{j,t+1}(Y_{t+1})$ on Y_{t+1} in what follows. We also adopt the convention that $R^+_{0,t}$ and $R^-_{0,t+1}$ are risk-free returns.

The optimal consumption-portfolio choice problem of an investor can be solved by backward induction. In the last period T, agents simply consume their total wealth, and therefore $V^-_T(W) = V^+_T(W) = u(W)$. For $t = 1, 2, \cdots, T-1$,we denote $\xi = [\xi_0, \xi_1, \xi_2, \cdots \xi_J]$ as the vector of the investment in the intertemporal asset market and write the corresponding consumption-portfolio choice problem as:

$$V^+_t\left(W^+_t\right) = \max_{C,\xi}\left\{u(C) + \beta\mathcal{I}\left[V^-_{t+1}\left(W^-_{t+1}\right)\right]\right\}$$
$$C + \sum\nolimits_{j=0}^{J}\xi_j = W^+_t$$
$$W^-_{t+1} = \sum\nolimits_{j=0}^{J}\xi_j R^-_{j,t+1}.$$

Similarly, the optimal portfolio choice problem on the markets for pure news events is

$$V^-_t\left(W^-_t\right) = \max_{\zeta}\mathcal{I}\left[V^+_t\left(W^+_t\right)\right]$$
$$W^+_t = W^-_t - \sum\nolimits_{j=0}^{J}\zeta_j + \sum\nolimits_{j=0}^{J}\zeta_j R^+_{j,t},$$

where $\zeta = [\zeta_0, \zeta_1, \zeta_2, \cdots \zeta_J]$ is a vector of investment in announcement returns.

We assume that for some initial wealth level, W_0 and a sequence of returns $\left\{\left\{R^+_{j,t}, R^-_{j,t+1}\right\}_{j=0,1,\cdots,J}\right\}_{t=1,2,\cdots T}$, an interior competitive equilibrium with sequential trading exists where all markets clear. We focus on the announcement premium implied by the property of the certainty equivalence functional $\mathcal{I}[\cdot]$.

Characterization of the announcement premium We first introduce some terminologies and notations. We use $\mathcal{F}^-_t$ to denote all the information available at time t^- before s_t is revealed, and $\mathcal{F}^+_t$ to denote all the information available at time t after the announcement of s_t. We assume all uncertainty in the economy is generated from some

nonatomic probability space $(\Omega, \mathcal{B}, P)$, where $\mathcal{B}$ is the associated Borel $\sigma-$algebra. We denote $L^2(\Omega, \mathcal{B}, P)$ as the set of random variables defined on $(\Omega, \mathcal{B}, P)$ with finite second moments. We say that $\mathcal{I}[\cdot]$ is monotone with respect to first (second) order stochastic dominance if $\mathcal{I}[V_1] \geq \mathcal{I}[V_2]$ whenever V_1 first (second) order stochastic dominates V_2. $\mathcal{I}[\cdot]$ is said to be strictly monotone with respect to first (second) order stochastic dominance if in addition, $\mathcal{I}[V_1] > \mathcal{I}[V_2]$ whenever V_1 strictly first (second) order stochastic dominates V_2. We first provide a theorem for the existence of SDF for pure news events.

Theorem 1. *(Existence of SDF for Pure-News-Events)*

Suppose both u and $\mathcal{I}$ are Lipschitz continuous, Fréchet differentiable with Lipschitz continuous derivatives.[3] *Suppose that $\mathcal{I}$ is strictly monotone with respect to first order stochastic dominance, then in any interior competitive equilibrium with sequential trading, $\forall t$, the risk-free announcement return $R^+_{0,t} = 1$. In addition, there exists a non-negative measurable function $m^*_t : \mathbf{R} \to \mathbf{R}$ such that*

$$E\left[m^*_t\left(V^+_t\right)\left\{R^+_{j,t}(s_t) - 1\right\}\middle|\, \mathcal{F}^-_t\right] = 1 \quad for \quad all \quad j = 1, 2, \cdots J. \tag{12}$$

Under the regularity condition (48) in Appendix C, $E\left[m^*_t\left(V^+_t\right)\middle|\, \mathcal{F}^-_t\right] = 1$ and (12) can be written as:

$$E\left[m^*_t\left(V^+_t\right) R^+_{j,t}(s_t)\middle|\, \mathcal{F}^-_t\right] = 1 \quad for \quad all \quad j = 0, 1, 2, \cdots J. \tag{13}$$

Intuitively, Theorem 1 provides sufficient differentiability conditions on the certainty equivalence functional $\mathcal{I}$ so that SDF can be constructed from marginal utilities. Under these conditions, Theorem 2 below links announcement premium to the properties of the certainty equivalence functional.

Theorem 2. *(Announcement Premium) Under the assumptions of Theorem 1,*

1. *$m^*_t\left(V^+_t\right) = 1$ for all V^+_t if and only $\mathcal{I}$ is the expectation operator.*
2. *$m^*_t\left(V^+_t\right)$ is a non-increasing function of continuation utility V^+_t if and only if the certainty equivalence functional $\mathcal{I}$ satisfies monotonicity with respect to SSD.*

By Theorem 2, m^* under expected utility is trivial: pure news events do not require any discounting. In general, the above theorem links the sign of the announcement premium to the properties of the certainty equivalence functional: monotonicity with respect to second order stochastic dominance is equivalent to the fact that m^*_t is non-increasing in continuation

[3] We use the notion of Fréchet differentianility because $\mathcal{I}$ is a nonlinear functional defined on the infinite dimensional space $L^2(\Omega, \mathcal{B}, P)$. We provide the details of the definition of Lipschitz continuity and Fréchet differentiability in Appendix C.

utility V_t^+. Because m_t^* is a probability density, this can be interpreted as pessimism: it overweights states with higher continuation utility and underweighs states with lower continuation utility. Because m_t^* prices assets on the market for pure news events, pessimism implies a premium for payoffs that are increasing in continuation utility. To see this point, let $f\left(V_t^+\right)$ be any announcement payoff that is increasing in V_t^+, then its pre-news price $P_t^-\left(X\right) = E\left[m_t^*\left(V_t^+\right) f\left(V_t^+\right) \middle| \mathcal{F}_t^-\right] \leq E\left[f\left(V_t^+\right) \middle| \mathcal{F}_t^-\right]$ because $m_t^*\left(V_t^+\right)$ is decreasing in V_t^+ and $f\left(V_t^+\right)$ is increasing in V_t^+. As a result,

$$\frac{E\left[f\left(V_t^+\right) \middle| \mathcal{F}_t^-\right]}{P_t^-\left(X\right)} \geq 1 = R_{0,t}^+,$$

where $R_{0,t}^+$ is the risk-free announcement return. In fact, pessimism of m_t^* is equivalent to a non-negative premium for all risky payoffs (See Lemma 7 in Appendix B). The condition of monotonicity with respect to SSD may be hard to verify in practice. Below we provide a sufficient condition for this property.

Remark 1. *Suppose $\mathcal{I}$ is monotone with respect to first order stochastic dominance, then quasi-concavity implies monotonicity with respective to second order stochastic dominance.*

The quasi-concavity condition is particularly interesting because in the special case where $u\left(\cdot\right)$ is affine as in Strzalecki [41], then quasi-concavity is equivalent to the uncertainty aversion axiom of Gilboa and Schmeidler [19]. An immediate consequence of the above remark is that all uncertainty-averse preferences require a non-negative announcement premium. As a result, the empirical evidence on announcement premium can be interpreted as providing a market-based evidence for uncertainty aversion.

3.5 Discussion of Theorem 2

Compensation for uncertainty aversion It is clear from Theorem 2 that the property of the SDF for pure news events, m^* depends only on the curvature of the certainty equivalence functional $\mathcal{I}\left[\cdot\right]$ and not on the property of the utility function u. This is in sharp contrast with the SDF arising from expected utility-based asset pricing models, which reflect agent's attitude toward risk and depends on the curvature of u. Because the quasi-concavity of $\mathcal{I}\left[\cdot\right]$ reflects agent aversion toward uncertainty, announcement premium separates compensation for uncertainty aversion from compensation for risk aversion.

Theorem 1 and 2 highlight the difference in the timing of the realization of the risk premium due to the curvature of the von Neumann-Morgenstern utility u and that of the announcement premium due to the curvature of the certainty equivalence functional $\mathcal{I}$. The risk premium due to the curvature of u realizes after the action of consumption is completed,

while investors receive the announcement premium right after the revelation of news and before consumption takes place. In the case of expect utility, it is well-known that the non-negativity of risk premium is equivalent to the concavity of u. If we view $E[u(\cdot)]$ as a mapping from the space of random variables to the real line, then the concavity of u is equivalent to $E[u(\cdot)]$ being non-decreasing in second order stochastic dominance. Theorem 2 provides an analagous result: a non-negative announcement premium is equivalent to $\mathcal{I}$ being monotone with respect to second order stochastic dominance.

Theorem 2 applies to a large class of non-expected utility models that exhibits uncertainty aversion. Below we provide another example of m^* under the smooth ambiguity averse preference of Klibanoff, Marinacci, and Mukerji [26, 27] in the two-period setup that we describe in Section 3.1. In general, the smooth ambiguity model can be represented in the form of (11) with the following choice of the certainty equivalence functional:

$$\mathcal{I}[V] = \int_{\Delta} \phi\left(E^x[V]\right) d\mu(x). \tag{14}$$

We use Δ to denote a set of probability measures indexed by x, P_x. The notation $E^x[\cdot]$ stands for expectation under the probability P_x and $\mu(x)$ is a probability measure over x. As shown by Klibanoff, Marinacci, and Mukerji [26, 27], the concavity of ϕ corresponds to uncertainty aversion. We consider the following parameteric form of u and ϕ: $u(C) = \frac{1}{1-\gamma}C^{1-\gamma}$, $\phi(V) = \frac{1}{1-\eta}\left[(1-\gamma)V\right]^{\frac{1-\eta}{1-\gamma}}$. Under this specification, concavity of ϕ corresponds to the case of $\eta \geq \gamma$.

Example 1. *The smooth ambiguity-averse preference*

To complete the decription of the certainty equivalent functional in (14), we need to specify $\{P_x\}_{x\in\Delta}$ and a probability measure, μ on Δ. We assume that under P_x, $\ln C_1$ is normally distributed with mean $\mu + x$ and variance σ^2, and $d\mu(x) = \frac{1}{\sqrt{2\pi}\tau}e^{-\frac{1}{2\tau^2}x^2}$. Also, we assume that the reference probability that econometricians use to compute equity premium is $P = \int P_x d\mu(x)$. In this case, the SDF for pure news events are given by:

$$m^* = \frac{\left[(1-\gamma)V\right]^{\frac{\gamma-\eta}{1-\gamma}\frac{\tau^2}{\sigma^2+\tau^2}}}{E\left[\left[(1-\gamma)V\right]^{\frac{\gamma-\eta}{1-\gamma}\frac{\tau^2}{\sigma^2+\tau^2}}\right]}.$$

Clearly, m^* assigns higher weights to low utility states under the assumption of uncertainty aversion, $\eta > \gamma$.

Bounds on probability distortions Theorem 2 implies that the SDF for pure news events is a density and can be interpreted as probability distortions. Just like intertemporal

asset returns provide restrictions on the SDF that prices these returns, announcement returns provide restrictions on the magnitude of the probability distortions. For example, the standard deviation of m_t^* must be higher than the maximum Sharpe ratio of any announcement return, as in Hansen and Jagannathan [20]. That is, for an arbitrary announcement return, R_t^+, we have:

$$\sigma_t\left[m_t^*\right] \geq \frac{E_t\left[R_t^+ - 1\right]}{\sigma_t\left[R_t^+\right]}, \tag{15}$$

where $\sigma_t[\cdot]$ denote the conditional standard deviation and $E_t[\cdot]$ is the conditional expectation with respect to $\mathcal{F}_t^-$. We can also bound the entropy of the SDF for pure news events as in Bansal and Lehmann [3]:

$$-E_t\left[\ln m_t^*\right] \geq E_t\left[\ln R_t^+\right]. \tag{16}$$

The SDF (for intertemporal cash flows) in non-expected utility based asset pricing models can typically be written in the form of $m^* \frac{u'(C_{t+1})}{u'(C_t)}$. Inequalities (15) and (16) therefore provide a guidance on the relative importance of the two components, m^* and $\frac{u'(C_{t+1})}{u'(C_t)}$. Using the moments of the market return during the period of 1961-2014 in Table 4, we can compute the Hansen-Jaganathan bound for SDF as $\sigma_t\left[m_t^* \frac{u'(C_{t+1})}{u'(C_t)}\right] \geq 40\%$. The Hansen-Jaganathan bound for the probability distortion component is tighter: $\sigma_t\left[m_t^*\right] \geq 55\%$. In the case of the entropy bound, $-E_t\left[\ln\left(m_t^* \frac{u'(C_{t+1})}{u'(C_t)}\right)\right] = -E_t\left[\ln m_t^*\right] - E_t\left[\ln\left(\frac{u'(C_{t+1})}{u'(C_t)}\right)\right]$. Therefore $-E_t\left[\ln\left(m_t^* \frac{u'(C_{t+1})}{u'(C_t)}\right)\right] \geq 6.19\%$ if we use the average market return during 1961-2014 to replace the right-hand-side of (16) and the bound on the probability distortion component is $-E_t\left[\ln m_t^*\right] \geq 3.36\%$ using the average announcement return during the same period. In both cases, the magnitude of announcement return implies a substantial amount of probability distortion.

Preference for Early Resolution of Uncertainty If we focus on recursive preferences with constant risk aversion and constant IES, then the monotonicity of $\mathcal{I}$ with respect to second order stochastic dominance is equivalent to preference for early resolution of uncertainty, as we show below.[4]

Example 2. *The Kreps-Porteus utility*

Here we show that the Kreps-Porteus utility requires a positive announcement premium if and only if it prefers early resolution of uncertainty. As we show in Appendix C, the Kreps-

[4]In general, however, preference for early resolution of uncertainty is neither sufficient nor necessary for a positive premium for pure news events. For example, the maxmin expected utility of Gilboa and Schmeidler [19] requires an announcement premium, but it is indifferent toward the timing of resolution of uncertainty, as shown in Strzalecki [41].

Porteus utility can be represented in the form of (11) with the following choices of u and $\mathcal{I}$:

$$u(C) = \frac{1}{1-\frac{1}{\psi}} C^{1-\frac{1}{\psi}}, \quad \mathcal{I}(V) = \frac{1}{1-\frac{1}{\psi}} \left\{ E\left[\left(\left(1-\frac{1}{\psi}\right) V \right)^{\frac{1-\gamma}{1-1/\psi}} \right] \right\}^{\frac{1-1/\psi}{1-\gamma}}. \tag{17}$$

It is straightforward to show that $\mathcal{I}$ is quasi-concave and therefore requires a positive announcement premium if and only if $\gamma \geq \frac{1}{\psi}$.

Assuming Kreps-Porteus utility, then the empirical evidence on announcement premium places a strong restriction on the agent's attitude toward the timing of resolution of uncertainty governed by the parameter $\gamma - \frac{1}{\psi}$. As we show in Appendix C, if we assume that the announcement return on aggregate wealth is log-normally distributed, then

$$\ln E_t\left[R^+_{W,t}\right] = \frac{\gamma - \frac{1}{\psi}}{1-\frac{1}{\psi}} Var_t\left[\ln R^+_{W,t}\right], \tag{18}$$

where $R^+_{W,t}$ denotes the announcement return on the claim to aggregate consumption. That is, moments in announcement returns can be used to inform the preference parameters of recursive utility.

Time-Non-Separable Preferences Note that our representation of preferences, (11) does not include some of the time-non-separable preferences, for example, the habit preference. Here we discuss some examples of time-non-separable preferences. We show that the external habit model does not require an announcement premium and the internal habit model generates a negative announcement premium.

Example 3. *(External habit: zero premium for pure news events)*

The external habit preference (for example, Campbell and Cochrane [8]) can be written as:

$$E\left[\sum_{t=0}^{T} \frac{1}{1-\gamma} (C_t - H_t)^{1-\gamma}\right],$$

where $\{H_t\}_{t=0}^{\infty}$ is the habit process. Consider the date-0 market for pure news events. The pre-news price of any payoff $\{X_t\}_{t=1}^{\infty}$ is

$$P_0^- = E\left[\sum_{t=0}^{T} \left(\frac{C_t - H_t}{C_0 - H_0}\right)^{-\gamma} X_t\right],$$

and the post-news price (after s_0 is announced) is

$$P_0^+ = E\left[\sum_{t=0}^{T}\left(\frac{C_t - H_t}{C_0 - H_0}\right)^{-\gamma} X_t \middle| s_0\right].$$

Clearly, as in the case of expected utility, $P_0^- = E\left[P_0^+\right]$ and there can be no announcement premium. The external habit model is similar to expected utility because the habit process is exogenous and agent do not take into account of the effect of current consumption choices on future habits when making consumption and investment decisions. The non-separability of preferences do affect consumption-investment choices in the internal habit model. However, as we show in Appendix C of the paper that the internal habit model of Boldrin, Christiano, and Fisher [6] generates a negative announcement premium.

4 Quantitative Results

In this section, we present a continuous-time model with the Kreps-Porteus utility and show that our model can quantitatively account for the dynamic pattern of announcement premium in the data. In our model, shocks to aggregate consumption are modeled as Brownian motions and arrive gradually over time. This setup allows us to distinguishes the announcement premium that is realized instantaneously upon pure news events and the risk premium that investors receive incrementally as shocks to consumption materialize slowly over time.

4.1 Model setup

The dynamics of consumption and dividends We consider a continuous-time representative agent economy. The growth rate of aggregate consumption during an infinitesimal interval Δ is specified as:

$$\ln C_{t+\Delta} - \ln C_t = x_t\Delta + \sigma\left(B_{C,t+\Delta} - B_{C,t}\right),$$

where x_t is a continuous time AR(1) process that is not observable to the agents of the economy. The law of motion of x_t is

$$x_{t+\Delta} = a_x\Delta\bar{x} + (1 - a\Delta)\, x_t + \sigma_x\left(B_{x,t+\Delta} - B_{x,t}\right),$$

where $\bar{x}$ is the long-run average growth rate of the economy, and $B_{C,t+\Delta} + B_{C,t}$ and $B_{x,t+\Delta} - B_{x,t}$ are normally distributed innovations (increments of Brownian motions). We

assume that $B_{C,t}$ and $B_{x,t}$ are independent. In the continuous-time notation,

$$\begin{aligned} \frac{dC_t}{C_t} &= x_t dt + \sigma dB_{C,t}, \\ dx_t &= a_x (\bar{x} - x_t) dt + \sigma_x dB_{x,t} \end{aligned} \tag{19}$$

Our benchmark asset is a claim to the following dividend process:

$$\frac{dD_t}{D_t} = [\bar{x} + \phi (x_t - \bar{x})] dt + \phi \sigma dB_{C,t}, \tag{20}$$

where we assume that the long-run average growth rate of consumption and dividend are the same, and we allow the leverage parameter $\phi > 1$ so that dividends are more risky than consumption as in Bansal and Yaron [4].

Timing of Information and Bayesian Learning There are two sources of information that the representative agent in the economy can use to update her beliefs about x_t. First, the realized consumption path contains information about x_t, and second, an additional signal about x_t is revealed at pre-scheduled discrete time points $T, 2T, 3T, \cdots$. For $n = 1, 2, 3, \cdots$, we denote the signal observed at time nT as s_n and assume $s_n = x_{nT} + \varepsilon_n$, where ε_n is normally distributed with mean zero and variance σ_S^2. Note that announcements of the signals at $t = T$, $2T$, $3T$, $\cdots$ are pure news events because they are not associated with the realization of any consumption shocks.

In the interior of $(0, T)$, the agent does not observe the true value of x_t and update her belief about x_t based on the observed consumption process according to Bayes' rule. We define $\hat{x}_t = E(x_t | C^t)$ to be the posterior mean of x_t and time t, and define $q_t = E\left[(x_t - \hat{x}_t)^2 \middle| C^t\right]$ to be the posterior variance of x_t.[5] The dynamics of x_t can be written as (Kalman-Bucy filter):

$$d\hat{x}_t = a [\bar{x} - \hat{x}_t] dt + \frac{q(t)}{\sigma} d\tilde{B}_{C,t}, \tag{21}$$

where the innovation process, $\tilde{B}_{C,t}$ is defined by:

$$d\tilde{B}_{C,t} = \frac{1}{\sigma} \left[\frac{dC_t}{C_t} - \hat{x}_t dt \right].$$

The posterior variance satisfies the Riccati equation:

$$dq(t) = \left[\sigma_x^2 - 2a_x q(t) - \frac{1}{\sigma^2} q^2(t) \right] dt. \tag{22}$$

[5]Here we use the notation C^t to denote the history of consumption up to time t.

The posterior distribution are updated immediately following the announcement of signals at time $T, 2T, \cdots$. At time T for example, agents update their beliefs using Bayes rule:

$$\hat{x}_T^+ = \frac{1}{q_T^+}\left[\frac{1}{\sigma_S^2}s + \frac{1}{q_T^-}\hat{x}_T^-\right]; \quad \frac{1}{q_T^+} = \frac{1}{\sigma_S^2} + \frac{1}{q_T^-}. \tag{23}$$

where s is the signal observed at time T, $\hat{x}_T^-$ and q_T^- are the posterior mean and variance of x_T before the announcement, and $\hat{x}_T^+$ and q_T^+ denote the posterior mean and variance of x_T after the announcement at time T. We plot the dynamics of posterior variance q_t in Figure 4 assuming that announcements are made every thirty days and they completely reveal the information about x_t, that is, $\sigma_S^2 = 0$. Upon announcements, the posterior variance drops immediately to zero, as indicated by the circles. After announcement, information arrives slowly and as a result, the posterior variance increases gradually over time before the next announcement.

Preferences and the stochastic discount factor We assume that the representative agent has a Kreps-Porteus utility with preference for early resolution of uncertainty, $\gamma > \frac{1}{\psi}$. In continuous-time, this preference is represented by the stochastic differential utility of Duffie and Epstein [12], which can be interpreted as the limit of the recursive relationship (11) over a small time interval Δ as $\Delta \to 0$:

$$V_t = \left(1 - e^{-\rho\Delta}\right) u\left(C_t\right) + e^{-\rho\Delta}\mathcal{I}\left[V_{t+\Delta} \middle| \hat{x}_t, q_t\right], \tag{24}$$

where ρ is the time discount rate, and $\mathcal{I}\left[\cdot \middle| \hat{x}_t, q_t\right]$ is the certainty equivalence functional conditioning on agents' posterior belief at time t, $(\hat{x}_t, q_t)$. To derive closed-form solutions, we focus on the case in which $\psi = 1$. The corresponding choice of utility function u and certainty equivalence functional $\mathcal{I}$ are: $u\left(C\right) = \ln C$ and $\mathcal{I}\left[V\right] = \frac{1}{1-\gamma}\ln E\left[e^{(1-\gamma)V}\right]$. This preference can also be interpreted as the multiplier robust control preference of Hansen and Sargent [21].

As well-known for representative agent economies with recursive preferences, any return $R_{t,t+\Delta}$ must satisfy the intertemporal Euler equation $E\left[SDF_{t,t+\Delta}R_{t,t+\Delta}\right] = 1$, where the stochastic discount factor for the time interval $(t, t+\Delta)$ given by

$$SDF_{t,t+\Delta} = e^{-\rho\Delta}\left(\frac{C_{t+\Delta}}{C_t}\right)^{-1}\frac{e^{(1-\gamma)V_{t+\Delta}}}{E_t\left[e^{(1-\gamma)V_{t+\Delta}}\right]}. \tag{25}$$

We can define $m_{t+\Delta}^* = \frac{e^{(1-\gamma)V_{t+\Delta}}}{E_t\left[e^{(1-\gamma)V_{t+\Delta}}\right]}$, which represents the probability distortion under the robust control interpretation of the model. With the above specification of preferences and

consumption, the value function has a closed-form solution, which we denote $V_t = V(\hat{x}_t, t, C_t)$, and

$$V(\hat{x}, t, C) = \frac{1}{a_x + \rho}\hat{x} + \frac{1}{1-\gamma}h(t) + \ln C, \tag{26}$$

where the function $h(t)$ is given in the appendix of the paper. We set the parameter values of our model to be consistent with standard long-run risk calibrations and list them in Table 1. We now turn to the quantitative implications of the model on the announcement premium.

4.2 The announcement premium

Equity premium on non-announcement days In the interior of $(nT, (n+1)T)$, the equity premium can be calculated using the SDF (25) as in standard learning models (for example, Veronesi [42] and Ai [2]). Let p_t denote the price-to-dividend ratio of the benchmark asset, and let r_t denote the risk-free interest rate at time t. Using a log-linear approximation of p_t, the equity premium for the benchmark asset over a small time interval $(t, t+\Delta)$ can be written as:[6]

$$E_t\left[\frac{p_{t+\Delta}D_{t+\Delta} + D_{t+\Delta}\Delta}{p_t D_t} - e^{r_t\Delta}\right] \approx \left[\gamma\sigma + \frac{\gamma-1}{a_x+\rho}\frac{q_t}{\sigma}\right]\left[\phi\sigma + \frac{\phi-1}{a_x+\rho}\frac{q_t}{\sigma}\right]\Delta. \tag{27}$$

The above expression has intuitive interpretations. The term $\left[\gamma\sigma + \frac{\gamma-1}{a_x+\rho}\frac{q_t}{\sigma}\right]$ is the sensitivity of the SDF with respect to consumption shocks, where $\gamma\sigma$ is market price of risk due to risk aversion, and the term $\frac{\gamma-1}{a_x+\rho}\frac{q_t}{\sigma}$ is associated with recursive utility and aversion to long-run risks. The second term $\left[\sigma + \frac{\phi-1}{a_x+\rho}\frac{q_t}{\sigma}\right]$ is the elasticity of asset return with respect to shocks in consumption, where $\phi\sigma$ is the sensitivity of dividend growth with respect to consumption growth, and $\frac{\phi-1}{a_x+\rho}\frac{q_t}{\sigma}$ is the response of price-dividend ratio to investors' belief about future consumption growth.

As we show in Figure 4, after the previous announcement and before the next announcement, because investors do not observe the true value of x_t, the posterior variance q_t increases over time. By equation (27), the equity premium also rises with q_t. In Figure 5, we plot the instantaneous equity premium $\left[\gamma\sigma + \frac{\gamma-1}{a_x+\rho}\frac{q_t}{\sigma}\right]\left[\phi\sigma + \frac{\phi-1}{a_x+\rho}\frac{q_t}{\sigma}\right]$ on non-announcement days. Clearly, the equity premium increases over time. This feature of our model captures the "pre-announcement drift" documented in empirical work.

As in typical continuous-time models, the amount of equity premium is proportional to the

[6]The equity premium depends on the state variable $\hat{x}$ in general. The log-linear approximation does not capture this dependence. We use the log-linear approximation to illustrate the intuition of the model. All figures and calibration results are obtained based on the global solution of the PDE obtained from the Markov chain approximation method.

length of holding period, Δ by equation (27). In fact, $SDF_{t,t+\Delta} \to 1$ as $\Delta \to 0$. Intuitively, the amount of risk diminishes to zero in an infinitesimally small time interval, and so does risk premium. The situation is very different on macroeconomic announcement days, which we turn to next.

Announcement returns Consider the stochastic discount factor in (25). As $\Delta \to 0$, the term $e^{-\rho\Delta}\left(\frac{C_{t+\Delta}}{C_t}\right)^{-\frac{1}{\psi}} \to 1$; however, $m^*_{t+\Delta}$ does not necessarily collapse to 1 unless the term V_t is a continuous function of t. We define $m^*_t = \lim_{\Delta\to 0^+} m^*_{t+\Delta}$. In periods with no news announcement, because consumption evolves continuously, so does the posterior belief $(\hat{x}_t, q_t)$. Therefore, $m^*_t = 1$. In contrast, at time $0, T, 2T, \cdots$, upon news announcement, $\hat{x}_t$ and q_t updates instantaneously, and

$$m^*_T = \frac{e^{\frac{1-\gamma}{a_x+\rho}\hat{x}^+_T}}{E\left[e^{\frac{1-\gamma}{a_x+\rho}\hat{x}^+_T}\middle|\, \hat{x}^-_T, q^-_T\right]}. \tag{28}$$

is the SDF for the pure news event at time T.

We focus on asset prices and returns around pure news events. We denote p^-_T as pre-news price-to-dividend ratio and p^+_T as the post-news price-to-dividend ratio at time T. No arbitrage implies that for all Δ,

$$p^-_T D_T = E^-_T\left[\int_0^\Delta SDF_{T,T+s}D_{T+s}ds + SDF_{T,T+\Delta}p_{t+\Delta}D_{T+\Delta}\right],$$

where E^-_T denote the expectation taken with respect to the information at time T before announcement. Taking the limit as $\Delta \to 0$, and use the fact that all terms in the SDF other than $m^*_{T+\Delta}$ collapse to 1, we have:

$$p^-_T = E^-_T\left[m^*_T \times p^+_T\right], \tag{29}$$

which is the continuous-time version of the asset pricing equation (10). As we show in Appendix D, using a first-order approximation, the announcement premium can be written as:

$$\ln\frac{E^-_T\left[p^+_T\right]}{p^-_T} \approx \frac{\gamma-1}{a_x+\rho}\frac{\phi-1}{a_x+\rho}\left(q^-_T - q^+_T\right). \tag{30}$$

We make the following observations.

1. In contrast to non-announcement periods, the equity premium does not disappear as $\Delta \to 0$. As long as $\phi > 1$ and the agent prefers early resolution of uncertainty ($\gamma > 1$), the announcement premium is positive.

In Figure 6, we plot the daily equity premium implied by our model with recursive utility (top panel) and that in a model with expected utility (bottom panel). Because we are calculating equity premium over a short time interval, the equity premium in non-announcement periods is negligible compared with announcement returns. The announcement premium is about 17 basis points in the top panel, while the equity premium is close to zero on announcement days. The quantitative magnitude of these returns is quite similar to their empirical counterpart in Table 4. Consistent with our theoretical results in Section 3, the announcement premium for expected utility is zero. The premium on announcement days for expected utility is 0.028 basis point, several orders of magnitude smaller than that for recursive utility.[7]

2. The announcement premium separates the probability distortion component of the SDF from compensation for risk aversion.

   The SDF in equation (25) has two components, the term $e^{-\rho\Delta}\left(\frac{C_{t+\Delta}}{C_t}\right)^{-1}$ that arises from standard log utility, and the term $\frac{e^{(1-\gamma)V_{t+\Delta}}}{E_t\left[e^{(1-\gamma)V_{t+\Delta}}\right]}$ that can be interpreted as probability distortion. As we take the limit $\Delta \to 0$, the intertemporal substitution of consumption term vanishes, and the SDF for pure news events, (28) depends only on probability distortion. In the log-linear approximation (30), the term $\frac{\gamma-1}{a_x+\rho}$ is due to investors' probability distortion with respect to x_T^+, and the term $\frac{\phi-1}{a_x+\rho}$ measures the sensitivity of price-to-dividend ratio with respect to probability distortions.

3. The magnitude of announcement premium is proportional to the variance reduction about the hidden state variable x_t upon the news announcement, $q_T^- - q_T^+$. The higher is the information content in news, the larger is the announcement premium. In addition, the announcement premium also increases with the persistence of the shocks. As the mean reversion parameter a_x gets smaller, the half life of the impact of x_t on consumption increases, and so does the announcement premium.

To better understand the nature of the announcement premium, we plot the price-dividend ratio in the model as a function of time under different assumptions of the posterior belief, $\hat{x}_t$. The dotted line is the price-to-dividend ratio assuming $\hat{x}_t = \bar{x}$, and the dash-dotted lines are plotted under values of $\hat{x}_t$ one standard deviation above and below $\bar{x}$, where standard deviation is calculated as the monthly standard deviation of the Brownian motion shock dB_t. Note that on average, announcements are associated with an immediately increase in the valuation ratio. After announcements, as discount rate increases, price-to-dividend ratio drops gradually until the next announcement. Because $\phi > 1$, the price-to-dividend ratio

[7]The equity premium on announcement days in the expected utility model is not literally zero because even though announcment premium is zeor, compensation for consumption risks is positive with $\Delta = \frac{1}{360}$.

is increasing in the posterior belief $\hat{x}_t$, and therefore the market equity requires a positive announcement premium by equation (30).

Comparison with alternative model specifications In Table 2, we present the model implied equity premium and announcement premium for our model with learning (left panel). For comparison, we also present under the column "Observable" the same moments for an otherwise identical model except that x_t is assumed to be fully observable. Under the column "Expected Utility", we report the moments of a model with expected utility by setting $\gamma = 1$ and keeping all other features identical to our benchmark learning model. Note that our model with learning produces a higher equity premium than the model with x_t fully observable. The average equity premium in the learning model is 5.4% per year, while the same moment is 4.58% in the model without learning. More importantly, in the model with learning, a large fraction of the equity premium are realized on the twelve scheduled announcement days: the total return on announcement days averages 2.53% per year. In the model where x_t is fully observable, announcements are not associated with any premium because their information content is fully anticipated.

As we discussed before, the announcement premium is zero under expected utility, and therefore the magnitude of equity premium on announcement days and that on non-announcement days are the same. In addition, under expected utility, risks in x_t are not priced and therefore the overall equity premium is a lot lower than that in economies with recursive utility.

4.3 Preference for early resolution of uncertainty

As we show in the previous section, within the class of recursive utility with constant relative risk aversion and constant IES, a positive announcement premium is equivalent to preference for early resolution of uncertainty. In this section, we use our calibrated model to provide a quantitative comparison of the magnitude of announcement premium and timing premium, defined as the welfare gain of early resolution of uncertainty.

We consider a macroeconomic announcement at time 0^+ that resolves all uncertainty in the economy from time 0^+ to time τ. Because $\gamma > \frac{1}{\psi}$, the resolution of uncertainty is associated with a non-trivial welfare gain. At the same time, investors who hold the market portfolio will receive an announcement premium at time 0^+ for the same reason. We now describe in detail the two related thought experiments.

Timing premium For simplicity, we focus on the special case of our model in which $\sigma_S^2 = 0$, that is, the true value of x_t is fully revealed upon announcement. As before, we use $V(\hat{x}, t, C)$ to denote the value function of the representative agent with posterior belief $\hat{x}$ at time t. We measure the welfare gain of information by the fraction of life-time consumption that the agent is willing to pay for the news announcement. Let $W_0^+(\tau)$ be the time-0^+ utility of an agent who knows all information about the economy (including information about $\{C_t\}_{t=0}^{\tau}$ and $\{x_t\}_{t=0}^{\tau}$) from time 0 to τ. Let $W_0^-(\tau)$ be the certainty equivalent of $W_0^+(\tau)$:

$$W_0^-(\tau) = \frac{1}{1-\gamma} \ln E_0^- \left[e^{(1-\gamma) W_0^+(\tau)} \right]. \tag{31}$$

That is, $W_0^-(\tau)$ is the utility of the representative agent who does not know any information about future but anticipates that all uncertainty from time 0 to τ will be resolved immediately at 0^+. Preference for early resolution implies that $V_0 < W_0^-(\tau)$, where $V_0 = V(x_0, 0, C_0)$ is the utility of the agent in the economy without advanced information as we defined in (24).[8] We define the timing premium $\lambda(\tau)$ as the maximum fraction of life-time consumption that the representative agent is willing to pay for the announcement:

$$\lambda(\tau) = 1 - e^{V_0 - W_0^-(\tau)}. \tag{32}$$

We call the dependence of $\lambda(\tau)$ on τ the term structure of timing premium.[9]

The utility associated with early resolution of uncertainty, $W_0^-(\tau)$ can be calculated as follows. At time 0^+, upon the news announcement, all uncertainty about consumption during the period $(0, \tau)$ is resolved. The agent's utility after news announcement at time 0^+ is computed as:

$$W_0^+(\tau) = \int_0^{\tau} e^{-\rho t} \rho \ln C_t dt + e^{-\rho \tau} V(x_\tau, 0, C_\tau). \tag{33}$$

In the above equation, the agent's utility function reduces to log on the interval $(0, \tau)$ because there is no uncertainty during this period. All uncertainty after time τ is incorporated in the continuation utility $V(x_\tau, 0, C_\tau)$. That is, we assume that after time τ, the dynamics of consumption and news announcement are identical to that in the benchmark economy we described in Section 4.1. At time 0^-, we can calculate the certainty equivalent of the continuation utility $W_0^+(\tau)$ by applying the certainty equivalence functional, (31).

Under our calculation, equation (32) is equivalent to the following definition of $\lambda(\tau)$: it is the fraction of life-time consumption reduction that makes the agent indifferent between not obtaining the information and knowing the information but accepting a consumption

[8] We use x_0 instead of $\hat{x}_0$ here because under the assumption that $\sigma_S^2 = 0$, the initial value of x_0 is known.

[9] Our definition of timing premium is a generalization of that in Epstein, Farhi, and Strzalecki [15]. Ai [1] provides a related decomposition of welfare gain in production economies.

reduction. Writing both V_0 and $W_0^-(\tau)$ as functions of state variables $(x_0, 0, C_0)$, formally, $\lambda(\tau)$ is the unique solution to the following equation:

$$W_0^-(\tau)(x_0, 0, (1-\lambda)C_0) = V(x_0, 0, C_0). \tag{34}$$

We provide a closed-form solution for $\lambda(\tau)$ in the appendix of the paper. Epstein, Farhi, and Strzalecki [15]'s definition of timing premium corresponds to the special case of $\lambda(\infty)$ in our framework.

Announcement premium We also calculate the market equity premium that investors receive upon the above described macroeconomic announcement. We denote $p_0^+(\tau)$ as the price-to-dividend ratio of the equity claim at time 0^+ upon the announcement that resolves all uncertainty during the period $(0, \tau)$. The pre-news price-to-dividend ratio is given by:

$$p_0^-(\tau) = \frac{E_0^-\left[e^{(1-\gamma)W_0^+(\tau)} p_0^+(\tau)\right]}{E_0^-\left[e^{(1-\gamma)W_0^+(\tau)}\right]}, \tag{35}$$

where $W_0^+(\tau)$ is the agent's continuation utility after announcement as defined in equation (33). The expected return associated with the announcement can be calculated as $\frac{E_0^-\left[p_0^+(\tau)\right]}{p_0^-(T)}$. We define the dependence of the announcement premium $ER(\tau) = \frac{E_0^-\left[p_0^+(\tau)\right]}{p_0^-(\tau)}$ on τ as the term structure of announcement premium.

We plot $\ln \frac{1}{1-\lambda(\tau)}$ and $\ln ER(\tau)$ in figure 8, where the top panel is the term structure at the one-year horizon, and the bottom panel illustrates the same up to five-years. The circles are the term structure of announcement premium (left scale), and the dotted line is the term structure of timing premium (right scale). Note that both the calculation of the timing premium and the announcement premium depends on the information content of the continuation utility $W_0^+(\tau)$ (equations (31) and (35)). It is therefore not surprising that both term structures are increasing in the amount of information in the announcement, and share a similar pattern.

The model we discussed in the last section corresponds to the short end of the term structure at $\tau = 0$. At zero, the premiums for news $\ln ER(0) = 0.17\%$, which is the same as the premium on announcement days reported in Figure 6. As τ increases, the announcement premium rises to about 4% at the one-year horizon. That is, if the news resolves all uncertainty in consumption for the next year, then the equity premium realized on the news announcement day is about 4%, which is comparable to the cumulative annual return on all announcement days generated in our previous model in the last section. Intuitively, equity

premium in our model has two components, the premium for news, which is mostly realized on news announcement days, and the premium for risks in consumption, which is realized gradually over time. If news announcement resolves all uncertainty for the next year, then the amount of equity premium realized immediately after news announcement is roughly equal to the cumulative annual equity premium for news in our previous model.

The term structure of timing premium, $\lambda(\tau)$ shows a similar pattern. It starts at 0 and slowly increases with τ. At the one-year horizon, the agent is willing to pay roughly 0.01% of his life-time consumption in order to receive the news that resolves all uncertainty for the next year.

In the bottom panel of figure 8, we plot the term structures up to five years. Because macroeconomic news rarely provide information for the aggregate economy beyond five years, this figure is informative about the magnitude of announcement premium and welfare gain at the horizon of macroeconomic news. If the news announcement resolves all uncertainty for the next five years, the equity premium realized on the news announcement day is an enormous 28%. The agent is willing to pay roughly 0.27% of his life-time consumption for the announcement due to preference for early resolution of uncertainty.

As in standard long-run risks models, the timing premium for resolving uncertainty until eternity is substantial. As $\tau \to \infty$, $\lambda(\tau) \to 28\%$ in our model, similar to the calculation in Epstein, Farhi, and Strzalecki [15]. However, as we discuss above, $\lambda(\tau)$ is substantially smaller at short horizons. In addition, the convergence is very slow: $\lambda(\tau)$ is about 13% at the thirty-year horizon and 18% at the fifty year horizon. Although it may be hard to gauge the timing premium for resolving all uncertainty until eternity by introspection, the magnitude of announcement premium in the data provides a market-based evidence for investors' willingness to pay for resolving uncertainty at the frequency of macro-economic news. In our calibration, the timing premium for macro-economic news is fairly small.

5 Conclusion

Motivated by the fact that a large fraction of the market equity premium are realized on a small number of trading days with significant macroeconomic news announcement, in this paper, we provide a theory and a quantitative analysis for premium for macroeconomic announcements. We provide a characterization for the set of preferences that give rise to the macroeconomic announcement premium. Our theoretical framework highlights that premium for macroeconomic news identifies equity market's compensation for uncertainty aversion as opposed to risk aversion. The empirical evidence on announcement premium can therefore be interpreted as a direct support for non-expected utility models with uncertainty aversion.

Our dynamic model quantitatively accounts for the pattern of equity premium around news announcement days. We also use our quantitative model as a laboratory to interpret the announcement premium as a market-based evidence for investors' preference for early resolution of uncertainty in the context of Epstein-Zin preferences. Our calibration implies a moderate amount of welfare gain from resolution of macroeconomic uncertainty at the frequency of macroeconomic news.

6 Appendix

A Details of the Empirical Evidence

In this section, we provide the details of our empirical evidence on the macroeconomic announcement premium.

Data description We focus on the following five major macroeconomic news announcements: unemployment/non-farm payroll (EMPL/NFP) and producer price index (PPI) published by the U.S. Bureau of Labor Statistics (BLS), the FOMC statements, the gross domestic product (GDP) reported by U.S. Bureau of Economic Analysis, and the Institute for Supply Management's Manufacturing Report (ISM) released by Bloomberg.[10]

The EMPL/NFL and PPI are both published at a monthly frequency and their announcement dates come from the BLS website. The BLS began announcing its scheduled release dates in advance in 1961 which is also the start date for our EMPL/NFL announcements sample. The PPI data series start in 1971.[11] There are a total of eight FOMC meetings each calendar year and the dates of FOMC meetings are taken from the Federal Reserve's web site. The FOMC statements begin in 1994 when the Committee started announcing its decision to the markets by releasing a statement at the end of each meeting. For meetings lasting two calendar days we consider the second day (the day the statement is released) as the event date. GDP is released quarterly beginning from 1997, which is the first year that full data are available, and the dates come from the BEA's website.[12] Finally, ISM is a monthly announcement with dates coming from Bloomberg starting from 1997. The last year for which we collect data on all announcements is 2014.

Our purpose is to demonstrate that a large fraction of equity premium is realized on a small number of trading days with significant macroeconomic news release. For this reason, we choose a set of macroeconomic announcements that occur only occasionally and contain significant news about the economy. With the exception of PPI, all other four series are

[10]Both unemployment and non-farm payroll information are released as part of the Employment Situation Report published by the BLS. We treat them as one announcement.

[11]While the CPI data is also available from the BLS back to 1961, once the PPI starts being published it typically precedes the CPI announcement. Given the large overlap in information between the two macro releases much of the "news" content in the CPI announcement will already be known to the market at the time of its release. For this reason we opt in favor of using PPI.

[12]GDP growth announcements are made monthly according to the following pattern: in April the advance estimate for Q1 GDP growth is released, followed by a preliminary estimate of the same Q1 GDP growth in May and a final estimate given in the June announcement. Arguably most uncertainty about Q1 growth is resolved once the advance estimate is published and most learning by the markets will occur prior to this release. For this reason we will focus only on the 4 advance estimate release dates every year.

ranked as top five macroeconomic announcements followed by Bloomberg users. We choose PPI because inflation is an important indicator of the macro-economy, and the news in PPI is not covered by other announcements in our sample. However, our results are robust to the exclusion of PPI.

Equity return on announcement days We first compare the average daily stock market return on news announcement days, which we denote as t, that on the day before the news announcement $(t-1)$, and that after the news announcement $(t+1)$. In Table 4, we present our results separately for each of the above news announcement and for all news, where standard errors are shown in parentheses below each number.[13] Excess market returns are taken from Kenneth French's web site. The vast majority of announcements are made on trading days. When this is not the case we assign the news release to the first trading day that follows the announcement.

For most of the news categories we consider, the announcement-day return is positive and significant, where the returns on non-announcement days, including $t-1$ and $t+1$ are generally not significantly different from zero. Between 1961 and 2014 we have a total of 13,592 trading days out of which 1,555 are announcement days and 12,037 are not. As shown in the table, the excess return is approximately 11 bp on announcement days (when all releases are stacked together) compared to about 1 bp on non-announcement days. In the 1997-2014 sub-sample (where all five macroeconomic announcements are available) there are a total of 4,530 trading days with macroeconomic releases coming out on 828 days and 3,702 being non-announcement days. The spread in excess returns between announcement and non-announcement days is even larger in this period with event day excess returns of 16 bp and 0 bp on non-event days.

To understand the economic significance of news announcement day returns, in Table **??**, we calculate the average annualized return of a strategy that long the market before the day of the pre-scheduled news announcements, hold it on the trading day with the news announcement, and sell immediately afterwards. We report the mean, the volatility, and the Sharpe ratio of the excess return of this strategy for the long sample in the top panel. We also report the same moments for the 1997-2014 subs-sample where data for all five macroeconomic news are available.

In the full 1961-2014 sample the annualized excess return on all news event days is 3.12%, which is about half of the 6.19% equity premium over this period. At the individual event level, the annualized excess returns achieved by trading on NFP/EMPL or PPI announcements are a little below 1%, while trading on FOMC and ISM achieves excess

[13] They are Newey-West standard errors (5-lags) of an OLS regression of excess returns on event dummies.

returns over 2%. The Sharpe ratio for this trading strategy is 0.51 when all events are used and as high as 0.80 when trading on FOMC announcements alone. In the 1997-2014 subsample the excess return with all events considered together is as high as 6.40%, which is roughly 85% of the 7.50% equity premium for this period. Not only does this trading strategy produce an excess return of similar size to the equity premium, but it also produces a Sharpe ratio of 0.76. The NFP/EMPL and ISM events on a stand alone basis give excess returns of approximately 2% at Sharpe ratios in the 0.42-0.44 range. The FOMC announcements stands out again with an excess return of 2.91% and a Sharpe ratio of 0.82.

Our empirical results are consistent with the previous research on stock market returns on macroeconomic announcement days. Savor and Wilson [39] document that stock market returns and Sharpe ratios are significantly higher on days with macroeconomic news releases. Lucca and Moench [31] show a pre-announcement drift before FOMC announcements, while Cieslak, Morse, and Vissing-Jorgensen [11] document the pattern of equity premium over FOMC meeting cycles.

High frequency returns To better identify the effect of news on the stock market, we also calculate high-frequency (30-minutes) equity returns around news announcement. Here we focus on the 1997-1014 period, in which the exact time of all five news releases are reported by Bloomberg. In Figure 1, we plot the average stock market returns over 30-minute intervals before and after the news event.

The return at time 0 is the 30-minute news event return. Employment/Non-farm payroll, GDP and PPI announcements are made at 8:30 AM before the market begins. In these cases we will consider the 30-minute news event return to be the return between 4:00 PM (close of trading) of the previous day and 9:30 AM when the market opens on the day of the announcement. The 30-minute event return for ISM announcements, which are made at 10:00AM, covers the interval between 9:30 AM and 10:00 AM of the announcement day. Finally, the timing of the FOMC news release varies. We add 30 minutes to the announcement time to account for the press conference after the FOMC meeting.[14] Once we normalize the news event as time 0, we define -2 as the interval between -60 minutes and -30 minutes relative to the news event, and -1 as the interval between -30 and 0 minutes relative to the news event. Time $+1$ and $+2$ are defined similarly.

The return goes up from -3.34 bp to 2.89 bp and then reaches 6.94 bp during the 30 minute window containing the macro news release. Once the news is incorporated by the market, the return falls to -1.12 bp and $-2.55bp$ following the news events.

[14]For example if a statement is released at 14:15 PM, we add 30 minutes for the press conference that follows and then we round the event time to 15:00 PM.

B Proof of Theorem 1 and 2

B.1 Definitions and Notations

We use $\mathbf{R}$ to denote the real line and $\mathbf{R}^n$ to denote the $n-$dimensional Euclidean space. In our setup, $\forall t = 1, 2 \cdots T$, Y_t is a $(\mathbf{R}, \mathcal{B})$ valued random variable, where we use $\mathcal{B}$ for the associated Borel $\sigma-$algebra. We define a filtration $\mathcal{F}_1^- = \sigma(Y_1)$, $\mathcal{F}_1^+ = \sigma(Y_1, s_1)$, $\mathcal{F}_2^- = \mathcal{F}_1^+ \vee \sigma(Y_2)$, $\mathcal{F}_2^+ = \mathcal{F}_1^+ \vee \sigma(Y_2, s_2)$, $\cdots$. A certainty equivalence functional $\mathcal{I}[\cdot]$ is a mapping $\mathcal{I} : L^2(\Omega, \mathcal{F}, P) \to \mathbf{R}$. We first state a definition of first order stochastic dominance (FSD) and second order stochastic dominance (SSD).

Definition 1. First order stochastic dominance: *X_1 first order stochastic dominates X_2, or $X_1 \geq_{FSD} X_2$ if there exist a random variable $Y \geq 0$ a.s. such that X_1 has the same distribution as $X_2 + Y$. Strict monotonicity, $X_1 >_{FSD} X_2$ holds if $P(Y > 0) > 0$ in the above definition.*

Definition 2. Second order stochastic dominance: *X_1 second order stochastic dominates X_2, or $X_1 \geq_{SSD} X_2$ if there exist a random variable Y such that $E[Y|X_1] = 0$ and X_2 has the same distribution as $X_1 + Y$. Strict monotonicity, $X_1 >_{SSD} X_2$ holds if $P(Y \neq 0) > 0$ in the above definition.*

Monotonicity with respect to FSD and SSD are defined as:

Definition 3. Monotonicity with respect to FSD (SSD): *The certainty equivalence functional $\mathcal{I}$ is said to be monotone with respect to FSD (SSD) if $\mathcal{I}[X_1] \geq \mathcal{I}[X_2]$ whenever $X_1 \geq_{FSD} X_2$ ($X_1 \geq_{SSD} X_2$). $\mathcal{I}$ is strictly monotone with respect to FSD (SSD) if $\mathcal{I}[X_1] > \mathcal{I}[X_2]$ whenever $X_1 >_{FSD} X_2$ ($X_1 >_{SSD} X_2$).*[15]

To construct the SDF for pure news events from marginal utilities, we need to use some concepts from standard functional analysis (see for example, Luenberger [32] and Rall [36]) to impose a differentiability condition on the certainty equivalence functional. Let $\mathcal{I} : L^2(\Omega, \mathcal{F}, P) \to R$ be a mapping from the normed vector space $L^2(\Omega, \mathcal{F}, P)$ to the real line. We use $\|\cdot\|$ to denote the L^2 norm and assume that $\mathcal{I}$ satisfy the following differentiability condition.

Definition 4. *(Fréchet Differentiable with Lipschitz Derivatives) The certainty equivalence functional $\mathcal{I}$ is Fréchet Differentiable if $\forall X \in L^2(\Omega, \mathcal{F}, P)$, there exist a unique continuous*

[15] Our definition is the same as the standard concept of risk aversion (see Rothschild and Stiglitz [38] and Werner [43]) except that the certainty equivalence function $\mathcal{I}$ is defined on the space of continuation utilities rather than consumption.

linear functional, $D\mathcal{I}[X] \in L^2(\Omega, \mathcal{F}, P)$ *such that for all* $\Delta X \in L^2(\Omega, \mathcal{F}, P)$

$$\lim_{\|\Delta X\| \to 0} \frac{\left|\mathcal{I}[X + \Delta X] - \mathcal{I}[X] - \int D\mathcal{I}[X] \cdot \Delta X dP\right|}{\|\Delta X\|} = 0.$$

A Fréchet differentiable certainty equivalence functional $\mathcal{I}$ *is said to have a Lipschitz derivatives if* $\forall X, Y \in L^2(\Omega, \mathcal{F}, P)$, $\|D\mathcal{I}[X] - D\mathcal{I}[Y]\| \leq K \|X - Y\|$ *for some constant* K.[16]

The above assumption is made for two purposes. First it allows us to apply the envelope theorems in Milgrom and Segal [35] to establish differentiability of the value functions. Second, it allows us to compute the derivatives of $\mathcal{I}$ to derive an appropriate SDF and use derivatives of $\mathcal{I}$ to integrate back to recover the certainty equivalence functional.[17] Intuitively, we use the following operation to relate the certainty equivalence function $\mathcal{I}$ and its derivatives. $\forall$ X,$Y \in L^2(\Omega, \mathcal{F}, P)$, to compute $\mathcal{I}(Y) - \mathcal{I}(X)$ from its derivatives, we define $g(t) = \mathcal{I}[X + t(Y - X)]$ for $t \in [0, 1]$ and write

$$\begin{aligned} \mathcal{I}[Y] - \mathcal{I}[X] &= g(1) - g(0) \\ &= \int_0^1 g'(t)\, dt \\ &= \int_0^1 \int_\Omega D\mathcal{I}[X + t(Y - X)](Y - X)\, dP dt. \end{aligned} \tag{36}$$

Fréchet Differentiable with Lipschitz Derivatives guarantees that the function $g(t)$ is continuously differentiable. The differentiability of g is straightforward (see for example, Luenberger [32]). To see that $g'(t)$ is continuous, note that

$$\begin{aligned} g'(t_1) - g'(t_2) &= \int_\Omega \{D\mathcal{I}[X + t_1(Y - X)] - D\mathcal{I}[X + t_2(Y - X)]\}(Y - X)\, dP \\ &\leq \|D\mathcal{I}[X + t_1(Y - X)] - D\mathcal{I}[X + t_2(Y - X)]\| \cdot \|Y - X\|. \end{aligned}$$

The Lipschitz continuity $D\mathcal{I}$ implies that

$$\|D\mathcal{I}[X + t_1(Y - X)] - D\mathcal{I}[X + t_2(Y - X)]\| \leq (t_1 - t_2) \|(Y - X)\|,$$

and the latter vanishes as $t_2 \to t_1$. In addition, applying the mean value theorem on g, for

[16] The standard definition of Fréchet Differentiability requires the existence of the derivative as a continuous linear functional. Because we focus on functions defined on $L^2(\Omega, \mathcal{F}, P)$, we apply the Riesz representation theorem and denote $D\mathcal{I}[X]$ as the representation of the derivative in $L^2(\Omega, \mathcal{F}, P)$.

[17] A weaker notion of differentiability, Gâtteaux differentiability is enough to derive a SDF for pure news events. However, the converse of Theorem 1 requires a stronger condition for differentiability, which is what we assume here.

some $\hat{t} \in (0,1)$, we have:

$$\mathcal{I}[Y] - \mathcal{I}[X] = \int_{\Omega} D\mathcal{I}\left[X + \hat{t}(Y - X)\right](Y - X)\, dP. \tag{37}$$

B.2 Differentiability of value functions

In this section, we establish the differentiability of the value functions recursively. In particular, we show that the value functions are elements of $\mathcal{D}$, where $\mathcal{D}$ is defined as

Definition 5. *$\mathcal{D}$ is the set of differentiable functions on the real line such that $\forall f \in \mathcal{D}$, i) f is Lipschitz continuous; ii) $\forall x \in \mathbf{R}$, $\frac{1}{h}[f(x+h-a) - f(x-a)]$ converges uniformly to $f'(x-a)$ in a. That is, $\forall \varepsilon > 0$, there exists a $\delta > 0$ such that $|h| < 0$ implies that $\left|\frac{1}{h}[f(x+h-a) - f(x-a)] - f'(x-a)\right| < \varepsilon$ for all $a \in \mathbf{R}$.*

We first establish that $V_t^+, V_t^- \in \mathcal{D}$ for all t. For any $v \in \mathcal{D}$, we define f_v and g_v as functions $of\,(W, \xi)$, where W is the wealth level, and $\xi \in \mathbf{R}^{J+1}$ is a portfolio strategy:

$$f_v(W,\xi) = u\left(W - \sum\nolimits_{j=0}^{J} \xi_j\right) + \beta \mathcal{I}\left[v\left(\sum\nolimits_{j=0}^{J} \xi_j R_j\right)\right], \tag{38}$$

$$g_v(W,\xi) = \mathcal{I}\left[v\left(W + \sum\nolimits_{j=0}^{J} \xi_j (R_j - 1)\right)\right]. \tag{39}$$

Because $R_j \in L^2(\Omega, \mathcal{F}, P)$ and v is Lipschitz continuous, $v\left(\sum_{j=0}^{J} \xi_j R_j\right)$ and $v\left(W - \sum_{j=0}^{J} \xi_j (R_j - 1)\right)$ are both square integrable. Therefore the above expressions are well-defined. A key condition to apply to the envelope theorem in Milgrom and Segal [35] to establish the differentiability of the value functions is equi-differentiability. The following lemma establishes the equi-differentiability of $\{f_v(W,\xi)\}_\xi$ and $\{g_v(W,\xi)\}_\xi$:

Lemma 1. *Suppose $u, v \in \mathcal{D}$, as $h \to 0$, both $\frac{1}{h}[f_v(W+h,\xi) - f_v(W,\xi)]$ and $\frac{1}{h}[g_v(W+h,\xi) - g_v(W,\xi)]$ converge uniformly for all ξ.*

Proof: *First,*

$$\frac{1}{h}[f_v(W+h,\xi) - f_v(W,\xi)] = \frac{1}{h}\left[u\left(W + h - \sum\nolimits_{j=0}^{J} \xi_j\right) - u\left(W - \sum\nolimits_{j=0}^{J} \xi_j\right)\right]$$

converges uniformly because $u \in \mathcal{D}$. Next, we need to show that

$$\frac{1}{h}[g_v(W+h,\xi) - g_v(W,\xi)] \to \frac{\partial}{\partial W} g_v(W,\xi) \tag{40}$$

and the convergence is uniform for all ξ*. Note that*

$$\frac{\partial}{\partial W}g_v(W,\xi) = \int D\mathcal{I}\left[v\left(W-\sum\nolimits_{j=0}^{J}\xi_j(R_j-1)\right)\right]\cdot v'\left(W-\sum\nolimits_{j=0}^{J}\xi_j(R_j-1)\right)dP$$

and

$$\begin{aligned} g_v(W+h,\xi)-g_v(W,\xi) &= \mathcal{I}\left[v\left(W+h+\sum\nolimits_{j=0}^{J}\xi_j(R_j-1)\right)\right]-\mathcal{I}\left[v\left(W+\sum\nolimits_{j=0}^{J}\xi_j(R_j-1)\right)\right] \\ &= \int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right](\bar{v}(1)-\bar{v}(0))\,dP, \quad \textit{for some } t\in(0,1), \end{aligned}$$

where we denote $\bar{v}(t) = tv\left(W+h-\sum_{j=0}^{J}\xi_j(R_j-1)\right)+(1-t)\,v\left(W-\sum_{j=0}^{J}\xi_j(R_j-1)\right)$ *and applied equation (37). Also, denote* $\bar{v}'(0) = v'\left(W-\sum_{j=0}^{J}\xi_j(R_j-1)\right)$*, then the right hand side of (40) can be written as* $\int_\Omega D\mathcal{I}[\bar{v}(0)]\,\bar{v}'(0)\,dP$*, we have:*

$$\begin{aligned} &\left|\frac{1}{h}\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right](\bar{v}(1)-\bar{v}(0))\,dP-\int_\Omega D\mathcal{I}[\bar{v}(0)]\,\bar{v}'(0)\,dP\right| \\ = &\left|\begin{array}{c}\frac{1}{h}\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right](\bar{v}(1)-\bar{v}(0))\,dP-\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\bar{v}'(0)\,dP \\ +\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\bar{v}'(0)\,dP-\int_\Omega D\mathcal{I}[\bar{v}(0)]\,\bar{v}'(0)\,dP\end{array}\right| \\ \leq &\int_\Omega\left|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\right|\left|\frac{1}{h}(\bar{v}(1)-\bar{v}(0))-\bar{v}'(0)\right|dP+\int_\Omega\left|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]-D\mathcal{I}[\bar{v}(0)]\right|\left|\bar{v}'(0)\right|dP \\ \leq &\left\|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\right\|\left\|\frac{1}{h}(\bar{v}(1)-\bar{v}(0))-\bar{v}'(0)\right\|+\left\|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]-D\mathcal{I}[\bar{v}(0)]\right\|\left\|\bar{v}'(0)\right\| \qquad (41) \end{aligned}$$

Because $v\in\mathcal{D}$*, for* h *small enough,* $\left|\frac{1}{h}(\bar{v}(1)-\bar{v}(0))-\bar{v}'(0)\right|\leq\varepsilon$ *with probability one and* $\left\|\frac{1}{h}(\bar{v}(1)-\bar{v}(0))-\bar{v}'(0)\right\|\leq\varepsilon$*. Also, because* $D\mathcal{I}$ *is Lipschitz continuous,* $\left\|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]-D\mathcal{I}[\bar{v}(0)]\right\|\leq K\left\|\bar{v}(1)-\bar{v}(0)\right\|\leq K^2h$*, where the second inequality is due to the Lipschitz continuity of* v*. This proves the uniform convergence of (41).*

Lemma 2. *Suppose* $u\in\mathcal{D}$*, then both* T^+ *and* T^- *map* $\mathcal{D}$ *into* $\mathcal{D}$*.*

Proof: *Note that* $T^+v(W) = \sup_\xi f_v(W,\xi)$ *and* $T^-v(W) = \sup_\xi g_v(W,\xi)$*, where* $f_v(W,\xi)$ *and* $g_v(W,\xi)$ *are defined in (38) and (39). It then follows from Lemma 1 that Theorem 3 in Milgrom and Segal [35] applies. Therefore, both* T^+v *and* T^-v *are differentiable, and*

$$\begin{aligned} \frac{d}{dW}T^+v(W) &= u'\left(W-\sum\nolimits_{j=0}^{J}\xi_j(W)\right) \\ \frac{d}{dW}T^-v(W) &= \int D\mathcal{I}\left[v\left(W-\sum\nolimits_{j=0}^{J}\xi_j(W)(R_j-1)\right)\right]\cdot v'\left(W-\sum\nolimits_{j=0}^{J}\xi_j(W)(R_j-1)\right)dP, \end{aligned}$$

where $\xi(W)$ *denotes the utility-maximizing portfolio at* W*.*

To see that $T^{+}v(W)$ is Lipschitz continuous, note that

$$f_v(W_1,\xi(W_2))-f_v(W_2,\xi(W_2)) \leq T^{+}v(W_1)-T^{+}v(W_2) \leq f_v(W_1,\xi(W_1))-f_v(W_2,\xi(W_1)). \tag{42}$$

Because $\forall\xi, \quad |f(W_1,\xi)-f(W_2,\xi)| = \left|u\left(W_1-\sum_{j=0}^{J}\xi_j\right)-u\left(W_2-\sum_{j=0}^{J}\xi_j\right)\right| \leq K|W_1-W_2|$, *where K is a Lipschitz constant for u, $|Tv(W_1)-Tv(W_2)| \leq K|W_1-W_2|$. We can prove that $T^{-}v(W)$ is Lipschitz continuous in a similary way:*

$$g_v(W_1,\xi(W_2))-g_v(W_2,\xi(W_2)) \leq T^{-}v(W_1)-T^{-}v(W_2) \leq g_v(W_1,\xi(W_1))-g_v(W_2,\xi(W_1)). \tag{43}$$

Note that $\forall\xi$,

$$\begin{aligned} |g_v(W_1,\xi)-g_v(W_2,\xi)| &= \left|\mathcal{I}\left[v\left(W_1+\sum_{j=0}^{J}\xi_j(R_j-1)\right)\right]-\mathcal{I}\left[v\left(W_2+\sum_{j=0}^{J}\xi_j(R_j-1)\right)\right]\right| \\ &\leq K\left\|v\left(W_1+\sum_{j=0}^{J}\xi_j(R_j-1)\right)-v\left(W_2+\sum_{j=0}^{J}\xi_j(R_j-1)\right)\right\| \\ &\leq K^2|W_1-W_2|, \end{aligned}$$

where the inequalities are due to the Lipschitz continuity of $\mathcal{I}$ and v, respectively.

Finally, equations (42) and (43) can be used to show that the family of functions $\{T^{+}v(W-a)\}_a$ and $\{T^{-}v(W-a)\}_a$ are equi-differentiable. For example, let $W_1 \to W_2$,

$$\frac{1}{W_1-W_2}\left[f_v(W_1,\xi)-f_v(W_2,\xi)\right]$$

converges uniformly by Lemma 1, and by equation (42), $\frac{1}{W_1-W_2}\left[T^{+}v(W_1)-T^{+}v(W_2)\right]$ must also converge uniformly.

Given that $u \in \mathcal{D}$, Lemma 2 can be used to establish the differentiability of $V_t^{+}(W)$ and $V_t^{-}(W)$ recursively. Finally, we note that if $u'(x) > 0$ for all $x \in \mathbf{R}$, then $V_t^{+}(W)$ and $V_t^{-}(W)$ must satisfy the same property by the envelope theorem.

B.3 Existence of SDF for pure news events

We write the time t^{-} portfolio selection problem of the agent as

$$\max_{\zeta} \mathcal{I}\left[V_t^{+}\left(W_t^{-}+\sum_{j=0}^{J}\zeta_j\left(R_{j,t}^{+}-1\right)\right)\middle|\,\mathcal{F}_t^{-}\right], \tag{44}$$

where $R_t^{+}=\left[R_{0,t}^{+},R_{1,t}^{+},R_{2,t}^{+},\cdots R_{J,t}^{+}\right]$ is a vector of announcement returns that depends on the announcement at time t, s_t, that is, they are $\mathcal{F}_t^{+}$ measurable. We use the notation

$\mathcal{I}\left[\cdot|\,\mathcal{F}_t^-\right]$ to emphasize that the certainty equivalence functional $\mathcal{I}$ maps $L^2\left(\Omega,\mathcal{F}_t^+,P\right)$ into $L^2\left(\Omega,\mathcal{F}_t^-,P\right)$. Clearly, no arbitrage implies that the risk-free announcement return $R_{0,t}^+=1$.

We denote $W_t^+=W_t^-+\sum_{j=0}^J\zeta_j\left(R_{j,t}^+-1\right)$. The value function $V^+\left(W_t^+\right)$ is determined by the the agent's portfolio choice problem at time t^+ after the announcement s_t is made:

$$V_t^+\left(W_t^+\right)=\max_{\xi}u\left(W_t^+-\sum\nolimits_{j=0}^J\xi_j\right)+\beta\mathcal{I}\left[V_{t+1}^-\left(\sum\nolimits_{j=0}^J\xi_jR_{j,t+1}^-\right)\middle|\,\mathcal{F}_t^+\right],\tag{45}$$

where $R_{t+1}^-=\left[R_{0,t+1}^-,R_{1,t+1}^-,R_{2,t+1}^-,\cdots R_{J,t+1}^-\right]$ is a vector of intertemporal returns.

Because the time-t^+ value function, V_t^+ is differentiable, and the certainty equivalent functional, $\mathcal{I}$ is Fréchet differentiable, $\mathcal{I}\left[V_t^+\left(W_t^-+\sum_{j=0}^J\zeta_j\left(R_{j,t}^+-1\right)\right)\middle|\,\mathcal{F}_t^-\right]$ is differentiable in ζ.[18] Therefore, the first order condition with respect to ζ_j implies that

$$E\left[D\mathcal{I}\left[V^+\left(W_t^+\right)\right]\frac{d}{dW}V_t^+\left(W_t^+\right)\left(R_{j,t}^+-1\right)\middle|\,\mathcal{F}_t^-\right]=0.\tag{46}$$

Also, the envelop condition for (45) implies

$$\frac{d}{dW}V_t^+\left(W_t^+\right)=u'\left(W_t^+-\sum\nolimits_{j=0}^J\xi_j\right)=u'\left(Y_t\right),$$

where the last equality uses the resource constraint because period-t consumption must equal to total endowment. Note that $u'\left(Y_t\right)>0$ and is $\mathcal{F}_t^-$ measurable; therefore, (46) implies:

$$E\left[D\mathcal{I}\left[V_t^+\left(W_t^+\right)\right]\left(R_{j,t}^+-1\right)\middle|\,\mathcal{F}_t^-\right]=0.\tag{47}$$

As we show in the next section, monotonicity of $\mathcal{I}$ guarantees that $D\mathcal{I}\geq 0$ with probability one. To derive equation (13), we need to assume a slightly stronger condition:

$$D\mathcal{I}\left[X\right]>0 \text{ with strictly positive probability for all } X.^{19}\tag{48}$$

In this case, the SDF in (13) can be constructed as:

$$m_t^*=\frac{D\mathcal{I}\left[V_t^+\left(W_t^+\right)\right]}{E\left[D\mathcal{I}\left[V_t^+\left(W_t^+\right)\right]\middle|\,\mathcal{F}_t^-\right]}.\tag{49}$$

From the above construction, it is clear that if $\mathcal{I}$ is expected utility, then m_t^* must be a

[18] See for example Proposition 1 in Chapter 7 of Luenberger [32].

[19] Note that monotonicity with respect to FSD implies that $D\mathcal{I}\left[X\right]\geq 0$ with probability one for all X. If condition (48) does not hold, we must have $D\mathcal{I}\left[X\right]=0$ with probability one. If $\mathcal{I}$ is strictly monotone with respect to FSD, then this cannot happen on an open set in L^2. Therefore, even without assuming (48), our result implies that the SDF for pure news events exists generically.

constant. Conversely, if m_t^* is a constant, then $\mathcal{I}$ is linear and represents expected utility.

The SDF for pure news event is constructed from the Fréchet Derivative of the certainty equivalence functional, and therefore a linear functional on $L^2\left(\Omega, \mathcal{F}_t^+, P\right)$. By the Riesz representation theorem, it has a representation as an element in $L^2\left(\Omega, \mathcal{F}_t^+, P\right)$, that is, its representation is measurable with respect to the signal s_t. Below we show that m_t^* can be represented as a measurable function of continuation utility: $m_t^* = m_t^*\left(V_t^+\right)$. That is, m_t^* depends on s_t only through the continuation utility. Here, with a slight abuse of notation, we also use m_t^* for a measurable function from $(\mathbf{R}, \mathcal{B})$ to $(\mathbf{R}, \mathcal{B})$.

Lemma 3. *If $\mathcal{I}$ is invariant with respect to distribution, then $D\mathcal{I}[X]$ can be represented by a measurable function of X.*

Proof: *Take any $X \in L^2\left(\Omega, \mathcal{F}_t^+, P\right)$, let T be a measure-preserving transformation such that the invariant $\sigma-$field of T differ from the $\sigma-$field generated by X (which we denote as $\sigma(X)$) only by measure zero sets (For the existence of such measure-preserving transformations, see exercise 17.43 in Kechris [25]). Let $D\mathcal{I}[X]$ be the $L^2\left(\Omega, \mathcal{F}_t^+, P\right)$ representation of the Fréchet Derivative of the certainty equivalence functional $\mathcal{I}$ at X. Below, we first show that $D\mathcal{I}[X] \circ T$ must also be a Fréchet Derivative of $\mathcal{I}$ at X. Because the Fréchet Derivative is unique, we must have $D\mathcal{I}[X] = D\mathcal{I}[X] \circ T$ with probability one; therefore, $D\mathcal{I}[X]$ must be measurable with respect to the invariant σ-field of T and therefore, also measurable with respect to $\sigma(X)$.*

Because $\mathcal{I}[\cdot]$ is Fréchet differentiable, to show $D\mathcal{I}[X] \circ T$ is the Fréchet Derivative of $\mathcal{I}$ at X, it is enough to verify that $D\mathcal{I}[X] \circ T$ is a Gâteaux derivative, that is,

$$\lim_{\alpha \to 0} \frac{1}{\alpha}\left[V(X + \alpha Y) - V(X)\right] = \int \left(D\mathcal{I}[X] \circ T\right) \cdot Y dP \tag{50}$$

for all $Y \in L^2\left(\Omega, \mathcal{F}_t^+, P\right)$.

Because T is measure preserving and X is measurable with respect to the invariance $\sigma-$field of T, $X = X \circ T$ with probability one. Therefore, $V(X + \alpha Y) = V(X \circ T + \alpha Y) = V(X + \alpha Y \circ T^{-1})$, where the second equality is due to the fact that T^{-1} is measure preserving, and $[X \circ T + \alpha Y] \circ T^{-1} = X + \alpha Y \circ T^{-1}$ has the same distribution with $X \circ T + \alpha Y$. As a result,

$$\begin{aligned}
\frac{1}{\alpha}\left[V(X + \alpha Y) - V(X)\right] &= \frac{1}{\alpha}\left[V\left(X + \alpha Y \circ T^{-1}\right) - V(X)\right] \\
&= \int D\mathcal{I}[X] \times Y \circ T^{-1} dP, \\
&= \int D\mathcal{I}[X] \circ T \cdot Y dP,
\end{aligned}$$

where the last equality uses the fact that $[D\mathcal{I}[X] \cdot Y \circ T^{-1}] \circ T = D\mathcal{I}[X] \circ T \cdot Y$ *have the same distribution with* $D\mathcal{I}[X] \times Y \circ T^{-1}$*. This proves (50).*

Note that our definition of monotonicity with respect to FSD implies invariance with respect to distribution. To see this, note that X has the same distribution of Y implies both $X \leq_{FSD} Y$ and $Y \geq_{FSD} X$. As a result, $\mathcal{I}[X] = \mathcal{I}[Y]$ by monotonicity. By Lemma 3, $D\mathcal{I}\left[V_t^+\left(W_t^+\right)\right]$ is measurable with respect to $V_t^+\left(W_t^+\right)$ and we can write $m_t^* = m_t^*\left(V_t^+\right)$ for some measurable function m_t^*.

B.4 Properties of SDF for pure news events

In this section, we establish two results. First, $m_t^*\left(V_t^+\right)$ is non-negative if and only if $\mathcal{I}$ is monotone with respect to FSD. Second, $m_t^*\left(V_t^+\right)$ is non-increasing in V_t^+ if and only if $\mathcal{I}$ is monotone with respect to SSD. As a result, $\mathcal{I}$ is monotone with respect to SSD if and only if the announcement premium is non-negative for all risky assets.

We first establish the non-negativity of the SDF, $m_t^*\left(V_t^+\right)$, which is Lemma 4 below.

Lemma 4. *$\mathcal{I}$ is monotone with respect FSD if and only if $D\mathcal{I}[X] \geq 0$ a.s.*

Proof: *Suppose $D\mathcal{I}[X] \geq 0$ a.s. for all $X \in L^2(\Omega, \mathcal{F}, P)$. Take any Y such that $Y \geq 0$ a.s., we have:*

$$\mathcal{I}[X+Y] - \mathcal{I}[X] = \int_0^1 \int_\Omega D\mathcal{I}[X+tY]\, Y\, dPdt \geq 0.$$

Conversely, suppose $\mathcal{I}$ is monotone with respect to FSD, we can prove $D\mathcal{I}[X] \geq 0$ a.s. by contradiction. Suppose the latter is not true and there exist an $A \in \mathcal{F}$ with $P(A) > 0$ and $D\mathcal{I}[X] < 0$ on A. Because $D\mathcal{I}$ is continuous, we can assume that $D\mathcal{I}[X + t\chi_A] < 0$ on A for all $t \in (0, \varepsilon)$ for ε small enough, where χ_A is the indicator function of A. Therefore,

$$\mathcal{I}[X+\chi_A] - \mathcal{I}[X] = \int_0^1 \int_\Omega D\mathcal{I}[X+t\chi_A]\, \chi_A dPdt < 0,$$

contradicting monotonicity with respect to FSD.

Next, we show that $\mathcal{I}$ is monotone with respect to SSD if and only if $m_t^*\left(V_t^+\right)$ is non-increasing in V_t^+. We first prove the following lemma.

Lemma 5. *$\mathcal{I}$ is monotone with respect SSD if and only if $\forall X \in L^2(\Omega, \mathcal{F}, P)$, for any $\sigma-$field $\mathcal{G} \subseteq \mathcal{F}$,*

$$\int D\mathcal{I}[X] \cdot (X - E[X|\mathcal{G}])\, dP \leq 0. \tag{51}$$

Proof: Suppose condition (51) is true, by the definition of SSD, for any X and Y such that $E[Y|X]=0$, we need to prove

$$\forall\lambda\in(0,1),\quad \mathcal{I}(X)\geq\mathcal{I}(X+Y).$$

Using (36),

$$\begin{aligned}\mathcal{I}(X+Y) &\geq \mathcal{I}(X)=\int_0^1\int_\Omega D\mathcal{I}[X+tY]\,Y\,dPdt\\ &= \int_0^1\frac{1}{t}\int_\Omega D\mathcal{I}[X+tY]\{tY+X-X-tE[Y|X]\}\,dPdt\\ &= \int_0^1\frac{1}{t}\int_\Omega D\mathcal{I}[X+tY]\{[X+tY]-E[X+tY|X]\}\,dPdt\\ &\leq 0,\end{aligned}$$

where the last inequality uses (51).

Conversely, assuming $\mathcal{I}$ is increasing in SSD, we prove (51) by contradiction. if (51) is not true, then by the continuity of $D\mathcal{I}[X]$, for some $\varepsilon>0$, $\forall t\in(0,\varepsilon)$,

$$\int D\mathcal{I}[(1-t)X+tE[X|\mathcal{G}]]\cdot(X-E[X|\mathcal{G}])\,dP>0.$$

Therefore,

$$\mathcal{I}[(1-\varepsilon)X+\varepsilon E[X|\mathcal{G}]]-\mathcal{I}[X]=\int_0^\varepsilon\int D\mathcal{I}[(1-t)X+tE[X|\mathcal{G}]]\{E[X|\mathcal{G}]-X\}\,dPdt<0.$$

However, $(1-\varepsilon)X+\varepsilon E[X|\mathcal{G}]\geq_{SSD}X$, a contradiction.[20]

Due to Lemma 3, $D\mathcal{I}[X]$ can be represented by a function of X, we denote $D\mathcal{I}[X]=\eta(X)$. To establish the equivalence between monotonicity with respect to SSD and the (negative) monotonicity of $m_t^*\left(V_t^+\right)$, we only need to prove that condition (51) is equivalent to $\eta(\cdot)$ being a non-increasing function, which is Lemma 6 below.

Lemma 6. *Condition (51) is equivalent to $\eta(X)$ being a non-increasing function of X.*

Proof: *First, we assume $\eta(X)$ is non-increasing. To prove (51), note that $E[X|\mathcal{G}]$ is*

[20] An easy way to prove the statement, $(1-\varepsilon)X+\varepsilon E[X|\mathcal{G}]\geq_{SSD}X$ is to observe that an equivalent definition of SSD is $X_1\geq_{SSD}X_2$ if $E[\phi(X_1)]\geq E[\phi(X_2)]$ for all concave functions ϕ (see Rothschild and Stiglitz [38] and Werner [43]). If $E[Z|V_1]=0$, then for any concave function ϕ, $\phi(V_1+\lambda Z_1)\geq\lambda\phi(V_1+Z)+(1-\lambda)\phi(V_1)$. Therefore, $E[\phi(V_1+\lambda Z_1)]\geq\lambda E[\phi(V_1+Z)]+(1-\lambda)E[\phi(V_1)]$
$\geq E[\phi(V_1+Z)]$, where the last inequality is true because $E[\phi(V_1)]\geq E[\phi(V_1+Z_1)]$.

measurable with respect to $\sigma(X)$*, and we can the Law of Iterated Expectation to write:*

$$\begin{aligned}\int D\mathcal{I}[X]\cdot(X-E[X|\mathcal{G}])\,dP &= E[\eta(X)\cdot(X-E[X|\mathcal{G}])]\\ &\leq E[\eta(E[X|\mathcal{G}])\cdot(X-E[X|\mathcal{G}])]\\ &= 0,\end{aligned}$$

where the inequality follows from the fact that $\eta(X)\leq\eta(E[X|\mathcal{G}])$ *when* $X\geq E[X|\mathcal{G}]$ *and* $\eta(X)\geq\eta(E[X|\mathcal{G}])$ *when* $X\leq E[X|\mathcal{G}]$*.*

Second, to prove the converse of the above statement by contradiction, we assume (51) is true, but there exist $x_1<x_2$*, both occur with positive probability such that* $\eta(x_1)<\eta(x_2)$*. We a random variable* Y*:*

$$Y=\begin{cases}0 & X=x_1 \ \ or \ \ x_2\\ X & otherwise\end{cases}.$$

and denote $P_1=P(X=x_1)$*,* $P_2=P(X=x_2)$*. Note that*

$$\begin{aligned}&\int D\mathcal{I}[X]\cdot(X-E[X|Y])\,dP\\ =\ &\int\eta(X)\cdot(X-E[X|Y])\,dP\\ =\ &P_1\eta(x_1)\left[x_1-\frac{P_1x_1+P_2x_2}{P_1+P_2}\right]+P_2\eta(x_2)\left[x_2-\frac{P_1x_1+P_2x_2}{P_1+P_2}\right]\\ >\ &0\end{aligned}$$

because $\eta(x_1)<\eta(x_2)$*, a contradiction.*

To formalize the equivalence between monotonicity with respect to SSD and a non-negative announcement premium for all risky assets, we consider payoffs that are functions of continuation utility, $f\left(V_t^+\right)$. We define a utility-dependent payoff to be risky if $f(v)$ is a non-decreasing function on the set $\{v:f(v)\neq0\}$.[21] We have the following result.

Lemma 7. $m_t^*\left(V_t^+\right)$ *is a non-increasing function of* V_t^+ *if and only if the announcement premium on any risky asset is non-negative.*

Proof: *If* $m_t^*\left(V_t^+\right)$ *is a non-decreasing function, then for any non-decreasing function* f*,*

$$E\left[m_t^*\left(V_t^+\right)f\left(V_t^+\right)\right]\leq E\left[m_t^*\left(V_t^+\right)\right]E\left[f\left(V_t^+\right)\right]=E\left[f\left(V_t^+\right)\right],$$

because $m_t^*\left(V_t^+\right)$ *and* $f\left(V_t^+\right)$ *is negatively correlated. Conversely, if* $m_t^*(v_1)<m_t^*(v_2)$ *for*

[21] Alternatively, we can define any payoff to be risky if its projection onto $L^2\left(\Omega,\sigma\left(V_t^+\right),P\right)$ is represented by a f function that satisfy the same properties.

some $v_1 < v_2$, both of which occur with positive probability, then the following payoff is risky:

$$f(v) = \begin{cases} v & v \in \{v_1, v_2\} \\ 0 & v \neq v_1, v_2 \end{cases},$$

and yet $E\left[m_t^*\left(V_t^+\right) f\left(V_t^+\right)\right] > E\left[f\left(V_t^+\right)\right]$, contradicting a non-negative premium for $f\left(V_t^+\right)$.

C Discussion of Theorem 2

C.1 Examples of uncertainty averse preferences

In this section, we show that most of the non-expected utility proposed in the literature can be represented in the form of (11).

- The recursive utility of Kreps and Porteus [28] and Epstein and Zin [17]. The recursive preference be generally represented as:

$$U_t = u^{-1}\left\{(1-\beta)\, u(C_t) + \beta\phi \circ h^{-1} E\left[h\left(U_{t+1}\right)\right]\right\}. \tag{52}$$

  For example, the well-known recursive preference with constant IES and constant risk aversion is the special case in which $u(C) = \frac{1}{1-1/\psi} C^{1-1/\psi}$ and $h(U) = \frac{1}{1-\gamma} C^{1-\gamma}$. With a monotonic transformation,

$$V = u(U), \tag{53}$$

  then the recursive relationship for V can be written in the form of (11) with the same u function and certainty equivalence functional:

$$\mathcal{I}(V) = \phi \circ h^{-1}\left(\int h \circ \phi^{-1}(V)\, dP\right).$$

  Denoting $f = h \circ u^{-1}$, the SDF for pure news events can be written as:

$$m^*(V) \propto f'(V), \tag{54}$$

  where we suppress the normalizing constant, which is chosen so that $m^*(V)$ integrates to one.

- The maxmin expected utility of Gilboa and Schmeidler [19]. The dynamic version of this preference is studied in Epstein and Schneider [16] and Chen and Epstein [10]. This

preference can be represented as the special case of (11) where the certainty equivalence functional is of the form:

$$\mathcal{I}(V) = \min_{m \in M} \int mV dP,$$

where M is a family of probability densities that is assumed to be closed in the weak* topology. As we show in Section 3.3 of the paper, the SDF for pure news events for this class of preference is the Radon-Nikodym derivative of the minimizing probability measure with respect to P.

- The variational preferences of Maccheroni, Marinacci, and Rustichini [33], the dynamic version of which is studied in Maccheroni, Marinacci, and Rustichini [34], features a certainty equivalence functional of the form:

$$\mathcal{I}(V) = \min_{E[m]=1} \int mV dP + c(m),$$

where $c(\pi)$ is a convex and weak$^*-$lower semi-continuous function. Similar to the maxmin expected utility, the SDF for pure news events for this class of preference is minimizing probability density.

- The multiplier preferences of Hansen and Sargent [21] and Strzaleck [40] is represented by the certainty equivalence functional:

$$\mathcal{I}(V) = \min_{E[m]=1} \int mV dP + \theta R(m),$$

where $R(m)$ denote the relative entropy of the density m with respect to the reference probability measure P, and $\theta > 0$ is a parameter. In this case, the SDF for pure news events is also the minimizing probability that can be written as a function of the continuation utility: $m^*(V) \propto e^{-\frac{1}{\theta}V}$.

- The second order expected utility of Ergin and Gul [18] can be written as (11) with the following choice of $\mathcal{I}$:

$$\mathcal{I}(V) = \phi^{-1}\left(\int \phi(V)\, dP\right),$$

where ϕ is a concave function. In this case, the SDF for pure news events can be written as a function of continuation utility:

$$m^*(V) \propto \phi'(V).$$

- The smooth ambiguity preferences of Klibanoff, Marinacci, and Mukerji [26] and Klibanoff, Marinacci, and Mukerji [27] can be represented as:

$$\mathcal{I}(V) = \phi^{-1}\left(\int_M \phi\left(\int_\Omega mV dP\right) d\mu(m)\right), \tag{55}$$

where μ is a probability measure on a set of probabilities densities M. The SDF for pure news events can be written as a function of V:

$$m^* \propto \int_M \phi'\left(\int mV dP\right) m d\mu(m). \tag{56}$$

C.2 Details of Example 1

Under the representation (55), the set of probabilities in Δ is identified by their corresponding density. We denote the density of P_x to be m_x. Therefore, $m_x(\ln C) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(\ln C - \mu - x)^2}$. Also, under the P measure, the density of $\ln C$ is $\frac{m_x(\ln C)}{\int m_x(\ln C)\mu(dx)}$. Using (56),

$$m^* \propto \frac{\int_{-\infty}^{\infty} \phi'\left(E^x\left[\frac{1}{1-\gamma}C^{1-\gamma}\right]\right) m_x(\ln C)\,\mu(dx)}{\int m_x(\ln C)\,\mu(dx)}. \tag{57}$$

Using the property of log-normal distribution,

$$E^x\left[\frac{1}{1-\gamma}C^{1-\gamma}\right] = \frac{1}{1-\gamma} e^{(1-\gamma)(\mu+x)+\frac{1}{2}(1-\gamma)^2\sigma^2}.$$

Therefore,

$$\phi'\left(E^x\left[\frac{1}{1-\gamma}C^{1-\gamma}\right]\right) = \left[(1-\gamma)\,E^x\left[\frac{1}{1-\gamma}C^{1-\gamma}\right]\right]^{\frac{\gamma-\eta}{1-\gamma}} = e^{(\gamma-\eta)\left[(\mu+x)+\frac{1}{2}(1-\gamma)\sigma^2\right]}.$$

Also, we can simplify the term $m_x(\ln C)\,\mu(dx)$ as

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(c-\mu-x)^2} \times \frac{1}{\sqrt{2\pi}\tau} e^{-\frac{1}{2\tau^2}x^2} dx \\ &= \frac{1}{\sqrt{2\pi}\rho} e^{-\frac{1}{2\rho^2}\left(x - \frac{\tau^2}{\sigma^2+\tau^2}(\ln C - \mu)\right)^2} \times \frac{1}{\sqrt{2\pi\left(\tau^2+\sigma^2\right)}} e^{-\frac{1}{2(\tau^2+\sigma^2)}(\ln C - \mu)^2} dx, \end{aligned}$$

with $\rho^2 = \frac{1}{\frac{1}{\sigma^2}+\frac{1}{\tau^2}}$. Note that the first part is a normal density for x with mean $\frac{\tau^2}{\sigma^2+\tau^2}(\ln C - \mu)$ and variance ρ^2, and the second part is a normal density of $\ln c$ with $N(\mu, \tau^2+\sigma^2)$.

Now we can compute the numerator of (57) as:

$$\begin{aligned}
&\int_{-\infty}^{\infty} e^{(\gamma-\eta)\left[(\mu+x)+\frac{1}{2}(1-\gamma)\sigma^2\right]}\frac{1}{\sqrt{2\pi}\rho}e^{-\frac{1}{2\rho^2}\left(x-\frac{\tau^2}{\sigma^2+\tau^2}(\ln C-\mu)\right)^2}dx\frac{1}{\sqrt{2\pi\left(\tau^2+\sigma^2\right)}}e^{-\frac{1}{2(\tau^2+\sigma^2)}(\ln C-\mu)^2}\\
= \;& e^{(\gamma-\eta)\frac{\tau^2}{\sigma^2+\tau^2}(\ln C-\mu)+\frac{1}{2}(\gamma-\eta)^2\rho^2+(\gamma-\eta)\left[\mu+\frac{1}{2}(1-\gamma)\sigma^2\right]}\times\frac{1}{\sqrt{2\pi\left(\tau^2+\sigma^2\right)}}e^{-\frac{1}{2(\tau^2+\sigma^2)}(\ln C-\mu)^2}
\end{aligned}$$

Also,

$$\int m_x\left(\ln C\right)\mu\left(dx\right)=\frac{1}{\sqrt{2\pi\left(\tau^2+\sigma^2\right)}}e^{-\frac{1}{2(\tau^2+\sigma^2)}(\ln C-\mu)^2}$$

Therefore,

$$m^*\propto e^{(\gamma-\eta)\frac{\tau^2}{\sigma^2+\tau^2}\ln C}.$$

We can represent m^* as a function of V :

$$m^*\left(V\right)\propto\left[\left(1-\gamma\right)V\right]^{\frac{\gamma-\eta}{1-\gamma}\frac{\tau^2}{\sigma^2+\tau^2}}.$$

As long as $\gamma<\eta$, m^* assigns a higher weight to low utility states.

C.3 Details of Example 2

Here we provide the details of the derivation of equation (18). Consider a dynamic economy as we described in Section 3.4. Suppose the conditional distribution of the growth rate aggregate endowment, $\frac{Y_{t+1}}{Y_t}$ is summarized by a vector of state variables x_t.[22] It is convenient to define U_t as the normalization of utility that is homogenous of degree one in consumption as in (52). Homogeneity implies that the post-news utility of the representative agent, U_t^+ must be of the form $U_t^+=u_t^+\left(x_t\right)Y_t$ for some function $u_t^+\left(\cdot\right)$. The following lemma express the wealth-to-consumption ratio as a function of $u_t^+\left(x\right)$:

Lemma 8. *After the date-t announcement at time t^+, the wealth-to-consumption ratio of the representative agent is $\frac{1}{1-\beta}u_t^+\left(x_t^+\right)^{1-1/\psi}$.*

Proof: *Let $\bar{U}_t^+\left(x,W\right)$ denote the value function of the representative agent in the competitive equilibrium described in Section 3.4 of the paper. Note that $\bar{U}_t^+\left(x,W\right)$ satisfy the following recursion:*

$$\bar{U}_t^+\left(x_t^+,W_t^+\right)=\left[\left(1-\beta\right)C_t^{1-\frac{1}{\psi}}+\beta\left(E\left[\bar{U}_{t+1}^-\left(x_{t+1}^-,W_{t+1}^-\right)^{1-\gamma}\middle|x_t^+\right]\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}.$$

[22] Under Bayes learning, x_t will be a function of the histories of both Y_t and s_t.

The envelop condition implies

$$\frac{\partial}{\partial W}\bar{U}_t^+\left(x_t^+, W_t^+\right) = (1-\beta)\left[\frac{C_t}{\bar{U}_t^+\left(x_t^+, W_t^+\right)}\right]^{-\frac{1}{\psi}}.$$

Homogeneity of the value function implies

$$\frac{\partial}{\partial W}\bar{U}_t^+\left(x_t^+, W_t^+\right) = \frac{\bar{U}_t^+\left(x_t^+, W_t^+\right)}{W_t^+}.$$

In equilibrium, $U_t^+\left(x_t^+, Y_t\right) = \bar{U}_t^+\left(x_t^+, W_t^+\right)$*; therefore,*

$$(1-\beta)\left[\frac{Y_t}{U_t\left(x_t^+, Y_t\right)}\right]^{-\frac{1}{\psi}} = \frac{U_t\left(x_t^+, Y_t\right)}{W_t^+},$$

which gives

$$\frac{W_t^+}{Y_t} = \frac{1}{1-\beta}u_t^+\left(x_t^+\right)^{1-1/\psi}.$$

Using the result from (54), the pre-announcement aggregate wealth can be computed as:

$$W_t^- = \frac{E_t^-\left[\left(\left(1-\frac{1}{\psi}\right)V_t^+\right)^{\frac{1/\psi-\gamma}{1-1/\psi}}W_t^+\right]}{E_t^-\left[\left(\left(1-\frac{1}{\psi}\right)V_t^+\right)^{\frac{1/\psi-\gamma}{1-1/\psi}}\right]},$$

where V_t^+ and $U_t^+\left(x_t^+, Y_t\right)$ are related through the monotonic transformation (53). Denote $w_t^+ = \frac{W_t^+}{Y_t}$ as the wealth-to-consumption ratio after announcement, based on the result of Lemma 8, we have:

$$W_t^- = \frac{E_t^-\left[\left[(1-\beta)\,w_t^+ Y_t^{1-\frac{1}{\psi}}\right]^{\frac{1/\psi-\gamma}{1-1/\psi}}w_t^+ Y_t\right]}{E_t^-\left[\left[(1-\beta)\,w_t^+ Y_t^{1-\frac{1}{\psi}}\right]^{\frac{1/\psi-\gamma}{1-1/\psi}}\right]}.$$

Therefore, the expected announcement return on aggregate wealth, $E\left[R_{W,t}^+\right]$, can be written

as:

$$\begin{aligned}\frac{E\left[W_t^+\right]}{W_t^-} &= \frac{E_t^-\left[w_t^+ Y_t\right] E_t^-\left[\left[\left(1-\beta\right) {w_t^+}^{1-\frac{1}{\psi}} {Y_t}^{1-\frac{1}{\psi}}\right]^{\frac{1/\psi-\gamma}{1-1/\psi}}\right]}{E_t^-\left[\left[\left(1-\beta\right) {w_t^+}^{1-\frac{1}{\psi}} {Y_t}^{1-\frac{1}{\psi}}\right]^{\frac{1/\psi-\gamma}{1-1/\psi}} w_t^+ Y_t\right]} \\ &= \frac{E_t^-\left[w_t^+\right] E_t^-\left[\left(w_t^+\right)^{\frac{1/\psi-\gamma}{1-1/\psi}}\right]}{E_t^-\left[\left(w_t^+\right)^{\frac{1-\gamma}{1-1/\psi}}\right]}\end{aligned}$$

Assuming that the log return on aggregate wealth, $\ln R_{W,t}^+$ is normally distributed, we have

$$\ln E\left[R_{W,t}^+\right] = \frac{\gamma - \frac{1}{\psi}}{1-\frac{1}{\psi}} Var\left[\ln R_{W,t}^+\right].$$

C.4 Details of Example 3

In this section, we provide more examples of time non-separable preferences. For simplicity, we focus on a two-period setup and consider utility functions of the following form:

$$u\left(C_0\right) + \beta E\left[u\left(C_1 + \alpha C_0\right)\right], \quad \alpha \in (0,1). \tag{58}$$

With $\alpha < 0$, this is the internal habit model (for example, Boldrin, Christiano, and Fisher [6]). We show in this case the announcement premium is negative. The class of preferences with $\alpha > 0$ is discussed in Dunn and Singleton [13] and Heaton [23]. In general, the premium for pure news event can be positive or negative with $\alpha > 0$.

If the representative agent's preference is given by (58), the pre-announcement price is:

$$P^-\left(X\right) = E\left[\frac{\beta u'\left(C_1 + \alpha C_0\right)}{u'\left(C_0\right) + \alpha\beta E\left[u'\left(C_1 + \alpha C_0\right)\right]} X\right].$$

The post-announcement price can be written as:

$$P^+\left(X\right) = \frac{\beta u'\left(C_1 + \alpha C_0\right)}{u'\left(C_0\right) + \alpha\beta u'\left(C_1 + \alpha C_0\right)} X.$$

Therefore, the expected announcement return is

$$\frac{E\left[P^{+}\left(X\right)\right]}{P^{-}\left(X\right)}=\frac{E\left[\frac{\beta u'(C_1+\alpha C_0)}{u'(C_0)+\alpha\beta u'(C_1+\alpha C_0)}X\right]}{\frac{E[\beta u'(C_1+\alpha C_0)X]}{u'(C_0)+\alpha\beta E[u'(C_1+\alpha C_0)]}}.$$

Note that

$$E\left[\frac{\beta u'\left(C_1+\alpha C_0\right)}{u'\left(C_0\right)+\alpha\beta u'\left(C_1+\alpha C_0\right)}X\right]=$$
$$\frac{E\left[\beta u'\left(C_1+\alpha C_0\right)X\right]}{u'\left(C_0\right)+\alpha\beta E\left[u'\left(C_1+\alpha C_0\right)\right]}+Cov\left\{\beta u'\left(C_1+\alpha C_0\right)X,\ \frac{1}{u'\left(C_0\right)+\alpha\beta u'\left(C_1+\alpha C_0\right)}\right\}.$$

We make the following observations.

1. The internal habit model: First, assume $\alpha<0$. This is the internal habit case. Because $\frac{1}{u'(C_0)+\alpha\beta u'(C_1+\alpha C_0)}$ decreases with C_1 for $\alpha<0$, the fact that that $P^{+}\left(X\right)$ is an increasing function of C_1 implies that $\beta u'\left(C_1+\alpha C_0\right)X$ must be increasing in C_1. Under the assumption $\alpha<0$ and that u is strictly concave, $\frac{1}{u'(C_0)+\alpha\beta u'(C_1+\alpha C_0)}$ decreases with C_1. As a result, $Cov\left\{\beta u'\left(C_1+\alpha C_0\right)X,\ \frac{1}{u'(C_0)+\alpha\beta u'(C_1+\alpha C_0)}\right\}<0$ and $\frac{E\left[P^{+}(X)\right]}{P^{-}(X)}<1$. That is, the premium for pure news events is negative.

2. Durable consumption goods: with $\alpha>0$, (58) is a special case of the durable consumption goods model of Dunn and Singleton [13]. In this case, $\frac{1}{u'(C_0)+\alpha\beta u'(C_1+\alpha C_0)}$ is an increasing function of C_1. Here the assumption that $P^{+}\left(X\right)$ is an increasing function of C_1 is not sufficient for the monotonicity of $\beta u'\left(C_1+\alpha C_0\right)X$ with respect to C_1. We have two cases.

   - Case 1: The term $\beta u'\left(C_1+\alpha C_0\right)X$ increases with C_1. In this case, $Cov\left\{\beta u'\left(C_1+\alpha C_0\right)X,\ \frac{1}{u'(C_0)+\alpha\beta u'(C_1+\alpha C_0)}\right\}>0$, and $\frac{E\left[P^{+}(X)\right]}{P^{-}(X)}>1$. This implies that the premium for pure news events is positive. An example of payoff that satisfies this condition is $X=\frac{C_1}{u'(C_1+\alpha C_0)}$.
   - Case 2: The term $\beta u'\left(C_1+\alpha C_0\right)X$ decreases with date-1 consumption, C_1. In this case $Cov\left\{\beta u'\left(C_1+\alpha C_0\right)X,\ \frac{1}{u'(C_0)+\alpha\beta u'(C_1+\alpha C_0)}\right\}<0$, and we have $\frac{E\left[P^{+}(X)\right]}{P^{-}(X)}<1$. That is, a negative premium for pure-news events. An example of payoff that satisfies this condition is $X=\frac{\sqrt{u'(C_0)+\alpha\beta u'(C_1+\alpha C_0)}}{u'(C_1+\alpha C_0)}$.

D Details of the Continuous-time model

D.1 Asset Pricing in the Learning Model

Value function of the representative agent In the interior of $(nT, (n+1)T)$, standard optimal filtering implies that the posterior mean and variance of x_t are given by equations (21) and (22). The posterior variance q_t has a closed form solution:

$$q_t = \frac{\sigma_x^2 \left(1 - e^{-2\hat{a}(t+t^*-nT)}\right)}{(\hat{a}-a)\, e^{-2\hat{a}(t+t^*-nT)} + a + \hat{a}}, \tag{59}$$

where $\hat{a}$ and t^* are defined as:

$$\hat{a} = \sqrt{a^2 + (\sigma_x/\sigma)^2}; \qquad t^* = \frac{1}{2\hat{a}} \ln \frac{\sigma_x^2 + (\hat{a}-a)\, q_{nT}^+}{\sigma_x^2 - (\hat{a}+a)\, q_{nT}^+}.$$

On the boundaries, q_{nT}^- and q_{nT}^+ satisfy equation (23):

$$\frac{1}{q_{nT}^+} = \frac{1}{\sigma_S^2} + \frac{1}{q_{nT}^-}. \tag{60}$$

Given a q_0, equations (59) and (60) completely determine q_t as a function of t. In calibrations, we focus on the steady state where $q(t) = q(t \bmod T)$ and adopt the convention $q(0) = q_{nT}^+$ and $q(T) = q_{nT}^-$, for $n = 1, 2, \cdots$.

Using the results from Duffie and Epstein [12], the representative consumer's preference is specified by a pair of aggregators $(f, \mathcal{A})$ such that the utility of the representative agent is the solution to the following stochastic differential equation (SDU):

$$d\bar{V}_t = [-f(C_t, \bar{V}_t) - \frac{1}{2}\mathcal{A}(V_t)||\sigma_V(t)||^2]dt + \sigma_V(t)dB_t,$$

for some square-integrable process $\sigma_V(t)$. We adopt the convenient normalization $\mathcal{A}(v) = 0$ (Duffie and Epstein [12]), and denote $\bar{f}$ the normalized aggregator. Under this normalization, $\bar{f}(C, V)$ is:

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

Due to homogeneity, the value function is of the form

$$\bar{V}(\hat{x}_t, t, C_t) = \frac{1}{1-\gamma} H(\hat{x}_t, t)\, C_t^{1-\gamma}, \tag{61}$$

where $H(\hat{x}_t, t)$ satisfies the following Hamilton–Jacobi–Bellman (HJB) equation:

$$-\frac{\rho}{1-\gamma}\ln H(\hat{x},t)\,H(x,t)+\left(\hat{x}-\frac{1}{2}\gamma\sigma^2\right)H(\hat{x},t)+\frac{1}{1-\gamma}H_t(\hat{x}_t,t)$$
$$+\left[\frac{1}{1-\gamma}a_x(\bar{x}-\hat{x})+q_t\right]H_x(\hat{x},t)+\frac{1}{2}\frac{1}{1-\gamma}H_{xx}(\hat{x},t)\frac{q_t^2}{\sigma^2}=0, \tag{62}$$

with the boundary condition that for all $n=1,2\cdots$

$$H\left(\hat{x}_{nT}^{-},nT\right)=E\left[H\left(\hat{x}_{nT}^{+},nT\right)\middle|\,\hat{x}_{nT}^{-},q_{nT}^{-}\right]. \tag{63}$$

The value function that has the representation (24) is a monotonic transformation of $\bar{V}$: $V_t=\frac{1}{1-\gamma}\ln\left[(1-\gamma)\bar{V}_t\right]$.

The solution to the partial differential equation (PDE) (62) together with the boundary condition (63) is separable and given by:

$$H(\hat{x},t)=e^{\frac{1-\gamma}{a_x+\rho}\hat{x}+h(t)},$$

where $h(t)$ satisfy the following ODE:

$$-\rho h(t)+h'(t)+f(t)=0, \tag{64}$$

where $f(t)$ is defined as:

$$f(t)=\frac{(1-\gamma)^2}{a_x+\rho}q(t)+\frac{1}{2}\frac{(1-\gamma)^2}{(a_x+\rho)^2}\frac{1}{\sigma^2}q^2(t)-\frac{1}{2}\gamma(1-\gamma)\sigma^2+a_x\bar{x}\frac{1-\gamma}{a_x+\rho}.$$

The general solution to (64) is of the form:

$$h(t)=h(0)\,e^{\rho t}-e^{\rho t}\int_0^t e^{-\rho s}f(s)\,ds.$$

We focus on the steady state in which $h(t)=h(t \bmod T)$ and use the convention $h(0)=h(0^+)$ and $h(T)=h(T^-)$. Under these notations, the boundary condition (63) implies $h(T)=h(0)+\frac{1}{2}\left(\frac{1-\gamma}{a_x+\rho}\right)^2[q(T)-q(0)]$.

Asset prices In the interior of $(nT,(n+1)T)$, the law of motion of the state price density, π_t satisfies the stochastic differential equation of the form:

$$d\pi_t=\pi_t\left[-r(\hat{x}_t,t)\,dt-\sigma_\pi(t)\,d\tilde{B}_{C,t}\right],$$

where

$$r\left(\hat{x},t\right)=\beta+\hat{x}-\gamma\sigma^{2}+\frac{1-\gamma}{a_{x}+\rho}q_{t}$$

is the risk-free interest rate, and

$$\sigma_{\pi}\left(t\right)=\gamma\sigma+\frac{\gamma-1}{a_{x}+\rho}\frac{q_{t}}{\sigma}$$

is the market price of the Brownian motion risk.

For $t\in\left(nT,\left(n+1\right)T\right)$, the price of the claim to the dividend process can then be calculated as:

$$p\left(\hat{x}_{t},t\right)D_{t}=E_{t}\left[\int_{t}^{(n+1)T}\frac{\pi_{s}}{\pi_{t}}D_{s}ds+\frac{\pi_{(n+1)T}}{\pi_{t}}p\left(\hat{x}_{(n+1)T}^{-},\left(n+1\right)T^{-}\right)D_{(n+1)T}\right].$$

The above present value relationship implies that

$$\pi_{t}D_{t}+\lim_{\Delta\to0}\frac{1}{\Delta}\left\{E_{t}\left[\pi_{t+\Delta}p\left(\hat{x}_{t+\Delta},t+\Delta\right)D_{t+\Delta}\right]-\pi_{t}p\left(\hat{x}_{t},t\right)D_{t}\right\}=0. \tag{65}$$

Equation (65) can be used to show that the price-to-dividend ration function must satisfy the following PDE:

$$1-p\left(\hat{x},t\right)\varpi\left(\hat{x},t\right)+p_{t}\left(\hat{x},t\right)-p_{x}\left(\hat{x},t\right)\nu\left(\hat{x},t\right)+\frac{1}{2}p_{xx}\left(\hat{x},t\right)\frac{q^{2}\left(t\right)}{\sigma^{2}}=0, \tag{66}$$

where the functions $\varpi\left(\hat{x},t\right)$ and $\nu\left(\hat{x},t\right)$ are defined by:

$$\begin{aligned}\varpi\left(\hat{x},t\right) &= \rho-\mu+\phi\bar{x}+\left(1-\phi\right)\hat{x}+\left(\phi-1\right)\left[\gamma\sigma^{2}+\frac{1-\gamma}{a_{x}+\rho}q\left(t\right)\right]\\ \nu\left(\hat{x},t\right) &= a_{x}\left(\hat{x}-\bar{x}\right)+\left(\gamma-\phi\right)q\left(t\right)+\frac{1-\gamma}{a_{x}+\rho}\left(\frac{q\left(t\right)}{\sigma}\right)^{2}.\end{aligned}$$

Alos, equation (35) can be used to derive the following boundary condition for $p\left(\hat{x},t\right)$:

$$p\left(\hat{x}_{T}^{-},T^{-}\right)=\frac{E\left[e^{\frac{1-\gamma}{a_{x}+\rho}\hat{x}_{T}^{+}}p\left(\hat{x}_{T}^{+},T^{+}\right)\middle|\,\hat{x}_{T}^{-},q_{T}^{-}\right]}{e^{\frac{1-\gamma}{a_{x}+\rho}\hat{x}_{T}^{-}+\frac{1}{2}\left(\frac{1-\gamma}{a_{x}+\rho}\right)^{2}\left[q_{T}^{-}-q_{T}^{+}\right]}}. \tag{67}$$

Again, we focus on the steady-state and denote $p\left(\hat{x},0\right)=p\left(\hat{x},nT^{+}\right)$, and $p\left(\hat{x},T\right)=p\left(\hat{x},nT^{-}\right)$. Under this condition PDE (66) together with the boundary condition can be used to determined the price-to-dividend ratio function.

We define $\mu_{R,t}$ to the instantaneous risk premium, that is,

$$\mu_{R,t} = \frac{1}{p\left(\hat{x}_t, t\right) D_t} \left\{ D_t dt + E_t \; d\left[p\left(\hat{x}_t, t\right) D_t \right] \right\}.$$

Standard results implies that in the interior of $(nT, (n+1)T)$, the instantaneous risk premium is given by:

$$\mu_{R,t} - r\left(\hat{x}, t\right) = \gamma\sigma^2 + \left[\frac{\gamma p_x\left(\hat{x}_t, t\right)}{p\left(\hat{x}_t, t\right)} + \frac{\gamma - 1}{a_x + \rho} \right] q\left(t\right) + \frac{\gamma - 1}{a_x + \rho} \frac{p_x\left(\hat{x}_t, t\right)}{p\left(\hat{x}_t, t\right)} \frac{q^2\left(t\right)}{\sigma^2}.$$

D.2 Numerical Solutions

To solve the PDE (66) with the boundary condition (67), we consider the following auxiliary problem:

$$p\left(x_t, t\right) = E\left[\int_t^T e^{-\int_t^s \varpi(x_u, u) du} ds + e^{-\int_t^T \varpi(x_u, u) du} p\left(x_T, T\right) \right], \tag{68}$$

where the state variable x_t follows the law of motion;

$$dx_t = -\nu\left(\hat{x}, t\right) dt + \frac{q\left(t\right)}{\sigma} dB_t. \tag{69}$$

Note that the solution to (68) and (67) satisfies the same PDE. Given an initial guess of the pre-new price-to-dividend ratio, $p^-\left(x_\tau, \tau\right)$, we can solve (68) by the Markov chain approximation method (Kushner and Dupuis [29]):

1. We first start with an initial guess of a pre-news price to dividend ratio function, $p\left(x_T, T\right)$.

2. We construct a locally consistent Markov chain approximation of of the diffusion process (69) as follows. We choose a small dx, let $Q = \left|\nu\left(\hat{x}, t\right)\right| dx + \left(\frac{q(t)}{\sigma}\right)^2$, and define the time increment $\Delta = \frac{dx^2}{Q}$ be a function of dx. Define the following Markov chain on the space of x:

$$\begin{aligned} \Pr\left(x + dx \,|x\right) &= \frac{1}{Q}\left[-\nu\left(\hat{x}, t\right)^+ dx + \frac{1}{2}\left(\frac{q\left(t\right)}{\sigma}\right)^2 \right], \\ \Pr\left(x - dx \,|x\right) &= \frac{1}{Q}\left[-\nu\left(\hat{x}, t\right)^- dx + \frac{1}{2}\left(\frac{q\left(t\right)}{\sigma}\right)^2 \right]. \end{aligned}$$

   One can verify that as $dx \to 0$, the above Markov chain converges to the diffusion process (69) (In the language of Kushner and Dupuis [29], this is a Markov chain that

is locally consistent with the diffusion process (69)).

3. With the initial guess of $p(x_T, T)$, for $t = T - \Delta$, $T - 2\Delta$, etc, we use the Markov chain approximation to compute the discounted problem in (68) recursively:

$$p(x_t, t) = \Delta + e^{-\varpi(x,t)\Delta} E\left[p(x_{t+\Delta}, t + \Delta)\right],$$

until we obtain $p(x, 0)$.

4. Compute an updated pre-news price to dividend ratio function, $p(x_T, T)$ using (67):

$$p\left(\hat{x}_T^-, T^-\right) = \frac{E\left[e^{\frac{1-\gamma}{a_x+\rho}\hat{x}_T^+} p\left(\hat{x}_T^+, 0\right) \middle| \hat{x}_T^-, q_T^-\right]}{e^{\frac{1-\gamma}{a_x+\rho}\hat{x}_T^- + \frac{1}{2}\left(\frac{1-\gamma}{a_x+\rho}\right)^2 \left[q_T^- - q_T^+\right]}}.$$

Go back to step 1 and iterate until the function $p(x_T, T)$ converges.

D.3 Preference for Early Resolution of Uncertainty

The term structure of timing premium We first compute the term $\int_0^\tau e^{-\rho t} \rho \ln C_t dt$ in equation (33). The law of motion of C_t (equation (19)) implies

$$\ln C_t = \ln C_0 + \int_0^t \left(x_s - \frac{1}{2}\sigma^2\right) ds + \int_0^t \sigma dB_s. \tag{70}$$

Therefore,

$$\begin{aligned} \int_0^\tau e^{-\rho t} \rho \ln C_t dt &= \rho \ln C_0 \times \int_0^\tau e^{-\rho t} dt + \rho \int_0^\tau e^{-\rho t} \int_0^t x_s ds dt \\ &\quad - \frac{1}{2}\rho\sigma^2 \int_0^\tau t e^{-\rho t} dt + \rho \int_0^\tau e^{-\rho t} \int_0^t \sigma dB_s dt. \end{aligned} \tag{71}$$

It is straightforward to show that

$$\begin{aligned} \rho \ln C_0 \times \int_0^\tau e^{-\rho t} dt &= \left(1 - e^{-\rho\tau}\right) \ln C_0, \\ \frac{1}{2}\rho\sigma^2 \int_0^\tau t e^{-\rho t} dt &= \frac{1}{2}\sigma^2 \left[\frac{1}{\rho}\left(1 - e^{-\rho\tau}\right) - \tau e^{-\rho\tau}\right]. \end{aligned} \tag{72}$$

Also, using the stochastic Fubini's theorem (Karatzas and Shreve [24], page 209),

$$\rho \int_0^\tau e^{-\rho t} \int_0^t \sigma dB_s dt = \sigma \int_0^\tau \left(e^{-\rho t} - e^{-\rho\tau}\right) dB_t. \tag{73}$$

The term $\rho \int_0^\tau e^{-\rho t} \int_0^t x_s ds dt$ can be computed as

$$\rho \int_0^\tau e^{-\rho t} \int_0^t x_s ds dt = \int_0^\tau \left(e^{-\rho t} - e^{-\rho \tau}\right) x_t dt = \int_0^\tau e^{-\rho t} x_t dt - e^{-\rho \tau} \int_0^\tau x_t dt. \tag{74}$$

Using the solution for the Ornstein–Uhlenbeck process x_t,

$$x_t = x_0 e^{-a_x t} + \left(1 - e^{-a_x t}\right) \bar{x} + \int_0^t e^{-a_x (t-s)} \sigma_x dB_s, \tag{75}$$

equation (74) can be written as:

$$\begin{aligned}
\int_0^\tau e^{-\rho t} x_t dt &= \int_0^\tau e^{-\rho t} \left[x_0 e^{-a_x t} + \left(1 - e^{-a_x t}\right) \bar{x} + \sigma_x \int_0^t e^{-a_x (t-s)} dB_{x,s} \right] dt \\
&= \frac{(x_0 - \bar{x})}{a_x + \rho} \left[1 - e^{-(a_x + \rho)\tau}\right] + \frac{\bar{x}}{\rho} \left[1 - e^{-\rho \tau}\right] \\
&\quad + \frac{\sigma_x}{a_x + \rho} \left[\int_0^\tau e^{-\rho t} dB_{xt} - e^{-(a_x + \rho)\tau} \int_0^\tau e^{-a_x t} dB_{xt} \right].
\end{aligned}$$

We can similarly compute the term $e^{-\rho \tau} \int_0^\tau x_t dt$ and combine terms to write (74) as:

$$\begin{aligned}
\int_0^\tau \left(e^{-\rho t} - e^{-\rho \tau}\right) dB_t &= (x_0 - \bar{x}) \left[\frac{1}{a_x + \rho} - \frac{1}{a_x} e^{-\rho \tau} + \left(\frac{1}{a_x} - \frac{1}{a_x + \rho} \right) e^{-(a_x + \rho)\tau} \right] \\
&\quad + \sigma_x \int_0^\tau \left[\frac{1}{a_x + \rho} e^{-\rho t} - \frac{1}{a_x} e^{-\rho \tau} + \left(\frac{1}{a_x} - \frac{1}{a_x + \rho} \right) e^{-(a_x + \rho)\tau} e^{\rho t} \right] dB_{x,t} \\
&\quad + \bar{x} \left[\frac{1}{\rho} \left(1 - e^{-\rho \tau}\right) - \tau e^{-\rho \tau} \right].
\end{aligned} \tag{76}$$

Using (72), (73), (74) and (76), we can sum up all terms in (71):

$$\begin{aligned}
\int_0^\tau e^{-\rho t} \rho \ln C_t dt &= \left(1 - e^{-\rho \tau}\right) \ln C_0 \\
&\quad + (x_0 - \bar{x}) \left[\frac{1}{a_x + \rho} - \frac{1}{a_x} e^{-\rho \tau} + \left(\frac{1}{a_x} - \frac{1}{a_x + \rho} \right) e^{-(a_x + \rho)\tau} \right] \\
&\quad + \bar{x} \left[\frac{1}{\rho} \left(1 - e^{-\rho \tau}\right) - \tau e^{-\rho \tau} \right] \\
&\quad + \sigma_x \int_0^\tau \left[\frac{1}{a_x + \rho} e^{-\rho t} - \frac{1}{a_x} e^{-\rho \tau} + \left(\frac{1}{a_x} - \frac{1}{a_x + \rho} \right) e^{-(a_x + \rho)\tau} e^{\rho t} \right] dB_{x,t} \\
&\quad - \frac{1}{2} \sigma^2 \left[\frac{1}{\rho} \left(1 - e^{-\rho \tau}\right) - \tau e^{-\rho \tau} \right] \\
&\quad + \sigma \int_0^\tau \left(e^{-\rho t} - e^{-\rho \tau}\right) dB_t.
\end{aligned} \tag{77}$$

Next, we compute the term $e^{-\rho\tau}V\left(x_\tau,0,C_\tau\right)=e^{-\rho\tau}\left[\frac{1}{a_x+\rho}x_\tau+\frac{1}{1-\gamma}h\left(0\right)+\ln C_\tau\right]$. Using the solution for C_t in (70) and the solution for x_t in (75), we can write the term $e^{-\rho\tau}\frac{1}{a_x+\rho}x_\tau$ as

$$\begin{aligned} e^{-\rho\tau}\frac{1}{a_x+\rho}x_\tau &= \frac{1}{a_x+\rho}e^{-(a_x+\rho)\tau}\left(x_0-\bar{x}\right)+\frac{1}{a_x+\rho}e^{-\rho\tau}\bar{x} \\ &\quad +\frac{\sigma_x}{a_x+\rho}e^{-(a_x+\rho)\tau}\int_0^\tau e^{a_x t}dB_{x,t}, \end{aligned} \tag{78}$$

and $e^{-\rho\tau}\ln C_\tau$ as

$$\begin{aligned} e^{-\rho\tau}\ln C_\tau &= e^{-\rho\tau}\ln C_0-\frac{1}{2}\sigma_x^2e^{-\rho\tau}\tau+\sigma_xe^{-\beta\tau}\int_0^\tau dB_\tau \\ &\quad +\frac{\left(x_0-\bar{x}\right)}{a_x}\left[e^{-\rho\tau}-e^{-(a_x+\rho)\tau}\right]+\bar{x}e^{-\beta\tau}\tau \\ &\quad +\frac{\sigma_x}{a_x}\left[e^{-\rho\tau}\int_0^\tau dB_{xt}-\int_0^\tau e^{-(a_x+\rho)\tau}e^{\rho t}dB_{x,t}\right] \end{aligned} \tag{79}$$

Now we can combine equations (77), (78), and (79) and cancel terms with opposite signs to get:

$$\begin{aligned} W_0^+\left(\tau\right) &= \int_0^\tau e^{-\rho t}\rho\ln C_t dt+e^{-\rho\tau}\left[\frac{1}{a_x+\rho}x_\tau+\frac{1}{1-\gamma}h\left(0\right)+\ln C_\tau\right] \\ &= \ln C_0+\frac{1}{a_x+\rho}x_0+\frac{1}{\rho}\left(1-e^{-\rho\tau}\right)\left[\frac{a_x\bar{x}}{a_x+\rho}-\frac{1}{2}\sigma^2\right]+\frac{1}{1-\gamma}e^{-\rho\tau}h\left(0\right) \\ &\quad +\int_0^\tau e^{-\rho t}\sigma dB_t+\frac{1}{a_x+\rho}\int_0^\tau e^{-\rho t}\sigma_x dB_{x,t} \end{aligned}$$

Note that $\int_0^\tau e^{-\rho t}\sigma dB_t$ and $\frac{1}{a_x+\rho}\int_0^\tau e^{-\rho t}\sigma_x dB_{x,t}$ and independently normally distributed random variables. Assuming the prior distribution of x_0 is $N\left(\hat{x}_0,q_0\right)$, the certainty equivalent of $W_0^+\left(\tau\right)$ can be calculated as:

$$\begin{aligned} \frac{1}{1-\gamma}E\left[e^{(1-\gamma)W_0^+(\tau)}\right] &= \frac{1}{a_x+\rho}\hat{x}_0+\frac{1}{2}\frac{1-\gamma}{\left(a_x+\rho\right)^2}q_0 \\ &\quad +\frac{1}{\rho}\left(1-e^{-\rho\tau}\right)\left[\frac{a_x\bar{x}}{a_x+\rho}-\frac{1}{2}\sigma^2\right]+\frac{1}{1-\gamma}e^{-\rho\tau}h\left(0\right) \\ &\quad +\frac{1}{2}\left(1-\gamma\right)\left[\sigma^2+\frac{\sigma_x^2}{\left(a_x+\rho\right)^2}\right]\frac{1}{2\rho}\left[1-e^{-2\rho\tau}\right] \end{aligned} \tag{80}$$

Because x_0 is known at time 0, the welfare gain of early resolution of uncertainty can be

computed as:

$$\begin{aligned}\ln\frac{1}{1-\lambda(\tau)} &= \frac{1}{\rho}\left(1-e^{-\rho\tau}\right)\left[\frac{a_x\bar{x}}{a_x+\rho}-\frac{1}{2}\sigma^2\right] \\ &\quad +\frac{1}{4\rho}(1-\gamma)\left[\sigma^2+\frac{\sigma_x^2}{(a_x+\rho)^2}\right]\left[1-e^{-2\rho\tau}\right]-\frac{1}{1-\gamma}\left(1-e^{-\rho\tau}\right)h(0)\end{aligned}$$

The term structure of announcement premium After the announcement, there is no uncertainty between time 0 and time τ. Therefore, the state price density is $e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-1}$ and the price of the asset is determined by:

$$p_0^+(\tau)D_0=\int_0^\tau e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-1}D_t dt+e^{-\rho\tau}\left(\frac{C_\tau}{C_0}\right)^{-1}p(x_\tau,0)D_\tau, \tag{81}$$

where $p_0^+(\tau)$ denote the price-to-dividend ratio after announcement, and $p(x_\tau,0)$ is the price-to-dividend ratio function in our benchmark economy (which is the solution to equation (66) along with its boundary conditions). Before announcement, the price-to-dividend ratio, $p_0^-(\tau)$ is determined by equation (35).

Using the law of motion of the dividend process (20), $D_t = e^{\bar{\mu}_d t}C_t^\phi$, where $\bar{\mu}_d = \mu_d - \phi\left[\bar{x}+\frac{1}{2}(\phi-1)\sigma^2\right]$ and equation (81) can be written as:

$$p_0^+(\tau)=\int_0^\tau e^{-(\rho-\bar{\mu}_d)t}\left(\frac{C_t}{C_0}\right)^{\phi-1}dt+p(x_\tau,0)e^{-(\rho-\bar{\mu}_d)\tau}\left(\frac{C_\tau}{C_0}\right)^{\phi-1}. \tag{82}$$

Therefore, the announcement premium, $ER(\tau)$ can be computed as:

$$\ln ER(\tau)=\ln E_0^-\left[p_0^+(\tau)\right]+\ln E_0^-\left[e^{(1-\gamma)W_0^+(\tau)}\right]-\ln E_0^-\left[e^{(1-\gamma)W_0^+(\tau)}p_0^+(\tau)\right].$$

We already provided the solution for $\ln E_0^-\left[e^{(1-\gamma)W_0^+(\tau)}\right]$ in the calculation of certainty equivalent (equation (80)). Below we provide the details of computing $\ln E_0^-\left[e^{(1-\gamma)W_0^+(\tau)}p_0^+(\tau)\right]$. $\ln E_0^-\left[p_0^+(\tau)\right]$ can be obtained as a special case of $\ln E_0^-\left[e^{(1-\gamma)W_0^+(\tau)}p_0^+(\tau)\right]$ with $\gamma=1$.

Using (82), we have:

$$\begin{aligned} & E\left[e^{(1-\gamma)W_0^+(\tau)}p_0^+(\tau)\right] \\ = \; & E\left[\int_0^\tau e^{(1-\gamma)W_0^+(\tau)-(\rho-\bar{\mu}_d)t}\left(\frac{C_t}{C_0}\right)^{\phi-1}dt + p(x_\tau,0)\,e^{(1-\gamma)W_0^+(\tau)-(\rho-\bar{\mu}_d)\tau}\left(\frac{C_\tau}{C_0}\right)^{\phi-1}\right] \\ = \; & E\left[\int_0^\tau e^{(1-\gamma)W_0^+(t)-(\rho-\bar{\mu}_d)t}\left(\frac{C_t}{C_0}\right)^{\phi-1}\xi_t dt + p(x_\tau,0)\,e^{(1-\gamma)W_0^+(\tau)-(\rho-\bar{\mu}_d)\tau}\left(\frac{C_\tau}{C_0}\right)^{\phi-1}\right], \end{aligned}$$

where we define $\xi_t = E_t\left[e^{(1-\gamma)\left[W_0^+(\tau)-W_0^+(t)\right]}\right]$. Note that conditioning on period-t information, $W_0^+(\tau) - W_0^+(t)$ is a normally distributed random variable and therefore $\xi_t = \frac{1}{4\rho}\left[e^{-2\rho t} - e^{-2\rho\tau}\right]\left[(1-\gamma)^2\sigma^2 + \frac{(1-\gamma)^2\sigma_x^2}{(a_x+\rho)^2}\right]$. We define a function $q(x,s)$ such that

$$\begin{aligned} q(x,s) \;=\; & E_s\left[\int_s^\tau e^{(1-\gamma)\left[W_0^+(t)-W_0^+(s)\right]-(\rho-\bar{\mu}_d)(t-s)}\left(\frac{C_t}{C_s}\right)^{\phi-1}\xi_t dt\right] \\ & + E_s\left[p(x_\tau,0)\,e^{(1-\gamma)\left[W_0^+(\tau)-W_0^+(s)\right]-(\rho-\bar{\mu}_d)(\tau-s)}\left(\frac{C_\tau}{C_s}\right)^{\phi-1}\right] \end{aligned}$$

Note that under our definition, $q(x,\tau) = p(x,0)$ and $q(x,0) = E_0^-\left[e^{(1-\gamma)W_0^+(\tau)}p_0^+(\tau)\right]$. Using the martingale method, it is straightforward to show that $q(x,s)$ satisfies the following PDE:

$$\varpi(x,t)\,q(x,t) = \xi(t) + q_t(x,t) + q_x(x,t)\left[a_x(\bar{x}-x) + e^{-\rho t}\frac{1-\gamma}{a_x+\rho}\sigma^2\right] + \frac{1}{2}q_{xx}(x,t)\,\sigma_x^2,$$

where $\varpi(x,t)$ is defined as:

$$\begin{aligned} \varpi(x,t) \;=\; & \rho - \mu_d + \phi\left[\bar{x} + \frac{1}{2}(\phi-1)\sigma^2\right] + x - \phi(x-\bar{x}) \\ & -\frac{1}{2}e^{-2\rho t}\left[(1-\gamma)^2\sigma^2 + \frac{(1-\gamma)^2\sigma_x^2}{(a_x+\rho)^2}\right] - (1-\gamma)(\phi-1)\,e^{-\rho t}\sigma^2. \end{aligned}$$

We can again apply the Markov chain approximation method in Section D.2 to solve the above PDE.

References

[1] Ai, Hengjie, 2007, Information Quality and Long-run Risks: Asset Pricing and Welfare Implications, *Unpublished*.

[2] Ai, Hengjie, 2010, Information Quality and Long-Run Risk: Asset Pricing Implications, *Journal of Finance* LXV, 1333–1367.

[3] Bansal, Ravi, and Bruce N. Lehmann, 1997, Growth-Optimal Portfolio Restrictions on Asset Pricing Models, *Macroeconomic Dynamics* 1, 333–354.

[4] Bansal, Ravi, and Amir Yaron, 2004, Risk for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *The Journal of Finance* 59, 1481–1509.

[5] Barro, Robert, 2009, Rare Disasters, Asset Prices, and Welfare Costs, *American Economic Review* 99, 243–264.

[6] Boldrin, Michele, Lawrence Christiano, and Jonas Fisher, 2001, Habit persistence, asset returns, and the business cycle, *American Economic Review* 91, 149–166.

[7] Breeden, Douglas T., 1979, An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities, *Journal of Financial Economics* 7, 265–296.

[8] Campbell, John Y., and John H. Cochrane, 1999, By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior, *Journal of Political Economy* 107, 205–255.

[9] Chateauneuf, Alain, and Jose Heleno Faro, 2009, Ambiguity through confidence functions, *Journal of Mathematical Economics* 45, 535–558.

[10] Chen, Zengjing, and Larry Epstein, 2002, Ambiguity, Risk, and Asset Returns in Continuous Time, *Econometrica* 70, 1403–1443.

[11] Cieslak, Anna, Adair Morse, and Annette Vissing-Jorgensen, 2015, Stock Returns over the FOMC Cycle, .

[12] Duffie, Darrell, and Larry G. Epstein, 1992, Stochastic Differential Utility, *Econometrica* 60, 353–394.

[13] Dunn, Kenneth, and Kenneth Singleton, 1986, Modeling The Term Structure of Interest Rates under Non-Separable Utility and Durability of Goods, *Journal of Financial Economics* 17, 27–55.

[14] Ellsberg, Daniel, 1961, Risk, Ambiguity, and the Savage Axioms, *The Quarterly Journal of Economics* 75, 643–669.

[15] Epstein, Larry, Emmanuel Farhi, and Tomasz Strzalecki, 2014, How much would you pay to resolve long-run risk?, *American Economic Review* 104, 2680–2697.

[16] Epstein, Larry, and Martin Schneider, 2003, Recursive multiple-priors, *Journal of Economic Theory* 113, 1–31.

[17] Epstein, Larry, and Stanley E. Zin, 1989, Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* 57, 937–969.

[18] Ergin, Haluk, and Faruk Gul, 2009, A Theory of Subjective Compound Lotteries, *Journal of Economic Theory* 144, 899–929.

[19] Gilboa, Itzhak, and David Schmeidler, 1989, Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics* 18, 141–153.

[20] Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of Security Market Data for Models of Dynamic Economies, *Journal of Political Economy* 99, 225–262.

[21] Hansen, Lars Peter, and Thomas Sargent, 2008, *Robustness*. (Princeton, New Jersey: Princeton University Press).

[22] Hansen, Lars Peter, and Thomas J. Sargent, 2001, Robust Control and Model Uncertainty, *American Economic Review, Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association* 91, 60–66.

[23] Heaton, John, 1993, The Interaction between Time-Nonseparable Preferences and Time Aggregation, *Econometrica* 61, 353–385.

[24] Karatzas, Ioannis, and Steven E. Shreve, 1991, *Brownian Motion and Stochastic Calculus, 2nd Ed.* (Springer-Verlag, NY, Inc).

[25] Kechris, Alexander S., 1995, *Classical Descriptive Set Theory*. (Springer-Verlag, NY, Inc).

[26] Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2005, A Smooth Model of Decision Making under Ambiguity, *Econometrica* 73, 1849–1892.

[27] Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2009, A Smooth Model of Decision Making under Ambiguity, *Journal of Economic Theory* 144, 930–976.

[28] Kreps, David M., and Evan L. Porteus, 1978, Temporal Resolution of Uncertainty. and Dynamic Choice Theory, *Econometrica* 46, 185–200.

[29] Kushner, Harold J, and Paul Dupuis, 2001, *Numerical Methods for Stochastic Control Problems in Continuous Time*. (Springer-Verlag, NY, Inc).

[30] Lucas, Robert, 1978, Asset Prices in an Exchange Economy, *Econometrica* 46, 1429–1446.

[31] Lucca, David O., and Emanuel Moench, 2015, The Pre-FOMC Announcement Drift, *The Journal of Finance* LXX.

[32] Luenberger, David G., 1997, *Optimization by Vector Space Methods.* (Wiley-Interscience; 1969 edition).

[33] Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini, 2006a, Ambiguity Aversion, Robustness, and the Variational Representation of Preferences, *Econometrica* 74, 1447–1498.

[34] Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini, 2006b, Dynamic variational preferences, *Journal of Economic Theory* 128, 4–44.

[35] Milgrom, Paul, and Ilya Segal, 2002, Envelope Theorems for Arbitrary Choice Sets, *Econometrica* 70, 583–601.

[36] Rall, Louis B., 1971, *Nonlinear Functional Analysis and Applications.* (Academic Press, New York).

[37] Rosoiu, Alexandru, 2016, Capital Market Reaction to Predetermined Macroeconomic Announcements, *Dissertation, Duke University.*

[38] Rothschild, Michael, and Joseph E. Stiglitz, 1970, Increasing Risk: A Definition, *Journal of Economic Theory* pp. 225–243.

[39] Savor, Pavel, and Mungo Wilson, 2013, How Much Do Investors Care About Macroeconomic Risk? Evidence from Scheduled Economic Announcement, *Journal of Financial and Quantitative Analysis* 48, 343–375.

[40] Strzaleck, Tomasz, 2011, Axiomatic Foundation of Multiplier Preferences, *Econometrica* 79, 47–73.

[41] Strzalecki, Tomasz, 2013, Temporal resolution of uncertainty and recursive models ambiguity aversion, *Econometrica* 81, 1039–1074.

[42] Veronesi, Pietro, 2000, How Does Information Quality Affect Stock Returns?, *Journal of Finance* 55, 807–837.

[43] Werner, Jan, 2009, Risk and Risk Aversion When States of Nature Matter, *Economic Theory* 41, 231–246.

Table 1

Calibrated Parameter Values

Parameter		Value
β	discount rate	0.01
γ	risk aversion	10
$\overline{x}$	average consumption growth	1.8%
σ	volatility of consumption growth	3.0%
a_x	mean reversion in x	0.10
σ_x	volatility of x (long-run risk)	0.26
ϕ	leverage	3
σ_S^2	Noise in announcement	0
$1/T$	Frequency of announcement per year	12

Table 1 presents the calibrated parameters of our learning model.

Table 2

Expected returns in the model with and without learning

	Learning	Observable	Expected Utility
Equity Premium	5.4%	4.58%	0.7%
Average return on annoucement Days	2.53%	0.09%	0.08%
Average risk-free interest rate	1.75%	2.56%	2.92%
Total volatility of equity return	11.18%	11.18%	8.29%
Total volatility on announcement days	6.53%	1.01%	0.33%
Total volatility on non-announcement days	9.05%	11.18%	8.29%
Volatility of risk-free interest rate	0.81%	0.81%	0.68%

Table 2 reports the total return on equity, return on announcement days, and the risk-free interest rate in the model with learning (left panel), those in a model where x_t is fully observable (middle panel), and those in model with learning and with expected utility (right panel). We simulate the model for 160 years and drop the first 100 years to grantee convergence to steady-state. We run 100 such simulations and report the sample average of the moments computed from these simulations.

Table 3

Average Daily Excess Returns around Announcement Days (Basis Points)

1961-2014				
	# Events	t-1	t	t+1
ALL	1624	1.77	11.21	0.84
		(0.66)	(3.99)	(0.27)
EMPL/NFP	648	4.24	6.18	-2.47
		(1.11)	(1.59)	(-0.65)
PPI	527	-3.09	7.61	-1.05
		(-0.73)	(1.57)	(-0.18)
FOMC	168	10.57	33.60	1.10
		(0.99)	(3.68)	(0.10)
GDP	144	-0.08	14.99	13.64
		(-0.01)	(1.50)	(1.32)
ISM	216	-0.60	18.42	5.76
		(-0.08)	(1.86)	(0.68)
NONE	11968		1.27	
			(1.40)	

1997-2014				
	# Events	t-1	t	t+1
ALL	897	-0.23	16.48	0.21
		(-0.05)	(3.76)	(0.05)
EMPL/NFP	216	-1.09	15.44	-7.15
		(-0.12)	(1.77)	(-0.87)
PPI	216	-8.38	7.53	-6.20
		(-1.10)	(0.83)	(-0.68)
FOMC	144	13.62	36.35	-1.34
		(1.11)	(3.49)	(-0.11)
GDP	144	-0.08	14.99	13.64
		(-0.01)	(1.50)	(1.32)
ISM	216	-0.60	18.42	5.76

Table 4

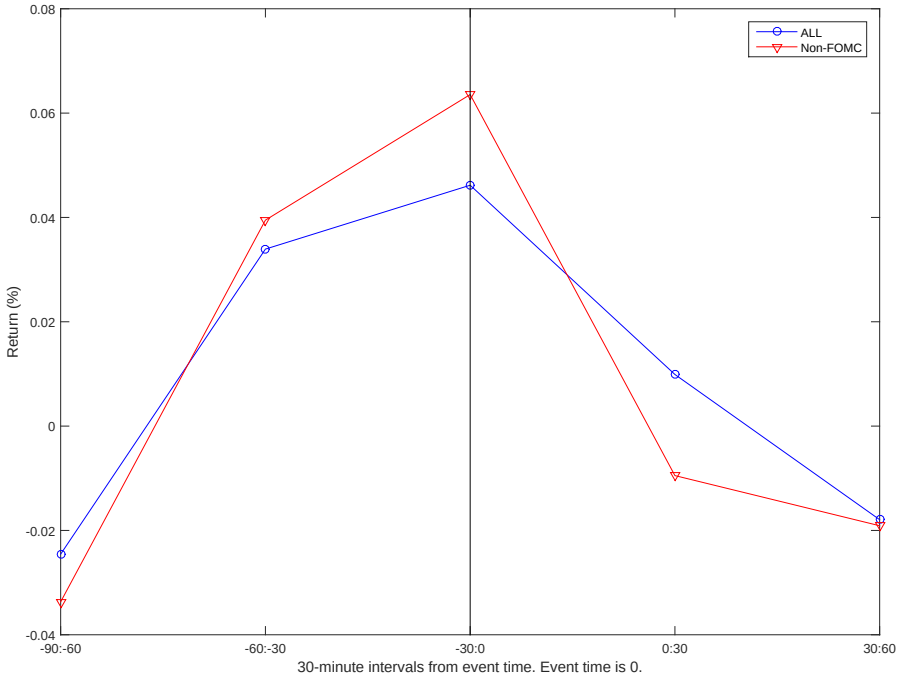
Mean Returns on Event Day (Annualized)

1961-2014					
	# Events per Year	Mean Return	Mean Excess Return	St. Dev. Return	Sharpe Ratio
Market	252	10.96%	6.19%	15.59%	0.40
ALL	30	3.85%	3.36%	6.23%	0.54
NFP/EMPL	12	0.97%	0.74%	3.42%	0.22
PPI	12	1.15%	0.91%	3.84%	0.24
FOMC	8	2.77%	2.69%	3.36%	0.80
GDP	8	1.27%	1.20%	3.41%	0.35
ISM	12	2.32%	2.21%	5.06%	0.44
NONE	222	7.11%	2.82%	14.29%	0.20

1997-2014					
	# Events per Year	Mean Return	Mean Excess Return	St. Dev. Return	Sharpe Ratio
Market	252	9.77%	7.44%	20.14%	0.37
ALL	50	8.70%	8.24%	9.37%	0.88
NFP/EMPL	12	1.96%	1.85%	4.45%	0.42
PPI	12	1.01%	0.90%	4.62%	0.20
FOMC	8	2.98%	2.91%	3.54%	0.82
GDP	8	1.27%	1.20%	3.41%	0.35
ISM	12	2.32%	2.21%	5.06%	0.44
NONE	202	1.08%	-0.78%	17.79%	-0.04

We annualize by multiplying daily returns by the average number of events per year and daily standard deviations by the square root of the same number (including and excluding FOMC announcements). This is consistent with a trading strategy where you are long the market on the announcement dates and do not trade the rest of the days in a given calendar year. The average number of announcements per year is given in column 2. The data sources for each release are described in table 3. The daily excess returns come from Kenneth French's website. The daily return is computed by adding the daily risk free rate to the excess return. Daily risk free rates are also taken from Kenneth French's website and are the simple daily rates that, over the number of trading days in the month, compound to 1-month TBill rates from Ibbotson and Associates Inc.

Figure 1. Average 30-minute returns around the time of macro announcements



The plot shows average returns over 30 minute intervals around the announcement time with all announcements stacked together. 0 is the 30-minute bin containing the announcement. Employment/Non-farm payroll, GDP and PPI announcements are all made at 8:30 AM before the market begins. In these cases we will consider the event return to be the return between 4:00 PM (close of trading) the previous day and 9:30 AM when the market opens on the day of the announcement. ISM announcements are made at 10:00. The event return will cover the interval 9:30 AM to 10:00 AM. Finally the timing of the FOMC statement release varies but we stay consistent with our procedure. For example a statement released at 14:15 will be assigned to the interval 14:00 - 14:30 while a statement made at 14:00 will be assigned to the 13:30 - 14:00 interval. The sample period is 1997-2013.

Figure 2. Consumption Dynamics in the Two-Period Model

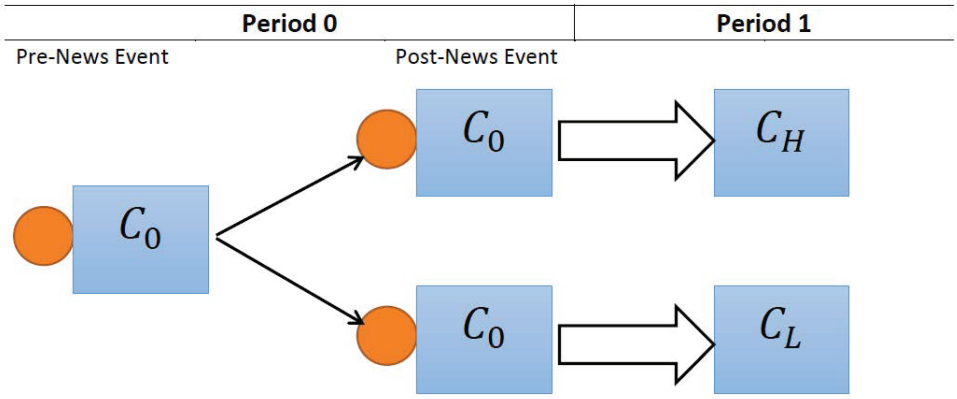


Figure 1: Consumption

Figure 2 This the consumption dynamics in the two-period model.

Figure 3. Asset Prices in the Two-Period Model

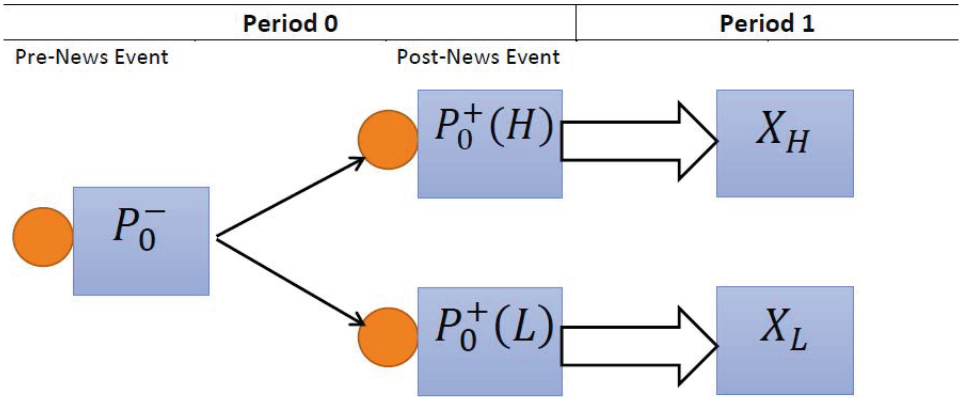


Figure 2: Asset Prices/Payoffs

Figure 3 This figure illustrates the asset prices in the two-period model with a market for pure news events.

Figure 4. Posterior variance of x_t

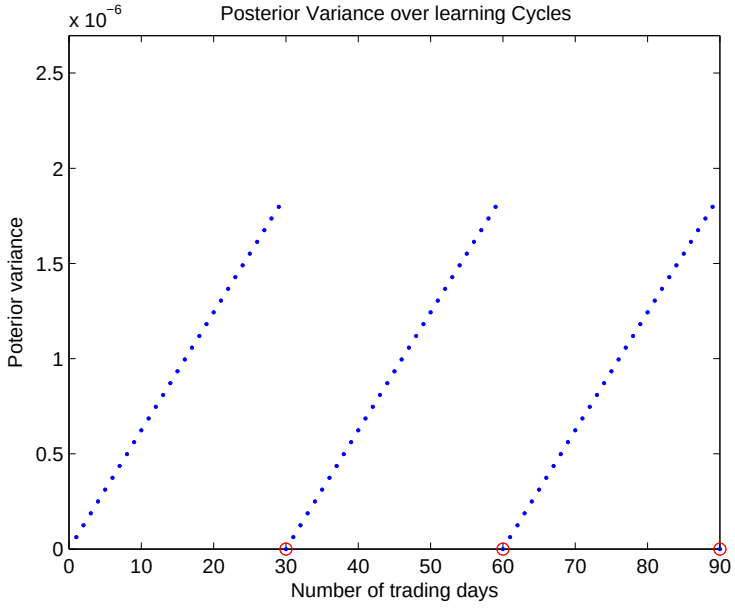


Figure 4 plots the posterior variance of x_t as a function of time.

Figure 5. Equity Premium on Non-Announcement Days

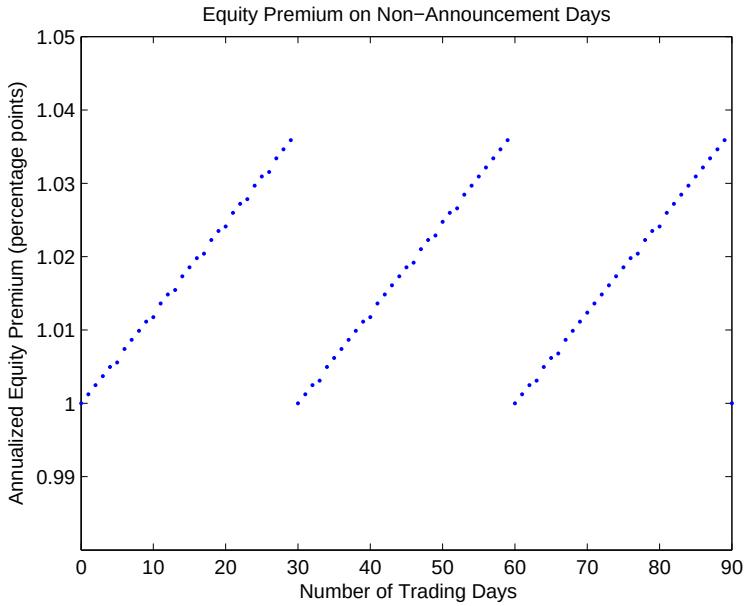


Figure 5 plots the annualized equity premium on non-announcement days in our model with recursive utility.

Figure 6. Announcement Premium for Recursive Utility and Expected Utility

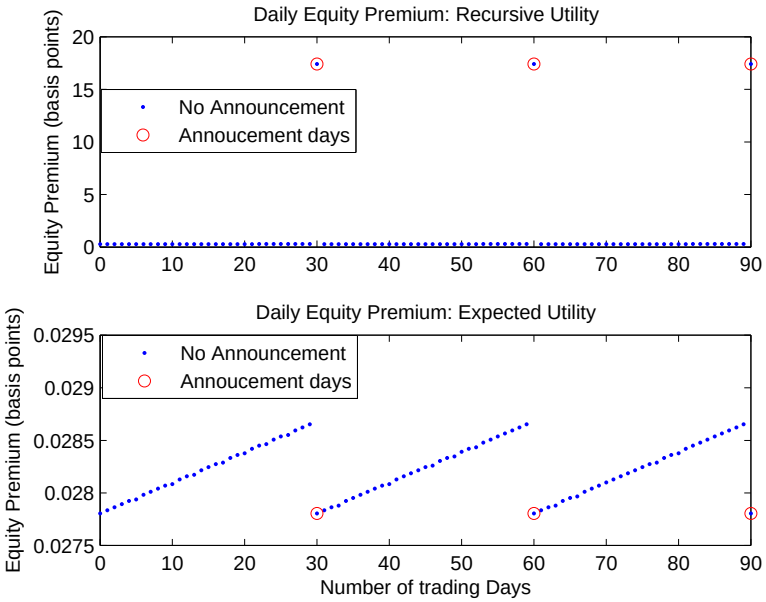


Figure 6 plots daily equity premium for recursive utility (top panel) and that for expected utility (bottom panel). Equity premium is measured in basis points. The dotted line is the equity premium on non-announcement days, and the circles indicate the equity premium on announcement days.

Figure 7. Dynamics of Price-to-Dividend Ratio

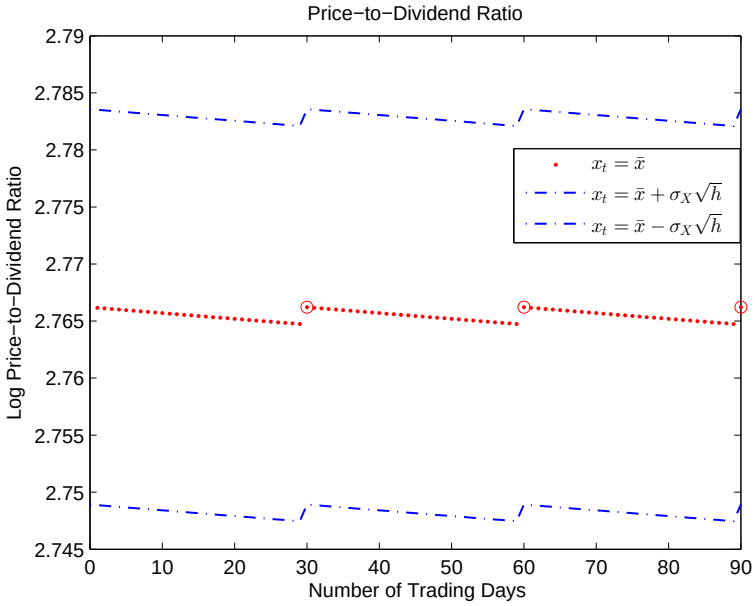


Figure 7 plots the evolution of log price-to-dividend ratio over the announcement cycles. The dotted line is the price-to-dividend ratio evaluated at the steady-state level of expected consumption growth rate, $\bar{x}$. The dash-dotted line is the price-to-dividend ratio for x_t one standard deviation above and below its steady-state level. The standard deviation is calculated at the monthly level, that is, $h = \frac{1}{12}$.

Figure 8. The Term Structure of News Premium and Welfare Gain of Uncertainty

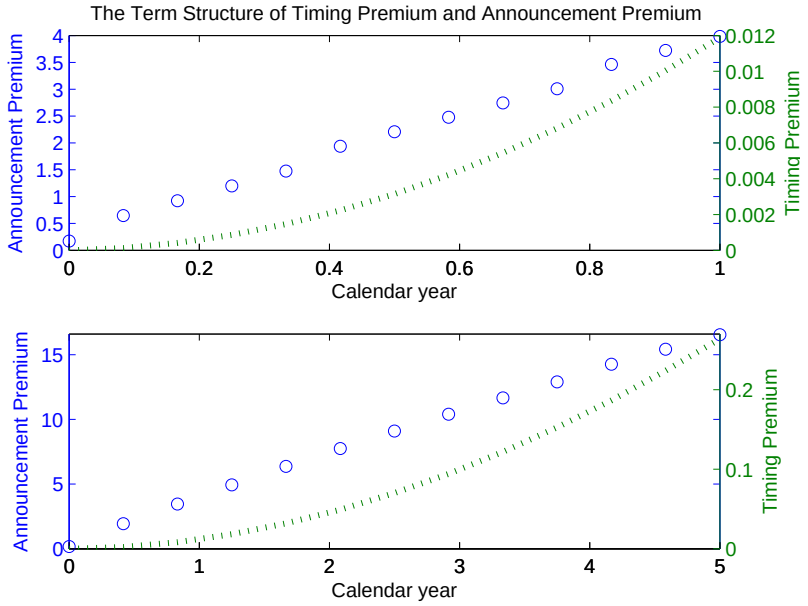


Figure 8 plots the term structure for premium for pure news events (circles, left scale) and the term structure of welfare gain of early resolution of uncertainty (dotted line, right scale) at the one-year (top panel) and five (bottom panel) year horizon.