Risk Preferences and The Macro Announcement Premium

Hengjie Ai and Ravi Bansal *

August 5, 2016

The paper develops a theory for equity premium around macroeconomic announcements. Stock returns realized around pre-scheduled macroeconomic announcements, such as the employment report and the FOMC statements, account for 55% of the market equity premium during the 1961-2014 period, and virtually 100% of it during the later period of 1997-2014, where more announcement data are available. We provide a characterization theorem for the set of intertemporal preferences that generate a positive announcement premium. Our theory establishes that the announcement premium identifies a significant deviation from expected utility and constitutes an asset market based evidence for a large class of non-expected models that features aversion to "Knightian uncertainty", for example, Gilboa and Schmeidler [30]. We also present a dynamic model to account for the evolution of equity premium around macroeconomic announcements.

JEL Code: D81, G12

Key words: equity premium, announcement, Knightian uncertainty, robustness

*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. The authors would like to thank Anmol Bhandari, Jacob Sagi, Jan Werner, V.V. Chari, and seminar participants at the Duke Finance workshop, HKUST, NBER Neemrana, SAIF, SED 2016, the University of Minnesota Economic Theory workshop, the University of Texas at Austin, and the University of Toronto for their helpful comments on the paper.

1 Introduction

In this paper, we develop a theory of equity premium for pre-scheduled macroeconomic announcements. We demonstrate that the macro announcement premium provides an asset-market-based evidence that highlights the importance of incorporating non-expected utility analysis in quantitative macro and asset pricing models.

Macro announcements, such as the publication of the employment report and the FOMC statements, resolve uncertainty about the future course of economic growth, and modern financial markets allow asset prices to react to these announcements instantaneously, especially with high-frequency trading. Empirically, a large fraction of the market equity premium in the United States is realized within a small number of trading days with significant macroeconomic announcements. During the period of 1961-2014, for example, the cumulative excess returns of the S&P 500 index on the thirty days per year with significant macroeconomic news announcements averaged 3.36%, which accounts for 55% of the total annual equity premium during this period (6.19%). The market return realized on announcement days constitutes 100% of the equity premium during the later period of 1997-2014, which more announcements are available.

To understand the above features of financial markets, we develop a theoretical model that allows uncertainty to resolve before the realizations of macroeconomic shocks and characterize the set of intermporal preferences for the representative consumer under which an announcement premium arises.

Our main result is that resolutions of uncertainty are associated with realizations of equity premium if and only if the investor's intertemporal preference can be represented by a certainty equivalence functional that increases with respect to second-order stochastic dominance, a property we define as generalized risk sensitivity. The above theorem has two immediate implications. First, the intertemporal preference has an expected utility representation if and only if the announcement premium is zero for all assets. Second, announcement premia can only be compensation for generalized risk sensitivity and cannot be compensation for the risk aversion of the Von Neumann–Morgenstern utility function.

Therefore, the macro announcement premium provides an asset-market-based evidence that identifies a key aspect of investors' preferences not captured by expected utility analysis: generalized risk sensitivity. Many non-expected utility models in the literature satisfy this property. For example, we show that the uncertainty aversion axiom of Gilboa and Schmeidler [30] provides a sufficient condition for generalized risk sensitivity. Therefore, the large magnitude of the announcement premium in the data can be interpreted as a strong empirical evidence for a broad class of non-expected utility models.

From an asset pricing perspective, the stochastic discount factor under non-expected utility generally has two components: the intertemporal marginal rate of substitution that appears in standard expected utility models and an additional term that can be interpreted as the density of a probability distortion. We demonstrate that the probability distortion is a valid stochastic discount factor for announcement returns. In addition, under differentiability conditions, generalized risk sensitivity is equivalent to the probability distortion being pessimistic, that is, it assigns higher weights to states with low continuation utility and lower weights to states with high continuation utility. Our results imply that the empirical evidence on the announcement premium is informative about the relative importance of the two components of the stochastic discount factor and therefore provides a strong discipline for asset pricing models.

We present a continuous-time model with learning to quantitatively account for the evolution of the equity premium before, at, and after macroeconomic announcements. In our model, investors update their beliefs about 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 two results in this environment. First, as in Breeden [12], 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, macro announcements result in non-trivial reductions of uncertainty, and are associated with realizations of a substantial amount of equity premium in an infinitesimally small window of time.

Second, the equity premium typically increases before macro announcements, peaks at the announcements, and drops sharply afterwards. The reduction in uncertainty right after announcements implies a simultaneous decline in the 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.

Related literature Our paper builds on the literature that studies decision making under non-expected utility. We adopt the general representation of dynamic preferences of Strzalecki [65]. Our framework includes most of the non-expected utility models in the literature as special cases. We show that examples of dynamic preferences that satisfy generalized risk sensitivity include the maxmin expected utility of Gilboa and Schmeidler [30], the dynamic version of which is studied by Chen and Epstein [16] and Epstein and Schneider [25]; the recursive preference of Kreps and Porteus [48] and Epstein and Zin [27]; the robust control preference of Hansen and Sargent [36, 37] and the related multiplier preference of Strzalecki [64]; the variational ambiguity-averse preference of Maccheroni, Marinacci, and Rustichini [54, 55]; the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji [46, 47]; and the disappointment aversion preference of Gul [31]. We also discuss the relationship between our notion of generalized risk sensitivity and related decision theoretic concepts studied in the above papers, for example, uncertainty aversion and preference for early resolution of uncertainty.

A vast literature applies the above non-expected utility models to the study of asset prices and the equity premium. We refer the readers to Epstein and Schneider [26] for a review of asset pricing studies with the maxmin expected utility model, Ju and Miao [40] for an application of the smooth ambiguity-averse preference, Hansen and Sargent [34] for the robust control preference, Routledge and Zin [62] for an asset pricing model with disappointment aversion, and Bansal and Yaron [9] and Bansal [6] for the long-run risk models that build on recursive preferences. The nonlinearity of the certainty equivalence functionals in the above models typically gives rise to an additional equity premium. However, the existing literature has not yet identified an asset-market-based evidence for the nonlinearity of the certainty equivalence functionals in the above models. Our results imply that the macro announcement premium provides such an evidence.

Our findings are consistent with the literature that identifies large variations in marginal utilities from the asset market data, for example, Hansen and Jagannathan [33], Bansal and Lehmann [7], and Alvarez and Jermann [3, 4]. Our theory implies that quantitatively, most of the variations in marginal utility must come from generalized risk sensitivity and not from risk aversion in expected utility models.

The above observation is likely to have sharp implications on the research on macroeconomic policies. Several recent papers study optimal policy design problems in non-expected utility models. For example, Farhi and Werning [29] and Karantounias [43] analyze optimal fiscal policies with recursive preferences, and Woodford [68], Karantounias [42], Hansen and Sargent [38], and Kwon and Miao [51, 50] focus on preferences that fear model uncertainty. In the above studies, the nonlinearity in agents' certainty equivalence functionals implies a forward-looking component of variations in their marginal utilities that affects policy makers' objectives. Our results imply that the empirical evidence of the announcement premium can be used to gauge the magnitude of this deviation from expected utility, and to quantify the importance of robustness in the design of macroeconomic policies.

Our empirical results are related to the previous research on stock market returns on macroeconomic announcement days. The previous literature documents that stock market returns and Sharpe ratios are significantly higher on days with macroeconomic news releases in the United States (Savor and Wilson [63]) and internationally (Brusa, Savor, and Wilson

[13]). Lucca and Moench [52] find similar patterns for equity returns on pre-scheduled FOMC announcement days. Cieslak, Morse, and Vissing-Jorgensen [18] provide an extensive study of the cyclical pattern of stock returns around FOMC announcements. Rosoiu [60] finds similar patterns in high-frequency data.

The rest of the paper is organized as follows. We document some stylized facts for the equity premium for macroeconomic announcements in Section 2. In Section 3, we present two simple examples to illustrate how the announcement premium can arise in models that deviate from expected utility. We present our theoretical results and discuss the notion of generalized risk sensitivity in Section 4. We present a continuous-time model in Section 5 to quantitatively account for the evolution of the equity premium around macroeconomic news announcement days. Section 6 concludes.

2 Stylized facts

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

1. A large fraction of the market equity premium are realized on a relatively small number of trading days with pre-scheduled macroeconomic news announcements (See also Savor and Wilson [63] and Lucca and Moench [52]).

   As shown in Table 1, during the 1961-2014 period, on average, thirty trading days per year have significant macroeconomic news announcements. The cumulative stock market excess return on the thirty news announcement days averages 3.36% per year, accounting for about 55% of the annual equity premium (6.19%) during this period. This pattern is even more pronounced if we focus on the later period of 1997-2014, where data on on all five announcements are available. In this period, the market equity premium is 7.44% per year, and the cumulative excess return of the S&P500 index on the fifty announcement days averages 8.24% per year. The equity premium on the rest of the trading days is not statistically different from zero.

Figure 1: **thirty-minute returns around announcements**

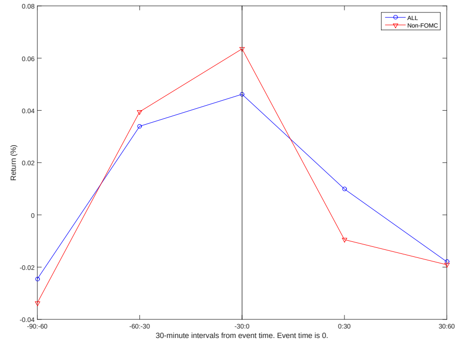


Figure 1 plots the average returns over 30 minute intervals around macro announcements, where time 0 is the announcement time. The solid line is the average return for all five announcements and the dashed line plots the average return for all announcements except the FOMC announcement. The sample period is 1997-2013.

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

   In Table 2, we document the average daily excess stock market return on announcement days (t), that on the days right before announcements ($t-1$) and that on the day right after announcements ($t+1$). The average announcement day return is 11.21 basis points during the entire sample period (top panel) and 16.48 basis points during the later period of 1997-2014, when more announcement data are available. The daily equity premium before and after announcement days is not statistically different from zero. In Figure 1, we illustrate the pattern of the evolution of market excess return in thirty-minute intervals around macro announcements by using high-frequency data for S&P 500 futures contracts during the period of 1997-2014. This evidence highlights that most of the announcement premium is realized during a short time interval around announcements.

3. The significance of the macro announcement premium is robust both intraday and overnight.

   Some announcements are pre-scheduled during financial market trading hours (e.g., FOMC announcements) and others are pre-scheduled prior to the opening of financial markets (e.g., non-farm payrolls). We define intraday return (or open-to-close return) as the stock market return from the open to the close of a trading day and overnight return (or close-to-open return) as the return from the close of a trading day to the open of the next trading day. We compute intraday and overnight returns for periods with and without prescheduled announcements and report our findings in Table 3. The average overnight return during the 1997-2014 period averages about 3.52 basis points per day, and the average intraday return is close to zero[1]. Remarkably, both intraday and overnight return are significantly larger on pre-scheduled announcements days than on non-announcements days. The average intraday return with announcement is 17.0 basis points, and the average overnight return with announcement is 9.32 basis points, while the average intraday and overnight return on non-announcement are not statistically different from zero. This new evidence reinforces the view that most of the equity premium realizes during periods of macroeconomic announcements.

3 Two illustrative examples

In this section, we set up a two-period model and discuss two simple examples to illustrate conditions under which resolutions of uncertainty are associated with realizations of the equity premium.

3.1 A two-period model

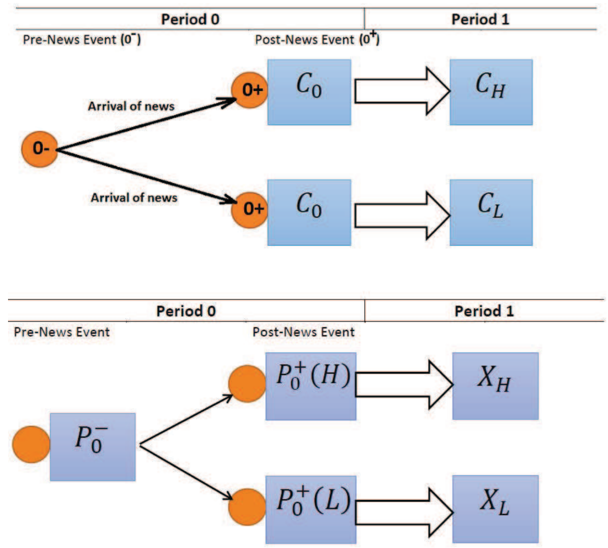
We consider a representative-agent economy with two periods (or dates), 0 and 1. There is no uncertainty in period 0, and the period-0 aggregate endowment is a constant, C_0. Aggregate endowment in period 1, denoted C_1, is a random variable. We assume a finite number of states: $n = 1, 2, \cdots N$, and denote the possible realizations of C_1 as $\{C_1(n)\}_{n=1,2,\cdots N}$. The probability of each state is $\pi(n) > 0$ for $n = 1, 2, \cdots, N$.

Period 0 is further divided into two subperiods. In period 0^-, before any information about about C_1 is revealed, the pre-announcement asset market opens, and a full set of Arrow-Debreu security is traded. The asset prices at this point are called pre-announcement prices and are denoted as P^-. Note that P^- cannot depend on the realization of C_1, which is not known at this point.

In period 0^+, the agent receives a news announcement s that carries information about C_1. Immediately after the arrival of s, the post-announcement asset market opens. Asset prices at this point, which are called post-announcement prices, depend on s and are denoted as $P^+(s)$. In period 1, the payoff of the Arrow-Debreu securities are realized and C_1 is consumed. In Figure 2, we illustrate the timing of information and consumption (top panel) and that of the asset markets (panel) assuming that $N = 2$ and that the news announcement s fully reveals C_1.

[1]The previous literature (for example, Kelly and Clark [45] and Polk, Lou, and Skouras [58]) documents that the overnight market return is on average higher than the intraday return in the United States.

Figure 2: **consumption and asset prices in the two-period model**



The announcement return of an asset, which we denote as $R_A(s)$, is defined as the return of a strategy that buy the asset before the pre-scheduled announcement and sell right afterwards:

$$R_A(s) = \frac{P^+(s)}{P^-}. \tag{1}$$

We say that the asset requires a positive announcement premium if $E[R_A(s)] > 1$. We also define the post-announcement return conditioning on announcement s as: $R_P(X|\, s) = \frac{X}{P^+(s)}$. Clearly, the total return of asset from period 0^- to period 1 is $R(X) = R_A(s)\, R_P(X|\, s)$.

We note two important properties of the news announcement in our model. First, it affects the conditional distribution of future consumption, but rational expectations imply that surprises in news must average to zero by the law of iterated expectation. Second, news announcements do not impact the current-period consumption. Empirically, as we show in Section 2, the stock market returns realized within the 30-minute intervals of announcements account for almost all of the equity premium at the annual level. The instantaneous response of consumption to news at this high frequency can hardly have any quantitative effect on the total consumption of the year.[2] Furthermore, our analysis remains valid in the continuous-time model presented in Section 5, where the length of a period is infinitesimal.

[2]Our assumption is broadly consistent with the empirical evidence that large movements in the stock market is typically not associated with significant immediate adjustment in aggregate consumption, for example, Bansal and Shaliastovich [8].

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 strictly increasing and continuously differentiable. For simplicity, we assume that s fully reveals C_1 in this example and the one in the following section, although our general result in Section 3.3 does not depend on this assumption. The pre-announcement price of an asset with payoff X is given by:

$$P^- = E\left[\frac{\beta u'(C_1)}{u'(C_0)} X\right]. \tag{2}$$

In period 0^+, because s full reveals the true state, the agent's preference is represented by

$$u(C_0) + \beta u(C_1(s)). \tag{3}$$

Therefore, for any s, the post-announcement price of the asset is

$$P^+(s) = \frac{\beta u'(C_1(s))}{u'(C_0)} X(s). \tag{4}$$

Clearly, the expected announcement return is $E[R_A(s)] = \frac{E[P^+(s)]}{P^-} = 1$. 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 preferences of Hansen and Sargent [35]:

$$\begin{aligned} & u(C_0) + \beta \min_m E[mu(C_1)] \\ & \textit{subject to}: \ E[m \ln m] \le \eta \\ & E[m] = 1. \end{aligned} \tag{5}$$

The above expression also can be interpreted as the maxmin expected utility of Gilboa and Schmeidler [30]. The agent treats the reference probability measure, under which equity premium is evaluated (by econometricians), as an approximation. As a result, the agent takes into account a class of alternative probability measures, represented by the density m close to the reference probability measure. The inequality $E[m \ln m] \le \eta$ requires that the relative entropy of the alternative probability models to be less than η.

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

$$P^{-}=E\left[m^{*}\frac{\beta u'\left(C_{1}\right)}{u'\left(C_{0}\right)}X\right], \tag{6}$$

where m^* is the density of the minimizing probability for (5) and can be expressed as a function of s:

$$m^{*}\left(s\right)=\frac{e^{-\frac{u\left(C_{1}(s)\right)}{\theta}}}{E\left[e^{-\frac{u\left(C_{1}\right)}{\theta}}\right]}. \tag{7}$$

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

In period 0^+, after the resolution of uncertainty, the agent's utility reduces to (3). As a result, the post-announcement price of the asset is the same as that in (4). Therefore, we can write the pre-announcement price as:

$$P^{-}=E\left[m^{*}\left(s\right)P^{+}\left(s\right)\right]. \tag{8}$$

Because m^* is a decreasing function of date-1 utility $u\left(C_1\right)$, it is straightforward to prove the following claim.

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

The intuition of the above result is clear. Because uncertainty is resolved after the announcement, asset prices are discounted using marginal utilities. Under expected utility, the pre-announcement price is computed using probability-weighted marginal utilities, and therefore the pre-announcement price must equal the expected post-announcement prices and there can be no announcement premium under rational expectation. Under the robust control preference, the pre-announcement price is not computed by using the reference probability, but rather by using the pessimistic probability that overweighs low-utility states and underweighs high-utility states as shown in equation (7). 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.

Because the probability distortion m^* discounts announcement returns, we will call it the announcement stochastic discount factor (SDF), or A-SDF to distinguish it from the standard SDF in intertemporal asset pricing models which are derived from agents' marginal rate of intertemporal substitution of consumption. In our model, there is no intertemporal consumption decision before and after the announcement. The term m^* reflects investors' uncertainty aversion and identifies the probability distortion relative to rational expectation.

4 Risk preferences and the announcement premium

The asset pricing equation (8) holds under much more general conditions. In this section, we consider a very general class of intertemporal preferences that has the recursive representation

$$V_t = u(C_t) + \beta \mathcal{I}[V_{t+1}], \tag{9}$$

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). We assume that $\mathcal{I}$ is normalized and weakly increasing in first-order stochastic dominance.[3] As shown by Strzalecki [65], representation (9) includes most of the dynamic preferences under uncertainty proposed in the literature.[4] Our main focus is to characterize the set of preferences under which the announcement premium is non-negative for payoffs that are increasing functions of continuation utility.

4.1 The announcement SDF

We continue to focus on the two-period model assuming i) a finite state space, ii) fully revealing announcements, and iii) equal probability of each state, that is, $\pi(s) = \frac{1}{N}$ for $s = 1, 2, \cdots, N$. Although none of the above assumptions are neccessary for our main theorem, they allow us to avoid cumbersome notations and illustrate the basic intuition for our results. We present and prove our general theorems without the above assumptions in a fully dynamic setup in Appendix C.

In the two-period model, because the signal s fully reveals the state of the world, $C_{1,s}$ is known at announcement. Because $\mathcal{I}$ is normalized and V_s is known after receiving announcement s, the agent's post-announcement utility can be written as: $V_s = u(C_0) + \beta u(C_{1,s})$. With a finite number of states, the certainty equivalence functional $\mathcal{I}$ can be viewed

[3] A certainty equivalence functional $\mathcal{I}$ is normalized if $\mathcal{I}[k] = k$ whenever k is a constant. It is weakly increasing in first-order stochastic dominance if $\mathcal{I}[X_1] \geq \mathcal{I}[X_2]$ whenever X_1 first-order stochastic dominates X_2. See Appendix C for details.

[4] As Strzalecki [65] shows, this representation includes the maxmin expected utility of Gilboa and Schmeidler [30], the second-order expected utility of Ergin and Gul [28], the smooth ambiguity preferences of Klibanoff, Marinacci, and Mukerji [46], the variational preferences of Maccheroni, Marinacci, and Rustichini [54], the multiplier preferences of Hansen and Sargent [34] and Strzalecki [64], and the confidence preferences of Chateauneuf and Faro [15]. In addition, our setup is more general than that of Strzalecki [65], because we do not require the function $u(\cdot)$ to be affine. In Appendix B, we provide expressions for the A-SDF for the above-mentioned decision models.

as a function from R^N to R. We use the vector notation and denote the agent's utility at time 0^- as $\mathcal{I}[V]$, where $V = [V_1, V_2, \cdots, V_N]$. Assuming $\mathcal{I}$ is differentiable, from date 0^- perspective, the agent's marginal utility with respect to C_0 is

$$\frac{\partial}{\partial C_0}\mathcal{I}[V] = \sum\nolimits_{s=1}^{N} \frac{\partial}{\partial V_s}\mathcal{I}[V] \cdot u'(C_0),$$

and the marginal utility with respect to $C_{1,s}$ is

$$\frac{\partial}{\partial C_{1,s}}\mathcal{I}[V] = \frac{\partial}{\partial V_s}\mathcal{I}[V] \cdot \beta u'(C_{1,s}).$$

Because the pre-announcement price of an asset can be computed as the marginal utility weighted payoffs, we can write

$$P^- = \sum_{s=1}^{N} \frac{\frac{\partial}{\partial C_{1,s}}\mathcal{I}[V]}{\frac{\partial}{\partial C_0}\mathcal{I}[V]} X_s = E\left[m^*(s)\,\beta\frac{u'(C_1(s))}{u'(C_0)}X(s)\right], \tag{10}$$

where

$$m^*(s) = \frac{1}{\pi(s)}\frac{\frac{\partial}{\partial V_s}\mathcal{I}[V]}{\sum_{s=1}^{N}\frac{\partial}{\partial V_s}\mathcal{I}[V]}. \tag{11}$$

At time 0^+, s fully reveals the state, and $P^+(s)$ is, again, given by equation (4). Clearly, the asset pricing equation (8) holds with the A-SDF m^* defined by (11). We summarize our results for the existence of A-SDF as follows.

Theorem 1. *(Existence of A-SDF)*

Assume that both u and $\mathcal{I}$ are continuously differentiable with strictly positive (partial) derivatives. Assume also that $\mathcal{I}$ is normalized.[5] *Then in any competitive equilibrium, there exists a strictly positive $m^* = \{m^*(s)\}_{s=1,2,\cdots,N}$ such that*

1. *m^* is a density, that is, $E[m^*] = 1$, and*
2. *for all announcement returns $R_A(s)$,*

$$E[m^*(s)\,R_A(s)] = 1. \tag{12}$$

To provide a precise statement about the sign of the announcement premium, we focus our attention on payoffs that are co-monotone with continuation utility. We define an announcement-contingent payoff f to be *co-monotone with continuation utility* if $[f(s) - f(s')][V(s) - V(s')] \geq 0$ for all s, s'. Intuitively, co-monotonicity captures the idea

[5] See Definition 5 in Appendix C.

that the payoff f is an increasing function of the continuation utility V. We are interested in identifying properties of the certainty equivalence functional $\mathcal{I}$ such that the following condition is true.

Condition 1. *The announcement premium is non-negative for all payoffs that are co-monotone with continuation utility V.*

In the case of expected utility, $\mathcal{I}[V] = \sum_{s=1}^{N} \pi(s) V_s$ is linear, and $m^*(s) = 1$ for all s by equation (11). As a result, equation (12) reduces to $E[R_A(s)] = 1$, and there can be no premium for any announcement return. In general, assuming $m^*(s)$ is a decreasing function of aggregate consumption, $C_{1,s}$, or equivalently, a decreasing functions of V_s, then by equation (12), the announcement premium $E[R_A(s)] - 1 = -cov[m^*(s), R_A(s)]$ must be positive for all returns $R_A(s)$ that are co-monotone with continuation utility V_s. The converse of the above statement is also true. That is, suppose $m^*(s)$ is a not a decreasing functions of V_s, then we can always find an return that is co-monotone with V_s, but requires a negative announcement premium.

As a result of the above observation, Condition 1 is equivalent to $m^*(s)$ being a decreasing function of V_s, which is equivalent to $\frac{\partial}{\partial V_s}\mathcal{I}[V]$ being a decreasing function of V_s by Equation (11). Under the assumption of equal probability for all states, the latter property is known to be equivalent to $\mathcal{I}$ being increasing in second-order stochastic dominance.[6] We can summarize our main results as follows.

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

1. *the announcement premium is zero for all assets if and only $\mathcal{I}$ is the expectation operator.*

2. *Condition 1 is equivalent to the certainty equivalence functional $\mathcal{I}$ being non-decreasing with respect to second-order stochastic dominance.*

The above theorem holds under much more general conditions. In Appendix D, we show that the conclusion of the above theorem remains true in a fully dynamic model without the assumption of fully revealing signals or a finite number of states with equal probability.[7]

[6] Under our assumptions, this property of $\mathcal{I}$ is also known as Schur concavity. See also Muller and Stoyan [57] and Chew and Mao [17].

[7] A stronger version of the above theorem is also true. That is, the announcement premium is positive for all payoffs that are strictly co-monotone with continuation utility V_s if and only if the certainty equivalence functional $\mathcal{I}$ is strictly increasing with respect to second-order stochastic dominance. A proof for this result is available upon request.

4.2 Generalized risk-sensitive preferences

Theorem 2 motivates the following definition of generalized risk sensitivity.

Definition 1. *(Generalized Risk Sensitivity)*

An intertemporal preference of the form (18) is said to satisfy (strictly) generalized risk sensitivity, if the certainty equivalence functional $\mathcal{I}$ is (strictly) monotone with respect to second-order stochastic dominance.

Under our definition, generalized risk-sensitive preference is precisely the class of preferences that requires a non-negative announcement premium for all assets with payoff co-monotone with investors' continuation utility.

The SDF for many non-expected utility models takes the form $m^*(s)\,\beta\frac{u'(C_1(s))}{u'(C_0)}$, as in Equation (10).[8] Theorem 2 has two important implications. First, generalized risk sensitivity is precisely the class of preferences under which m^* is a decreasing function of continuation utility and therefore enhances risk compensation. Second, generalized risk sensitivty can be identified from empirical evidence on announcement returns. In fact, as we show in Section 4.3, the empirical evidence on the announcement premium can be used to gauge the quantitative importance of the term m^*.

To clarify the notion of generalized risk sensitivity, we next discuss its relationship with two other related properties of choice behavior under uncertainty that are known to be associated with higher risk compensations, uncertainty aversion (Gilboa and Schmeidler [30]), and the preference for early resolution of uncertainty (Kreps and Porteus [48]).

Generalized risk sensitivity and uncertainty aversion Most of the dynamic uncertainty-averse preferences studied in the literature can be viewed as a special case of the general representation (9). In the special case in which $u(\cdot)$ is affine as in Strzalecki [65], quasiconcavity is equivalent to the uncertainty aversion axiom of Gilboa and Schmeidler [30]. We make the following observations about the relationship between uncertainty aversion and generalized risk sensitivity.

1. *Quasiconcavity of $\mathcal{I}$ is sufficient, but not neccessary, for generalized risk sensitivity.*

   A direct implication of the above result is that all uncertainty-averse preferences can be viewed as different ways to formalize generalized risk sensitivity, and they all require

[8] For example, Hansen and Sargent [34] use a risk-sensitive operator to motivate the term m^* as a decreasing function of continuation utility. In this sense, our notion of generalizes the risk-sensitive operator of Hansen and Sargent [34].

a non-negative announcement premium (for all assets with payoffs co-monotone with continuation utility). These preferences include the maxmin expected utility of Gilboa and Schmeidler [30], the second-order expected utility of Ergin and Gul [28], the smooth ambiguity preferences of Klibanoff, Marinacci, and Mukerji [46], the variational preferences of Maccheroni, Marinacci, and Rustichini [54], the multiplier preferences of Hansen and Sargent [34] and Strzalecki [64], and the confidence preferences of Chateauneuf and Faro [15].

In Appendix D, we provide a proof for the sufficiency of quasiconcavity for generalized risk sensitivity. To illustrate that quasiconcavity is not neccessary, in the same appendix, we also provide an example that satisfies generalized risk sensitivity, but not quasiconcavity.

2. *If $\mathcal{I}$ is of the form $\mathcal{I}[V] = \phi^{-1}(E[\phi(V)])$, where ϕ is a strictly increasing function, then generalized risk sensitivity is equivalent to quasiconcavity, which is also equivalent to the concavity of ϕ.*

   The certainty equivalence function of many intertemporal preferences takes the above form, for example the the second-order expected utility of Ergin and Gul [28] and the recursive preferences of Kreps and Porteus [48] and Epstein and Zin [27]. For these preferences, g-risk sensitivity is equivalent to the concavity of ϕ.

3. *Within the class of smooth ambiguity-averse preferences, uncertainty aversion is equivalent to generalized risk sensitivity.*

   The smooth ambiguity-averse preference of Klibanoff, Marinacci, and Mukerji [46, 47] can be represented in the form of (9) with the following choice of the certainty equivalence functional:

$$\mathcal{I}[V] = \phi^{-1}\left\{\int_{\Delta} \phi\left(E^{x}[V]\right) d\mu(x)\right\}. \tag{13}$$

   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. In Appendix D, we show that generalized risk sensitivity is equivalent to the concavity of ϕ, which is also equivalent to uncertainty aversion.

Generalized risk sensitivity and preference for early resolution of uncertainty A well-known class of model that gives rise to an additional term m^* in the SDF in (10) is the recursive preference with constant relative risk aversion and constant intertemporal elasticity of substitution (IES). It is also well-known that the resultant m^* is decreasing in continuation utility under preference for early resolution of uncertainty, that is, if the relative risk aversion γ is greater than the reciprocal of IES, $\frac{1}{\psi}$. However, for the general representation (9), generalized risk sensitivity is neither necessary nor sufficient for preference for early resolution of uncertainty. A simple example of a preference with indifference toward the timing of resolution of uncertainty, but requires a strictly positive announcement premium, is the constraint robust control preference in Section 3.3.[9] We summarize our main results below.

1. *Concavity of the certainty equivalence functional $\mathcal{I}$ is sufficient for both generalized risk sensitivity and preference for early resolution of uncertainty.*

   Note that concavity implies quasiconcavity and therefore generalized risk sensitivity. Theorem 2 of Strzalecki [65] also implies that these preferences satisfy preference for early resolution of uncertainty. As a result, Theorems 2 and 3 of Strzalecki [65] imply that the variational preference of Maccheroni, Marinacci, and Rustichini [54] is both risk sensitive and prefers early resolution of uncertainty.

2. *If $\mathcal{I}$ of the form $\mathcal{I}[V] = \phi^{-1}(E[\phi(V)])$ or is the smooth ambiguity preference, $\mathcal{I}[V] = \int_{\Delta} \phi(E^x[V])\,d\mu(x)$, where ϕ is strictly increasing and twice continuously differentiable, then generalized risk sensitivity implies preference for early resolution of uncertainty if either of the following two conditions hold.*

   (a) There exists $A \geq 0$ such that $-\frac{\phi''(a)}{\phi'(a)} \in [\beta A, A]$ for all $a \in R$.

   (b) $u(C) \geq 0$ for all C, and $\beta\left[-\frac{\phi''(k+\beta a)}{\phi'(k+\beta a)}\right] \leq -\frac{\phi''(a)}{\phi'(a)}$ for all a, $k \geq 0$.

   The above two conditions are the same as Conditions 1 and 2 in Strzalecki [65]. Intuitively, they require that the Arrow-Pratt coefficient of the function ϕ does not vary too much. In both cases, generalized risk sensitivity implies the concavity of ϕ. By Theorem 4 of Strzalecki [65], either of the above conditions implies preference for early resolution of uncertainty.

   Because the recursive utility with constant relative risk aversion and constant IES can be represented in the form of (9) with $u(C) = \frac{1}{1-\frac{1}{\psi}}C^{1-\frac{1}{\psi}}$, and $\mathcal{I}[V] = \phi^{-1}(E[\phi(V)])$,

[9] To verify that this preference is indifferent toward the timing of resolution of uncertainty, note that the time 0^- utility of the agent is

$$V^E = \min E[m\{u(C_0) + \beta u(C_1)\}],$$

if uncertainty is resolved at 0^+, and the utility at time 0^- is

$$V^L = u(C_0) + \beta \min E[m\{u(C_1)\}],$$

if uncertainty is resolved at time 1. Clearly, $V^E = V^L$.

where $\phi(x) = \left[\left(1 - \frac{1}{\psi}\right) x\right]^{\frac{1-\gamma}{1-1/\psi}}$. It follows from Condition b that $\mathcal{I}$ is quasi-concave and therefore requires a positive announcement premium if and only if $\gamma \geq \frac{1}{\psi}$. That is, for this class of preferences, preference for early resolution of uncertainty and generalized risk sensitivity are equivalent. Therefore, assuming a recursive utility with constant relative risk aversion and constant IES, the empirical evidence for announcement premium can be also interpreted as evidence for preference for early resolution of uncertainty.

3. *In general, preference for early resolution of uncertainty is neither sufficient nor necessary for a positive announcement premium.*

   In Appendix D, we provide an example of a generalized risk-sensitive preference that violates preference for early resolution of uncertainty, as well as an example of an utility function that prefers early resolution of uncertainty, but does not satisfy generalized risk sensitivity.

4. *The only class of preferences that requires a positive announcement premium and is indifferent toward the timing of resolution of uncertainty is the maxmin expected utility of Gilboa and Schmeidler [30].*

   The maxmin expected utility of Gilboa and Schmeidler [30] is the only class of preference of the form (9) that is indifferent toward the timing of resolution of uncertainty (Strzalecki [65]). As the example in Section 2.3 shows, this class of preference exhibits generalized risk sensitivity.

4.3 Asset pricing implications

Decomposition of returns by the timing of its realizations In general, equity returns can be decomposed into an announcement return and a post-announcement return. Using the notations we setup in Section 3 of the paper, the return of an asset can be computed as:

$$R(X) = \frac{X}{P^-} = R_P(X|s) R_A(s),$$

where $R_A(s)$ is the announcement return defined in (1), and $R_P(X|s) = \frac{X}{P^+(s)}$ is the post-announcement return (conditioning on s). The optimal portfolio choice problem on the post-announcement asset market implies that for each s, there exists $y^*(C_1|s)$, which is a function of C_1, such that

$$E\left[y^*(C_1|s) R_P(X|s)|s\right] = 1, \tag{14}$$

for all post-announcement returns. In our simple model in which s fully reveals C_1, $y^* = \frac{\beta u'(C_1)}{u'(C_0)}$. In general, y^* depends on agents' intertemporal marginal rate of substitution on the post-announcement asset market. Combining Equations (12) and (14), and applying the law of iterated expectation, we have

$$E\left[m^* y^* \cdot R(X)\right] = 1.^{10} \tag{15}$$

Equation (15) appears in many intertemporal asset pricing models. Equations (12) and (14) provide a decomposition of intertemporal returns into an announcement return and a post-announcement return and a decomposition of the SDF. We make the following comments.

1. Theorem 2 implies that the announcement premium must be compensation for generalized risk sensitivity and cannot be compensation for risk aversion associated with the Von Neumann–Morgenstern utility function u, because the A-SDF, m^*, depends only on the curvature of the certainty equivalence functional $\mathcal{I}[\cdot]$, and not on u. The announcement premium is determined by the properties of $\mathcal{I}[\cdot]$, whereas the post-announcement premium, which is not realized until the action of consumption is completed, refects the curvature of u.

2. The large magnitude of announcement return in the data implies that the probability distortion in m^* must be significant. Just as 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, we can bound the entropy of the A-SDF as in Bansal and Lehmann [7] and Backus, Chernov, and Zin [5]: $-E[\ln(m^* y^*)] = -E[\ln m^*] - E[\ln y^*] \geq E[\ln R]$. Using the average market return during the 1997-2014 period, we have $-E[\ln(m^* y^*)] \geq E[\ln R] = 7.44\%$. The entropy bound on the probability distortion component provides a tighter restriction: $-E[\ln m^*] \geq E[\ln R_A] = 8.24\%$ using the average announcement return during the same period. These bounds imply that the volatility of the SDF y^* is likely to be small.

3. The Hansen-Jaganathan bound (Hansen and Jagannathan [33]) for the SDF's leads to a similar conclusion. Equation (12) implies that for any announcement return, R_A, $\sigma[m^*] \geq \frac{E[R_A - 1]}{\sigma[R_A]}$. The Sharpe ratio for announcement returns reported in Table 1 can be used to compute the Hansen-Jaganathan bound for A-SDF: $\sigma[m^*] \geq 55\%$ for the 1961-2014 period, and $\sigma[m^*] \geq 88\%$ if we focus on the later period of 1997-2014, where more announcement data are available.

[10] We show in Appendix C that this decomposition holds in the fully dynamic model. See equation (43).

In contrast, the moments of the total stock market return in the same period in fact implies a weaker bound on SDF that prices the overall return, m^*y^*. Using equation (15), we have

$$\sigma\left[m^*y^*\right] \geq \frac{E\left[R\left(X\right)-1\right]}{\sigma\left[R\left(X\right)\right]}. \tag{16}$$

Inequality (16) implies $\sigma\left[m^*y^*\right] \geq 40\%$ if we focus on the period of 1961-2014 in Table 1, and $\sigma\left[m^*y^*\right] \geq 37\%$ for the latter period of 1997-2014.

4. Assuming log-normality, equation (12) implies $\ln E\left[R_A\right] = -Cov\left[\ln R_A, \ln m^*\right]$, and equation (15) implies

$$\begin{aligned} \ln E\left[R\right] - \ln R_f &= -Cov\left[\ln R, \ln m^* + \ln y^*\right] \qquad (17) \\ &= -Cov\left[\ln R, \ln m^*\right] - Cov\left[\ln R, \ln y^*\right]. \end{aligned}$$

Because $\ln R = \ln R_A\left(s\right) + \ln R_P$ and $-Cov\left[\ln R_A, \ln m^*\right] = \ln E\left[R_A\right]$ by equation (12), we have:

$$\ln E\left[R\right] - \ln R_f = \ln E\left[R_A\right] - Cov\left[\ln m^*, \ln R_P\right] - Cov\left[\ln y^*, \ln R\right]. \tag{18}$$

As we explained earlier, most of the equity premium in the data is announcement premium. Therefore, the term $-Cov\left[\ln m^*, \ln R_P\right] - Cov\left[\ln y^*, \ln R\right]$ must be close to zero. Note that $m^*\left(s\right)$ is a function of the announcement, while $R_P\left(X|s\right)$ is the post-announcement return; therefore, the term $Cov\left[\ln m^*, \ln R_P\right]$ is likely to be close to zero. This implies that the compensation for the curvature of the Von Neumann–Morgenstern utility function u as captured by the term $-Cov\left[\ln y^*, \ln R\right]$ must be close to zero.

5. The external habit preference does not generate an announcement premium, and the internal habit model produces a negative announcement premium.

   The external habit preference (for example, Constantinides [20] and Campbell and Cochrane [14]) can be written as: $E\left[\sum_{t=0}^{T}\frac{1}{1-\gamma}\left(C_t - H_t\right)^{1-\gamma}\right]$, where $\{H_t\}_{t=0}^{\infty}$ is the habit process. Consider the date-0 market for announcements. The pre-announcement price of any payoff $\{X_t\}_{t=1}^{\infty}$ is

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

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

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

Clearly, like in the case of expected utility, $P_0^- = E\left[P_0^+(s)\right]$, and there can be no announcement premium. The external habit model is similar to expected utility because the habit process is exogenous, and agents do not take into account of the effect of current consumption choices on future habits when making consumption and investment decisions. We further show in Appendix D that the internal habit model (for example, Boldrin, Christiano, and Fisher [11]) generates a negative announcement premium.

5 A quantitative model of announcement premium

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 the 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 distinguish the announcement premium that is instantaneously realized upon news announcements and the risk premium that investors receive incrementally as shocks to consumption materialize slowly over time.

5.1 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 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 [9].

Timing of information and Bayesian learning The representative agent in the economy can use two sources of information to update 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 s_n as the signal observed at time nT 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 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 updates belief about x_t based on the observed consumption process according to Bayes' rule. We define $\hat{x}_t = E(x_t | C^t)$ as the posterior mean of x_t and time t, and define $q_t = E\left[(x_t - \hat{x}_t)^2 \middle| C^t\right]$ as the posterior variance of x_t.[11] 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, $q(t)$ satisfies the Riccati equation:

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

The posterior distribution is updated immediately following the announcement of signals

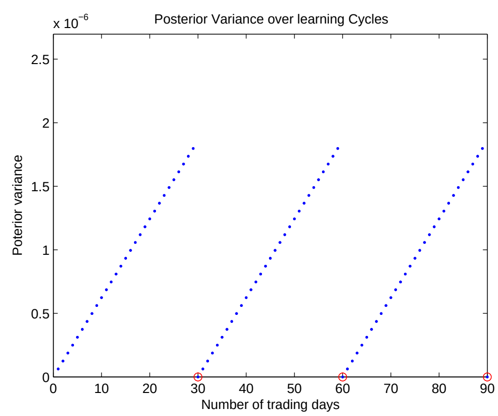
[11]Here we use the notation C^t to denote the history of consumption up to time t.

at time $T, 2T, \cdots$. At time T, for example, the agent updates beliefs using Bayes rule:

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

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 3 with the assumption that announcements are made every thirty days and they completely reveal the information about x_t, that is, $\sigma_S^2 = 0$. At announcements, the posterior variance drops immediately to zero, as indicated by the circles. After announcement, information slowly arrives and as a result, the posterior variance gradually increases over time before the next announcement.

Figure 3: **posterior variance of x_t**



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 [22] that can be interpreted as the limit of the recursive relationship (9) 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 choices 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 [34].

As is 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

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

Consistent with our earlier notation, we define the A-SDF $m^*_{t+\Delta} = \frac{e^{(1-\gamma)V_{t+\Delta}}}{E_t\left[e^{(1-\gamma)V_{t+\Delta}}\right]}$, and the intertemporal SDF $y^*_{t+\Delta} = e^{-\rho\Delta}\left(\frac{C_{t+\Delta}}{C_t}\right)^{-1}$. Clearly, $m^*_{t+\Delta}$ is a probability density and 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 $V_t = V\left(\hat{x}_t, t, C_t\right)$, and

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

where the function $h\left(t\right)$ is given in Appendix E 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 4. We now turn to the quantitative implications of the model on the announcement premium.

5.2 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) like in standard learning models (for example, Veronesi [66] 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:[12]

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

[12] In general, the equity premium depends on the state variable $\hat{x}$. 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 by the Markov chain approximation method.

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 3, 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 4, 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.

Figure 4: **Equity Premium on Non-Announcement Days**

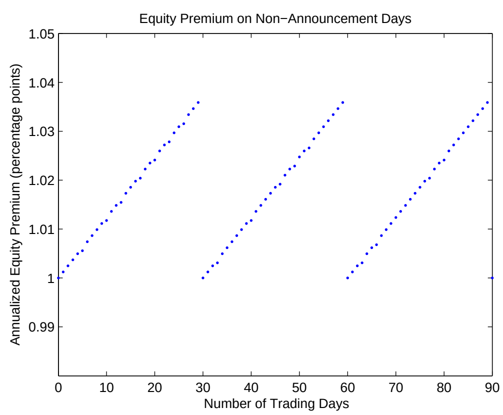


Figure 4 plots the annualized equity premium on non-announcement days in our model with recursive utility.

As in typical continuous-time models, the amount of equity premium is proportional to the 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 $y^*_{t+\Delta} = e^{-\rho\Delta}\left(\frac{C_{t+\Delta}}{C_t}\right)^{-\frac{1}{\psi}} \to 1$; however, the A-SDF $m^*_{t+\Delta}$ 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 announcements, 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 A-SDF for the announcement at time T.

We focus on asset prices and returns around announcements. We denote the time-T pre-announcement price-to-dividend ratio as p_T^- and the time-T post-announcement price-to-dividend ratio as p_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^- denotes the expectation taken with respect to the information at time T before announcements. Taking the limit as $\Delta \to 0$, and using the fact that $y_{t+\Delta}^* \to 1$, 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 (8). As we show in Appendix E, 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 5, 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 calculate 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, and 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 1. Consistent with our theoretical results in Section 2, the announcement premium for expected utility is zero. The premium on the announcement days for expected utility is 0.028 basis point, several orders of magnitude smaller than that for recursive utility.[13]

Figure 5: **announcement premium for recursive utility and expected utility**

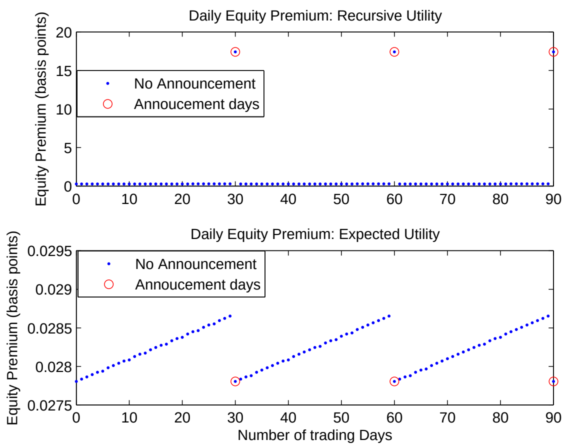


Figure 5 plots daily equity premium for recursive utility (top panel) and that for expected utility (bottom panel). Equity premium is measured in basis points. The dotted line is the equity premium on non-announcement days, and the circles indicate the equity premium on announcement days.

2. The announcement premium identifies the probability distortion component of the SDF. 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 A-SDF $\frac{e^{(1-\gamma)V_{t+\Delta}}}{E_t\left[e^{(1-\gamma)V_{t+\Delta}}\right]}$ that can be interpreted as probability distortion. As we take the limit $\Delta \to 0$, the intertemporal substitution of consumption term vanishes, and the A-SDF, (28), only depends 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 in the posterior belief of 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 becomes smaller, the half life of the impact of x_t on consumption increases, and so does the announcement premium.

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

Figure 6: **dynamics of price-to-dividend ratio**

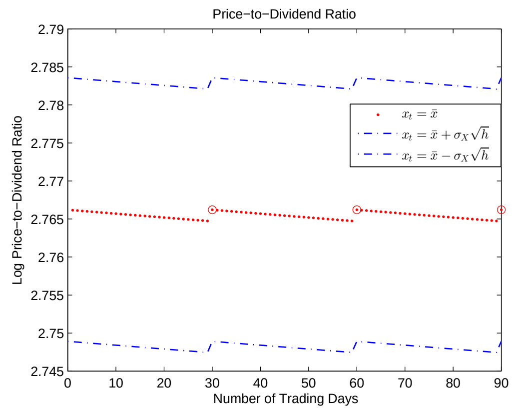


Figure 6 plots the evolution of log price-to-dividend ratio over the announcement cycles. The dotted line is the price-to-dividend ratio evaluated at the steady-state level of expected consumption growth rate, $\bar{x}$. The dash-dotted line is the price-to-dividend ratio for x_t one standard deviation above and below its steady-state level. The standard deviation is calculated at the monthly level, that is, $h = \frac{1}{12}$.

To better understand the nature of the announcement premium, in Figure 6, 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 the discount rate increases, the price-to-dividend ratio drops gradually until the next announcement. Because $\phi > 1$, the price-to-dividend ratio 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 5, 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 returns on announcement days averages 2.53% per year. In the model in which x_t is fully observable, announcements are not associated with any premium because their information content is fully anticipated.

As we have discussed before, the announcement premium is zero under expected utility, and therefore the magnitude of the equity premium on announcement days and that on non-announcement days is 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.

5.3 Preference for early resolution of uncertainty

As we have shown 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 discuss the relationship between the announcement premium and timing premium, defined as the welfare gain of early resolution of uncertainty, in the context of our continuous-time model.

We define timing premium as in Epstein, Farhi, and Strzalecki [24].[14] Consider a macroeconomic announcement at time 0^+ that resolves all uncertainty in the economy from time 0^+ to ∞. Let W^+ be the agent's utility at time 0^+ when all uncertainty is resolved, that is,

$$W_0^+(\tau) = \int_0^\infty e^{-\rho t}\rho \ln C_t dt. \tag{31}$$

Let W^- be the certainty equivalent of W^+:

$$W^- = \frac{1}{1-\gamma}\ln E^-\left[e^{(1-\gamma)W^+}\right]. \tag{32}$$

That is, W^- is the utility of the representative agent who does not know any information about the future, but anticipates that all uncertainty from time 0 to ∞ 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).

The timing premium λ is defined as the maximum fraction of life-time consumption that

[14] See also Ai [1] for a decomposition of the welfare of gain of early resolution of uncertainty in production economies.

the representative agent is willing to pay for the announcement:

$$1-\lambda = e^{V_0 - W_0^-(\tau)} = \frac{e^{V_0}}{\left\{E^-\left[e^{(1-\gamma)W^+}\right]\right\}^{\frac{1}{1-\gamma}}}. \tag{33}$$

Now consider an asset market for the announcement at time 0^+. Let p^+ as the price-to-dividend ratio of the equity claim at time 0^+ upon the announcement. The pre-announcement price-to-dividend ratio is given by

$$p^- = \frac{E_0^-\left[e^{(1-\gamma)W^+}p^+\right]}{E_0^-\left[e^{(1-\gamma)W^+}\right]}, \tag{34}$$

where W^+ is the agent's continuation utility after announcement as defined in equation (31). The expected return associated with the announcement can be calculated as

$$\frac{E_0^-\left[p_0^+\right]}{p^-} = \frac{E_0^-\left[p_0^+\right] E_0^-\left[e^{(1-\gamma)W^+}\right]}{E_0^-\left[e^{(1-\gamma)W^+}p^+\right]}.$$

It is not hard to show that $\lambda = \ln \frac{E_0^-\left[p_0^+\right]}{p^-} = 0$ for $\gamma = 1$. That is, the timing premium and announcement premium are both zero for expected utility. In addition, as long as p^+ is positively correlated with $e^{(1-\gamma)W^+}$, both the timing premium and the announcement premium increase with the measure of preference for early resolution of uncertainty, $\gamma - 1$.

As shown earlier in the paper, generalized risk sensitivity is necessary to account for the large observed announcement premium. Within the class of Kreps-Porteus preferences, the empirical evidence on the large announcement premium *ipso facto* implies a significant preference for early resolution of uncertainty. As there is no direct measurement on what investors are willing to pay for early resolution, it is hard to evaluate the model along this dimension. In addition to the announcement premium, previous literature has demonstrated that recursive preferences with a strong generalized risk sensitivity is important to account for risk premia across different assets classes (see Bansal and Yaron [9], Hansen, Heaton, and Li [32], Colacito and Croce [19], and Barro [10]) and other macro-economic facts (for example, Kaltenbrunner and Lochstoer [41] and Croce [21]). The announcement premium evidence along with this earlier papers underscore the importance of generalized risk sensitivity within this class of recursive preferences.

6 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 show that a positive announcement premium is equivalent to generalized risk sensitivity, that is, investors' certainty equivalence functional increases with respect to second order stochastic dominance. We demonstrate that generalized risk sensitivity is precisely the class of preferences in which deviations from expected utility enhances the volatility of the stochastic discount factor. As a result, our theoretical framework implies that the announcement premium can be interpreted as an asset-market-based evidence for a broad class of non-expected utility models that feature aversion to "Knightian uncertainty". We also present a dynamic model to quantitatively account for the pattern of equity premium around news announcement days.

Appendix

A Details of the Empirical Evidence

Here, we describe the details of our empirical evidence on the macroeconomic announcement premium.

Data description We focus on the top five macroeconomic news ranked by investor attention among all macroeconomic announcements at the monthly or lower frequency. They are 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.[15]

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.[16] 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.[17] 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.

Equity return on announcement days Table 1 reports the mean, standard deviation, and Sharpe ratio of the annual return of the market, and the same moments for the return on announcement days. The announcement returns are calculated as the culmulative market returns on announcement days within a year. This is equivalent to the 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.

In Table 2, we 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 $(t+1)$. We present our results separately for each of the five news announcement and for all news, where standard errors are shown in parentheses.[18] Excess market returns are taken from Kenneth French's web

[15] 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.

[16] 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.

[17] 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.

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

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.

High frequency returns In Figure 1, we plot the average stock market returns over 30-minute intervals before and after news announcements. Here we use high frequency data for the $S\&P$ 500 SPDR that runs from 1997 to 2013 and comes from the TAQ database. Each second the median price of all transactions occurring that second is computed9. The price at lower frequency intervals (for example 30-min) is then constructed as the price for the last (most recent) second in that interval when transactions were observed. The exact time at which the news are released are reported by Bloomberg. In Figure 1, 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.[19]

B Examples of Dynamic Preferences and A-SDF

In this section, we show that most of the non-expected utility proposed in the literature can be represented in the form of (9).

- The recursive utility of Kreps and Porteus [48] and Epstein and Zin [27]. The recursive preference be generally represented as:

$$U_t = u^{-1}\left\{(1-\beta)\,u\,(C_t) + \beta\phi \circ h^{-1}E\left[h\,(U_{t+1})\right]\right\}. \tag{35}$$

  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,

$$V = u\,(U)\,, \tag{36}$$

  then the recursive relationship for V can be written in the form of (9) with the same u function in equation (35) and the 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 A-SDF can be written as:

$$m^*\,(V) \propto f'\,(V)\,, \tag{37}$$

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

[19] 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.

- The maxmin expected utility of Gilboa and Schmeidler [30]. The dynamic version of this preference is studied in Epstein and Schneider [25] and Chen and Epstein [16]. This preference can be represented as the special case of (9) 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 A-SDF 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 [54], the dynamic version of which is studied in Maccheroni, Marinacci, and Rustichini [55], 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 A-SDF for this class of preference is minimizing probability density.

- The multiplier preferences of Hansen and Sargent [34] and Strzalecki [64] 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 A-SDF 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 [28] can be written as (9) 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 A-SDF can be written as a function of continuation utility:

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

- The smooth ambiguity preferences of Klibanoff, Marinacci, and Mukerji [46] and Klibanoff, Marinacci, and Mukerji [47] can be represented as:

$$\mathcal{I}(V) = \phi^{-1}\left(\int_M \phi\left(\int_\Omega mV dP\right) d\mu(m)\right), \tag{38}$$

where μ is a probability measure on a set of probabilities densities M. The A-SDF can be written as a function of V:

$$m^* \propto \int_M \phi'\left(\int mV dP\right) m d\mu(m). \tag{39}$$

- The disappointment aversion preference can be represented as a concave utility function u, and the certainty equivalence functional $\mathcal{I}$ implicitly defined as follows: $\mu = \mathcal{I}[V]$, where μ is the unique solution to the following equation:

$$\phi(\mu) = \int \phi(V)\, dP - \theta \int_{\mu \geq V} [\phi(\mu) - \phi(V)]\, dP,$$

where ϕ is a strictly increasing and concave function.

C The Announcement SDF

In this section, we provide a formal statement of Theorem 1 and 2 in a fully dynamic model with a continuum state space. We first setup some notations and definitions.

C.1 Definitions and Notations

We use $\mathbf{R}$ to denote the real line and $\mathbf{R}^n$ to denote the $n-$dimensional Euclidean space. Let $(\Omega, \mathcal{F}, P)$ be a non-atomic probability space, where $\mathcal{F}$ is the associated Borel $\sigma-$algebra. A certainty equivalence functional $\mathcal{I}[\cdot]$ is a mapping $\mathcal{I}: L^2(\Omega, \mathcal{F}, P) \to \mathbf{R}$, where $L^2(\Omega, \mathcal{F}, P)$ is the set of square integrable random variables defined on $(\Omega, \mathcal{F}, P)$. We first state a definition of first order stochastic dominance (FSD) and second order stochastic dominance (SSD).

Definition 2. 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 3. 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.*[20]

Monotonicity with respect to FSD and SSD are defined as:

Definition 4. 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 $X_1 >_{FSD} X_2$ ($X_1 >_{SSD} X_2$).*

We also assume that the certainty equivalence functional $\mathcal{I}$ is normalized as in Strzalecki [65]:

Definition 5. Normalized: *$\mathcal{I}$ is normalized if $\mathcal{I}[k] = k$ whenever k is a constant.*

To construct the A-SDF from marginal utilities, we need some concepts from standard functional analysis (see for example, Luenberger [53] and Rall [59]) to impose a differentiability condition on the certainty equivalence functional. We use $\|\cdot\|$ to denote the L^2 norm on $L^2(\Omega, \mathcal{F}, P)$ and assume that $\mathcal{I}$ satisfy the following differentiability condition.

Definition 6. *(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 linear functional, $D\mathcal{I}[X] \in L^2(\Omega, \mathcal{F}, P)$ such that for all $\Delta X \in L^2(\Omega, \mathcal{F}, P)$*

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

[20] Our definition of SSD is the same as the standard concept of increasing risk (see Rothschild and Stiglitz [61] and Werner [67]). However, it is important to note that in our model, the certainty equivalence function $\mathcal{I}$ is defined on the space of continuation utilities rather than consumption.

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.[21]

The above assumption is made for two purposes. First it allows us to apply the envelope theorems in Milgrom and Segal [56] to establish differentiability of the value functions. Second, it allows us to compute the derivatives of $\mathcal{I}$ to construct the A-SDF and use derivatives of $\mathcal{I}$ to integrate back to recover the certainty equivalence functional.[22] Intuitively, we use the following operation to relate the certainty equivalence function $\mathcal{I}$ and its derivatives. $\forall$ X,$Y \in L^2(\Omega, \mathcal{F}, P)$, we can define $g(t) = \mathcal{I}[X + t(Y - X)]$ for $t \in [0, 1]$ and compute $\mathcal{I}(Y) - \mathcal{I}(X)$ as

$$\begin{aligned} \mathcal{I}[Y] - \mathcal{I}[X] &= g(1) - g(0) \\ &= \int_0^1 g'(t)\, dt \\ &= \int_0^1 \int_\Omega D\mathcal{I}[X + t(Y - X)](Y - X)\, dP dt. \end{aligned} \tag{40}$$

We note that Fréchet Differentiability with Lipschitz Derivatives guarantees that the function $g(t)$ is continuously differentiable. The differentiability of g is straightforward (see for example, Luenberger [53]). 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 $t_2 \to t_1$. This proves the validity of (40).

For later reference, it is useful to note that we can apply the mean value theorem on g, and write for some $\hat{t} \in (0, 1)$,

$$\mathcal{I}[Y] - \mathcal{I}[X] = \int_\Omega D\mathcal{I}\left[X + \hat{t}(Y - X)\right](Y - X)\, dP. \tag{41}$$

C.2 A Dynamic Model with Announcements

In this section, we describe a fully dynamic model with announcements. Consider a representative agent, pure-exchange economy where time is finite and indexed by $t = 1, 2, \cdots, T$. The endowment process is denoted $\{Y_t\}_{t=1}^T$, where $Y_t \geq 0$ for all t. In period t after Y_t is realized, agents receive a public announcement s_t that carries information about the future path of $\{Y_s\}_{s=t+1}^T$ but do not affect current-period endowment. We define a filtration $\mathcal{F}_1^- = \sigma(Y_1)$, $\mathcal{F}_1^+ = \sigma(Y_1, s_1)$, $\mathcal{F}_2^- = \mathcal{F}_1^+ \vee \sigma(Y_2)$, $\mathcal{F}_2^+ = \mathcal{F}_1^+ \vee \sigma(Y_2, s_2)$, $\cdots$.

[21] 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 $L^2(\Omega, \mathcal{F}, P)$.

[22] A weaker notion of differentiability, Gâtteaux differentiability is enough to gurantee the existence of A-SDF. However, the converse of Theorem 1 requires a stronger condition for differentiability, which is what we assume here.

In a 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: $\{R_A(j,t)(s_t)\}_{j=0,1,\cdots,J}$. Here, $R_A(j,t)(s_t)$ 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 post announcement asset market opens and the agents trade a vector of $J+1$ returns with payoff contingent on the realization of next period endowment: $\{R_P(j,t)(Y_{t+1})\}_{j=0,1,\cdots,J}$. That is, the returns, $R_A(j,t)(s_t)$ require one unit of investment at time t^+ after s_t is revealed and provide payoff at time $t+1^-$ that is a function of Y_{t+1}, To save notation, we suppress the dependence of $R_A(j,t)(s_t)$ on s_t and the dependence of $R_P(j,t)(Y_{t+1})$ on Y_{t+1} in the rest of this section. We also adopt the convention that $R_A(0,t)$ and $R_P(0,t)$ are risk-free returns.

We use V_t^+ to denote the value function of the representative agent's life-time utility at time t^+ after the signal s_t is announced, and V_t^- to denote the agent's value function at time t^- before s_t is known. The optimal consumption-portfolio choice problem of the agent 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 investment in the post-announcement asset market and write the corresponding consumption-portfolio choice problem as:

$$V_t^+(W) = \max_{C,\xi} \left\{u(C) + \beta \mathcal{I}\left[V_{t+1}^-(W')\right]\right\}$$
$$C + \sum\nolimits_{j=0}^J \xi_j = W$$
$$W' = \sum\nolimits_{j=0}^J \xi_j R_P(j,t).$$

Similarly, the optimal portfolio choice problem on the pre-announcement market is

$$V_t^-(W) = \max_{\zeta} \mathcal{I}\left[V_t^+(W')\right]$$
$$W' = W - \sum\nolimits_{j=0}^J \zeta_j + \sum\nolimits_{j=0}^J \zeta_j R_A(j,t),$$

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 $\left\{\{R_P(j,t), R_A(j,t)\}_{j=0,1,\cdots,J}\right\}_{t=1,2,\cdots T-1}$, 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 $\mathcal{I}[\cdot]$.

C.3 Existence of A-SDF

We first state and prove our result on the existence of A-SDF. Below, we state a generalization of Theorem 1 in a fully dynamic model with a non-atomic probability space.

Theorem 3. *(Existence of A-SDF)*

Suppose both u and $\mathcal{I}$ are Lipschitz continuous, Fréchet differentiable with Lipschitz continuous derivatives. 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 $R_A(0,t) = 1$. In addition, there exists a non-negative measurable function $m_t^* : \mathbf{R} \to \mathbf{R}$ such that

$$E\left[m_t^*\left(V_t^+\right)\{R_A(j,t) - 1\}\middle|\mathcal{F}_t^-\right] = 1 \quad for \quad all \quad j = 1, 2, \cdots J. \tag{42}$$

Under the regularity condition (54) below, $E\left[\left.m_t^*\left(V_t^+\right)\right|\mathcal{F}_t^-\right]=1$ and (42) can be written as:

$$E\left[\left.m_t^*\left(V_t^+\right)R_A\left(j,t\right)\right|\mathcal{F}_t^-\right]=1 \quad for \quad all \quad j=0,1,2,\cdots J. \tag{43}$$

Differentiability of value function To prove Theorem 3, we first 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 7. *$\mathcal{D}$ is the set of differentiable functions on the real line such that $\forall f\in\mathcal{D}$, i) f is Lipschitz continuous; ii) $\forall x\in\mathbf{R}$, $\frac{1}{h}\left[f\left(x+h-a\right)-f\left(x-a\right)\right]$ converges uniformly to $f'\left(x-a\right)$ in a. That is, $\forall\varepsilon>0$, there exists $\delta>0$ such that $|h|<\delta$ implies that $\left|\frac{1}{h}\left[f\left(x+h-a\right)-f\left(x-a\right)\right]-f'\left(x-a\right)\right|<\varepsilon$ 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,ξ), where W is the wealth level, and $\xi\in\mathbf{R}^{J+1}$ is a portfolio strategy:

$$\begin{aligned} f_v\left(W,\xi\right) &= u\left(W-\sum\nolimits_{j=0}^{J}\xi_j\right)+\beta\mathcal{I}\left[v\left(\sum\nolimits_{j=0}^{J}\xi_jR_j\right)\right], && (44)\\ g_v\left(W,\xi\right) &= \mathcal{I}\left[v\left(W+\sum\nolimits_{j=0}^{J}\xi_j\left(R_j-1\right)\right)\right]. && (45)\end{aligned}$$

Because $R_j\in L^2\left(\Omega,\mathcal{F},P\right)$ and v is Lipschitz continuous, $v\left(\sum_{j=0}^{J}\xi_jR_j\right)$ and $v\left(W-\sum_{j=0}^{J}\xi_j\left(R_j-1\right)\right)$ are both square integrable and equations (44) and (45) are well-defined. To apply the envelope theorem in Milgrom and Segal [56], we first need to establish the equi-differentiability of the family of functions $\left\{f_v\left(W,\xi\right)\right\}_\xi$ and $\left\{g_v\left(W,\xi\right)\right\}_\xi$:

Lemma 1. *Suppose $u,v\in\mathcal{D}$, as $h\to0$, both $\frac{1}{h}\left[f_v\left(W+h,\xi\right)-f_v\left(W,\xi\right)\right]$ and $\frac{1}{h}\left[g_v\left(W+h,\xi\right)-g_v\left(W,\xi\right)\right]$ converge uniformly for all ξ.*

Proof: *First,*

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

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

$$\frac{1}{h}\left[g_v\left(W+h,\xi\right)-g_v\left(W,\xi\right)\right]\to\frac{\partial}{\partial W}g_v\left(W,\xi\right) \tag{46}$$

and the convergence is uniform for all ξ. Note that

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

and

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

where we denote $\bar v\left(t\right)=tv\left(W+h-\sum_{j=0}^{J}\xi_j\left(R_j-1\right)\right)+\left(1-t\right)v\left(W-\sum_{j=0}^{J}\xi_j\left(R_j-1\right)\right)$ and applied equation (41). Also, denote $\bar v'\left(0\right)=v'\left(W-\sum_{j=0}^{J}\xi_j\left(R_j-1\right)\right)$, then the right hand side of (46) can be *written as* $\int_\Omega D\mathcal{I}\left[\bar{v}(0)\right]\bar{v}'(0)\,dP$*, we have:*

$$
\begin{aligned}
& \left|\frac{1}{h}\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\left(\bar{v}(1)-\bar{v}(0)\right)dP - \int_\Omega D\mathcal{I}\left[\bar{v}(0)\right]\bar{v}'(0)\,dP\right| \\
= & \left|\begin{array}{c} \frac{1}{h}\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\left(\bar{v}(1)-\bar{v}(0)\right)dP - \int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\bar{v}'(0)\,dP \\ +\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\bar{v}'(0)\,dP - \int_\Omega D\mathcal{I}\left[\bar{v}(0)\right]\bar{v}'(0)\,dP \end{array}\right| \\
\leq & \int_\Omega \left|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\right|\left|\frac{1}{h}\left(\bar{v}(1)-\bar{v}(0)\right)-\bar{v}'(0)\right|dP + \int_\Omega \left|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]-D\mathcal{I}\left[\bar{v}(0)\right]\right|\left|\bar{v}'(0)\right|dP \\
\leq & \left\|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\right\|\left\|\frac{1}{h}\left(\bar{v}(1)-\bar{v}(0)\right)-\bar{v}'(0)\right\| + \left\|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]-D\mathcal{I}\left[\bar{v}(0)\right]\right\|\left\|\bar{v}'(0)\right\|
\end{aligned} \tag{47}
$$

Because $v \in \mathcal{D}$*, for* h *small enough,* $\left|\frac{1}{h}\left(\bar{v}(1)-\bar{v}(0)\right)-\bar{v}'(0)\right| \leq \varepsilon$ *with probability one and* $\left\|\frac{1}{h}\left(\bar{v}(1)-\bar{v}(0)\right)-\bar{v}'(0)\right\| \leq \varepsilon$*. Also, because* $D\mathcal{I}$ *is Lipschitz continuous,* $\left\|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]-D\mathcal{I}\left[\bar{v}(0)\right]\right\| \leq K\left\|\bar{v}(1)-\bar{v}(0)\right\| \leq K^2 h$*, where the second inequality is due to the Lipschitz continuity of* v*. This proves the uniform convergence of (47).*

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 (44) and (45). It then follows from Lemma 1 that Theorem 3 in Milgrom and Segal [56] applies. Therefore, both* T^+v *and* T^-v *are differentiable, and*

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

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

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

$$
f_v\left(W_1,\xi(W_2)\right)-f_v\left(W_2,\xi(W_2)\right) \leq T^+v(W_1)-T^+v(W_2) \leq f_v\left(W_1,\xi(W_1)\right)-f_v\left(W_2,\xi(W_1)\right). \tag{48}
$$

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

$$
g_v\left(W_1,\xi(W_2)\right)-g_v\left(W_2,\xi(W_2)\right) \leq T^-v(W_1)-T^-v(W_2) \leq g_v\left(W_1,\xi(W_1)\right)-g_v\left(W_2,\xi(W_1)\right). \tag{49}
$$

Note that $\forall\xi$,

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

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

Finally, equations (48) and (49) can be used to show that the family of functions $\left\{T^+v(W-a)\right\}_a$ *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 (48), $\frac{1}{W_1-W_2}\left[T^{+}v(W_1) - T^{+}v(W_2)\right]$ *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.

Existence of A-SDF In this section, we establish the existence of SDF as stated in Theorem 3. We write the time t^{-} portfolio selection problem of the agent as

$$\max_{\zeta} \mathcal{I}\left[V_t^{+}\left(W + \sum\nolimits_{j=0}^{J} \zeta_j \left(R_A(j,t) - 1\right)\right)\middle| \mathcal{F}_t^{-}\right], \tag{50}$$

where we use the notation $\mathcal{I}\left[\cdot\middle|\mathcal{F}_t^{-}\right]$ 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 $R_A(0,t) = 1$. The value function $V^{+}(W)$ is determined by the the agent's portfolio choice problem at time t^{+} after the announcement s_t is made:

$$V_t^{+}(W) = \max_{\xi} u\left(W - \sum\nolimits_{j=0}^{J} \xi_j\right) + \beta \mathcal{I}\left[V_{t+1}^{-}\left(\sum\nolimits_{j=0}^{J} \xi_j R_P(j,t)\right)\middle| \mathcal{F}_t^{+}\right]. \tag{51}$$

Because the time-t^{+} value function, V_t^{+} is differentiable, and the certainty equivalent functional, $\mathcal{I}$ is Fréchet differentiable, $\mathcal{I}\left[V_t^{+}\left(W + \sum_{j=0}^{J} \zeta_j \left(R_A(j,t) - 1\right)\right)\middle| \mathcal{F}_t^{-}\right]$ is differentiable in ζ.[23] Therefore, the first order condition with respect to ζ_j implies that

$$E\left[D\mathcal{I}\left[V^{+}(W')\right] \frac{d}{dW} V_t^{+}(W') \left(R_A(j,t) - 1\right)\middle| \mathcal{F}_t^{-}\right] = 0, \tag{52}$$

where we denote $W' = W + \sum_{j=0}^{J} \hat{\zeta}_j \left(R_A(j,t) - 1\right)$ and $\hat{\zeta}$ is the optimal portfolio choice. Also, the envelop condition for (51) implies

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

where the last equality uses the market clearing condition because period-t consumption must equal to total endowment. Note that $u'(Y_t) > 0$ and is $\mathcal{F}_t^{-}$ measurable; therefore, (52) implies:

$$E\left[D\mathcal{I}\left[V_t^{+}(W)\right]\left(R_A(j,t) - 1\right)\middle| \mathcal{F}_t^{-}\right] = 0. \tag{53}$$

As we show in the next section, monotonicity of $\mathcal{I}$ guarantees that $D\mathcal{I} \geq 0$ with probability one. To derive equation (43), we need to assume a slightly stronger condition:

$$D\mathcal{I}[X] > 0 \text{ with strictly positive probability for all } X.^{24} \tag{54}$$

[23] See for example Proposition 1 in Chapter 7 of Luenberger [53].

[24] Note that monotonicity with respect to FSD implies that $D\mathcal{I}[X] \geq 0$ with probability one for all X. If

In this case, the SDF in (43) can be constructed as:

$$m_t^* = \frac{D\mathcal{I}\left[V_t^+(W)\right]}{E\left[\, D\mathcal{I}\left[V_t^+(W)\right]\middle|\, \mathcal{F}_t^-\right]}. \tag{55}$$

Now we constructed the A-SDF as the Fréchet Derivative of the certainty equivalence functional. Because $D\mathcal{I}\left[V_t^+(W)\right]$ is a linear functional on $L^2\left(\Omega, \mathcal{F}_t^+, P\right)$, it has a representation as an element in $L^2\left(\Omega, \mathcal{F}_t^+, P\right)$ by the Riesz representation theorem. To complete the proof of Theorem 3, we only need to show that m_t^* can be represented as a measurable function of continuation utility: $m_t^* = m_t^*\left(V_t^+\right)$ for some measurable function $m_t^* : \mathbf{R}, \to \mathbf{R}$. That is, m_t^* depends on s_t only through the continuation utility. Note that our definition of monotonicity with respect to FSD implies invariance with respect to distribution, that is, $\mathcal{I}[X] = \mathcal{I}[Y]$ whenever X and Y have the same distribution (If X has the same distribution of Y then both $X \leq_{FSD} Y$ and $Y \geq_{FSD} X$ are true). The following lemma establishes that invariance with respect to distribution implies the measurability of m_t^* with repect to V_t^+.

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 [44]). 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,

$$\lim_{\alpha \to 0} \frac{1}{\alpha}\left[V(X + \alpha Y) - V(X)\right] = \int (D\mathcal{I}[X] \circ T) \cdot Y dP \tag{56}$$

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, $V(X + \alpha Y) = V(X \circ T + \alpha Y) = V\left(X + \alpha Y \circ T^{-1}\right)$, 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 $\left[D\mathcal{I}[X] \cdot Y \circ T^{-1}\right] \circ T = D\mathcal{I}[X] \circ T \cdot Y$ have the same distribution

condition (54) does not hold, we must have $D\mathcal{I}[X] = 0$ 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 (54), our result implies that the A-SDF exists generically.

with $D\mathcal{I}[X] \times Y \circ T^{-1}$. This proves (56).

C.4 Generalized Risk Sensitivity and the Announcement Premium

In this section, we establish the link between generalized risk sensitivity and the existence of announcement premium. We first state a generalization of Theorem 2.

Theorem 4. *(Announcement Premium) Under the assumptions of Theorem 3,*

1. *$m_t^*\left(V_t^+\right) = 1$ for all V_t^+ if and only $\mathcal{I}$ is the expectation operator.*
2. *The following conditions are equivalent:*
   - *(a) The certainty equivalence functional $\mathcal{I}$ satisfies generalized risk sensitivity.*
   - *(b) The A-SDF $m_t^*\left(V_t^+\right)$ is a non-increasing function of continuation utility V_t^+.*
   - *(c) The announcement premium is non-negative for all payoffs of the form $f\left(V_s\right)$, where $f \geq 0$ and is strictly increasing at non-zero points.*

Because the A-SDF, m_t^* is constructed from the derivative of the certainty equivalence functional, it is clear from equation (55) that if $\mathcal{I}$ is expected utility, then m_t^* must be a constant. Conversely, if m_t^* is a constant, then $\mathcal{I}$ is linear and represents expected utility.

We first prove the equivalence between (a) and (b) by the following lemmas. Lemma 4 establishes that $m_t^*\left(V_t^+\right)$ is non-negative if and only if $\mathcal{I}$ is monotone with respect to FSD. Lemma 5 and 6 jointly establish that generalized risk sensitivity of $\mathcal{I}$ is equivalent to $m_t^*\left(V_t^+\right)$ being a non-increasing function of V_t^+.

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 $\mathcal{G} \subseteq \mathcal{F}$,*

$$\int D\mathcal{I}[X] \cdot (X - E[X|\mathcal{G}])\, dP \leq 0. \tag{57}$$

Proof: *Suppose condition (57) 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 (40),

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

where the last inequality uses (57).

Conversely, assuming $\mathcal{I}$ *is increasing in SSD, we prove (57) by contradiction. if (57) 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.*[25]

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 (57) is equivalent to $\eta(\cdot)$ being a non-increasing function, which is Lemma 6 below.

Lemma 6. *Condition (57) is equivalent to* $\eta(X)$ *being a non-increasing function of* X.

Proof: *First, we assume* $\eta(X)$ *is non-increasing. To prove (57), note that* $E[X|\mathcal{G}]$ *is 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* $X\leq E[X|\mathcal{G}]$.

[25] 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 [61] and Werner [67]). 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 $E[\phi(V_1)]\geq E[\phi(V_1+Z_1)]$.

Second, to prove the converse of the above statement by contradiction, we assume (57) 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

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

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

The following lemma establishes the equivalence between (b) and (c).

Lemma 7. *$m_t^*(V_t^+)$ is a non-increasing function of V_t^+ is equivalent to Condition 1.*

Proof: *If $m_t^*(V_t^+)$ is a non-decreasing function, then for any payoff f that is co-monotone with V_t^+, we have*

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

because $m_t^*(V_t^+)$ and $f(V_t^+)$ is negatively correlated. Conversely, if $m_t^*(v_1) < m_t^*(v_2)$ for some $v_1 < v_2$, both of which occur with positive probability, then define the payoff $f(\cdot)$ as

$$f(v) = \begin{cases} 1 & v = v_2 \\ 0 & v \neq v_2 \end{cases},$$

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

D Generalized Risk-Sensitive Preferences

D.1 Generalized risk sensitivity and uncertainty aversion

Quasi-concavity is sufficient but not necessary for generalized risk sensitivity We first present a lemma establishes that quasiconcavity implies generalized risk sensitivity.

Lemma 8. *Suppose $\mathcal{I} : L^2(\Omega, \mathcal{F}, P) \to \mathbf{R}$ is continuous invariant with respect to distribution, then quasiconcavity implies generalized risk sensitivity.*

Proof: Suppose $\mathcal{I}$ is continuous, invariant with respect to distribution, and quasiconcave. Let $X_1 \geq_{SSD} X_2$, we need to show that $\mathcal{I}[X_1] \geq \mathcal{I}[X_2]$. By the definition of second order stochastic dominance, there exist a random variable Y such that $E[Y|X_1] = 0$ and X_2 has the same distribution as $X_1 + Y$. Because $\mathcal{I}$ is invariant with respect to distribution, $\mathcal{I}[X_1 + Y] = \mathcal{I}[X_2]$. Let $T : \Omega \to \Omega$ be any measure preserving *transformation such that the invariant $\sigma-$field of T differs from the $\sigma-$field generated by X only by sets of measure zero (see exercise 17.43 in Kechris [44]), then quasiconcavity implies that*

$$\mathcal{I}\left[\frac{1}{2}(X_1+Y)+\frac{1}{2}(X_1+Y)\circ T\right] \geq \min\left\{\mathcal{I}[X_1+Y],\ \mathcal{I}[(X_1+Y)\circ T]\right\}.$$

Note that because T is measure preserving and $\mathcal{I}$ is distribution invariant, we have $\mathcal{I}[X_1+Y] = \mathcal{I}[(X_1+Y)\circ T]$. Therefore, $\mathcal{I}\left[\frac{1}{2}(X_1+Y)+\frac{1}{2}(X_1+Y)\circ T\right] \geq \mathcal{I}[X_1+Y]$. It is therefore straightforward to show that $\mathcal{I}\left[\frac{1}{N}\sum_{j=0}^{N-1}(X_1+Y)\circ T^j\right] \geq \mathcal{I}[X_1+Y]$ for all N by induction. Note that $\frac{1}{N}\sum_{j=0}^{N-1}(X_1+Y)\circ T^j \to E[X_1+Y|X_1] = X_1$ by Birkhoff's ergodic theorem (note that the invariance $\sigma-$field of T is $\sigma(X)$ by construction). Continuity of I then implies $I[X_1] \geq \mathcal{I}[X_1+Y] = \mathcal{I}[X_2]$, that is, $\mathcal{I}$ satisfies generalized risk sensitivity.

It is clear from Lemma 8 that under continuity, the following condition is sufficient for generalized risk sensitivty:

$$\mathcal{I}[\lambda X+(1-\lambda)Y] \geq \mathcal{I}[X] \quad for \ \ all\ \lambda\in[0,1] \quad if\ X\ and\ Y\ have\ the\ same\ distribution. \tag{58}$$

Clearly, this condition is weaker than quasiconcavity.[26]

Next, we provide a counterexample of $\mathcal{I}$ that satisfies generalized risk sensitivity but is not quasiconcave. We continue to use the two-period example in Section 3, where we assume $\pi(H)=\pi(L)=\frac{1}{2}$. Given there are two states, random variables can be represented as vectors. We denote $\mathbf{X}=\{(x_H,x_L): 0\leq x_H, x_L\leq B\}$ to be the set of random variables bounded by B. Let $\mathcal{I}$ be the certainty equivalence functional defined on X such that

$$\forall X\in\mathbf{X},\ \ \mathcal{I}[X]=\phi^{-1}\left\{\min_{m\in M} E[m\phi(X)]\right\}, \ \ with\ \ \phi(x)=e^x, \tag{59}$$

where $M = \left\{(m_H,m_L): m_H+m_L=1,\ \max\left\{\frac{m_H}{m_L},\frac{m_L}{m_H}\right\}\leq\eta\right\}$ is a collection of density of probability meausres and the parameter $\eta\geq e^B$. Note that $\mathcal{I}$ defined in (59) is not concave because $\phi(x)$ is a strictly convex function. Below we show that $\mathcal{I}$ satisfy generalized risk sensitivity, but is not quasiconcavity.

Using (58), to establish generalized risk sensitivity, we need to show that for any X, $Y\in\mathbf{X}$ such that X and X have the same distribution, $\mathcal{I}[\lambda X+(1-\lambda)Y]\geq\mathcal{I}[X]$. Without loss of generality, we assume $X=[x_H,x_L]$ with $x_H>x_L$. Because Y has the same distribution with X, $Y=[x_L,x_H]$. We first show that for all $\lambda\geq\frac{1}{2}$,

$$\mathcal{I}[\lambda X+(1-\lambda)Y]\geq\mathcal{I}[X].$$

Because ϕ is strictly increasing, it is enough to prove that for all $\lambda\in\left[\frac{1}{2},1\right]$,

$$\frac{d}{d\lambda}\phi(\mathcal{I}[\lambda X+(1-\lambda)Y])\leq 0. \tag{60}$$

Because $x_H>x_L$, for all $\lambda\geq\frac{1}{2}$, $\lambda x_H+(1-\lambda)x_L\geq\lambda x_L+(1-\lambda)x_H$ and

$$\phi(\mathcal{I}[\lambda X+(1-\lambda)Y])=\frac{1}{2}m_H^*\phi(\lambda x_H+(1-\lambda)x_L)+\frac{1}{2}m_L^*\phi(\lambda x_L+(1-\lambda)x_H),$$

[26] Lemma 8 requires the underlying probability space to be non-atomic. The statement in (58) remains true if we assume that the underlying probability space is finite with equality probility for all states as in the main text of the paper.

where $m_H + m_L = 1$ and $\frac{m_H}{m_L} = \frac{1}{\eta}$. Therefore,

$$\begin{aligned}\frac{d}{d\lambda}\phi\left(\mathcal{I}\left[\lambda X + (1-\lambda)Y\right]\right) &= \frac{1}{2}\left[m_H^*\phi'\left(\lambda x_H + (1-\lambda)x_L\right) - m_L^*\phi'\left(\lambda x_L + (1-\lambda)x_H\right)\right](x_H - x_L) \\ &= \frac{1}{2}(x_H - x_L)\left\{m_H^* e^{\lambda x_H + (1-\lambda)x_L} - m_L^* e^{\lambda x_L + (1-\lambda)x_H}\right\}.\end{aligned}$$

Note that

$$\frac{m_H^* e^{\lambda x_H + (1-\lambda)x_L}}{m_L^* e^{\lambda x_L + (1-\lambda)x_H}} = \frac{1}{\eta}e^{(2\lambda-1)(x_H - x_L)} \leq \frac{1}{\eta}e^B \leq 1.$$

This proves (60). Similarly, one can prove $\mathcal{I}\left[\lambda X + (1-\lambda)Y\right] \geq \mathcal{I}[Y]$ for all $\lambda \in \left[0, \frac{1}{2}\right]$. This established generalized risk sensitivity.

To see $\mathcal{I}$ is not quasiconcave, consider $X_1 = [1, 0]$, and $X_2 = [x, x]$, where $x = \ln\frac{\eta+e}{\eta+1}$. One can verify that $\mathcal{I}[X_1] = \mathcal{I}[X_2]$, but $\mathcal{I}\left[\frac{1}{2}X_1 + \frac{1}{2}X_2\right] > \mathcal{I}[X_1]$, contradicting quasiconcavity.

Second order expected utility Certainty equivalence functionals of the form $\mathcal{I}[V] = \phi^{-1}\left(E\left[\phi(V)\right]\right)$, where ϕ is strictly increasing is called second order expected utility in Ergin and Gul [28]. For this class of preferences, generalized risk sensitivity is equivalent to quasiconcavity, which is also equivalent to the concavity of ϕ. To see this, suppose ϕ is concave, it is straightforward to show that $\mathcal{I}[\cdot]$ quasiconcave and satisfies generalized risk sensitivity by Lemma 8. Conversely, suppose $\mathcal{I}[\cdot]$ satisfies generalized risk sensitivity then $E[\phi(X)] \geq E[\phi(Y)]$ whenever $X \geq_{SSD} Y$. By remark B on page 240 of Rothschild and Stiglitz [61], ϕ is concave.

Generalized sensitivity is equivalent to quasiconcavity for smooth ambiguity preferences Using the results in Klibanoff, Marinacci, and Mukerji [46, 47], it straightforward to show that for the class of smooth ambiguity preference, concavity of ϕ is equivalent to the quasiconcavity of $\mathcal{I}$. As a result, quasiconcavity implies generalized risk sensitivity by Lemma 8. The nontrivial part of the above claim is that generalized risk sensitivity implies the concavity of ϕ. To see this is true, note that invariance with respect to distribution implies that the probability measure $\mu(x)$ must satisfy the following property: for all $A \in \mathcal{F}$,

$$\int\int_A dP_x d\mu(x) = P(A).$$

Clearly, generalized risk sensitivity implies that $I[E[V]] \geq I[V]$, for all $V \in L^2(\Omega, \mathcal{F}, P)$. That is,

$$\int \phi\left(E^x[V]\right)d\mu(x) \leq \phi\left(E[V]\right).$$

The fact that the above inequality has to hold for all V and $E[V] = \int E^x[V]\,d\mu(x)$ implies that ϕ must be concave.

D.2 Generalized risk sensitivity and preference for early resolution of uncertainty

Below, we provide details of Remark 3 in the discussion of preference for early resolution of uncertainty in Section 4.2. In order to show that generalized risk sensitivity is neither necessary nor sufficient for preference for early resolution of uncertainty, we provide two examples. The first example is a preference that satisfies generalized risk sensitivity but strictly prefers later resolution of uncertainty and the second example is a certainty equivalence functional that prefers early resolution of uncertainty but is strictly decreasing in second order stochastic dominance.

Example 1. *Consider the following utility function in the two period example:*

$$u(C) = C - b, \quad where \quad b = 2; \quad \mathcal{I}(X) = \left(E\sqrt{X}\right)^2.$$

It straight forward to check that $\mathcal{I}$ is quasi-concave therefore satisfy generalized risk sensitivity. Below we verify that this utility function prefers late resolution of uncertainty when the following consumption plan is presented: $C_0 = 1$, $C_H = 3.21$, and $C_L = 3$, where the distribution of consumption is given by $\pi(H) = \pi(L) = \frac{1}{2}$.

The utility with early resolution of uncertainty is given by:

$$W_{Early} = \mathcal{I}[u(C_0) + u(C_1)].$$

It is straightforward to show that:

$$u(C_0) + u(C_H) = 0.21; \quad u(C_0) + u(C_L) = 0$$

Therefore,

$$W_{Early} = \left[0.5 \times \sqrt{0.21} + 0.5 \times \sqrt{0}\right]^2 = 0.0525$$

The utility for late resolution of uncertainty is given by:

$$W_{Late} = u(C_0) + \mathcal{I}[u(C_1)] = 0.1025.$$

Example 2. *Consider the following preference:*

$$u(C) = C - b \text{ with } b = 2, \qquad I(X) = \sqrt{E[X^2]}, \quad and \quad \beta = 1.$$

Because X^2 is a strictly convex function, the certainty equivalence functional $\mathcal{I}$ is strictly decreasing in second order stochastic dominance. To see that the agent prefers early resolution of uncertainty, we consider the same numerical example as in Example 1. It is straightforward to verify that the utility for early resolution of uncertainty is

$$W_{Early} = \mathcal{I}[u(C_0) + u(C_1)] = 0.1485,$$

and the utility for later resolution is:

$$W_{Late} = u(C_0) + \mathcal{I}[u(C_1)] = 0.11.$$

D.3 Asset Pricing Implications

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:

$$u(C_0) + \beta E[u(C_1 + \alpha C_0)], \quad \alpha \in (0, 1). \tag{61}$$

With $\alpha < 0$, this is the internal habit model (for example, Boldrin, Christiano, and Fisher [11]). We show in this case the announcement premium is negative. The class of preferences with $\alpha > 0$ is discussed in Dunn and Singleton [23] and Heaton [39]. In general, the announcement premium can be positive or negative with $\alpha > 0$.

If the representative agent's preference is given by (61), the pre-announcement price and post-announcement price can be represented by

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

repectively. Therefore, the expected announcement return is

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

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 announcement premium is negative.

2. Durable consumption goods: with $\alpha > 0$, (61) is a special case of the durable consumption goods model of Dunn and Singleton [23]. 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 announcement premium 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 announcement premium. An example of payoff that satisfies this condition is $X = \frac{\sqrt{u'(C_0)+\alpha\beta u'(C_1+\alpha C_0)}}{u'(C_1+\alpha C_0)}$.

E Details of the Continuous-time model

E.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:

$$q_t = \frac{\sigma_x^2 \left(1 - e^{-2\hat{a}(t+t^*-nT)}\right)}{(\hat{a}-a)\, e^{-2\hat{a}(t+t^*-nT)} + a + \hat{a}}, \tag{62}$$

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):

$$\frac{1}{q_{nT}^+} = \frac{1}{\sigma_S^2} + \frac{1}{q_{nT}^-}. \tag{63}$$

Given a q_0, equations (62) and (63) 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 [22], 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 [22]), 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

$$\bar{V}\,(\hat{x}_t, t, C_t) = \frac{1}{1-\gamma} H\,(\hat{x}_t, t)\, C_t^{1-\gamma}, \tag{64}$$

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

$$\begin{aligned} -\frac{\rho}{1-\gamma} \ln H\,(\hat{x}, t)\, H\,(x, t) + \left(\hat{x} - \frac{1}{2}\gamma\sigma^2\right) H\,(\hat{x}, t) + \frac{1}{1-\gamma} H_t\,(\hat{x}_t, t) \\ + \left[\frac{1}{1-\gamma} a_x\,(\bar{x} - \hat{x}) + q_t\right] H_x\,(\hat{x}, t) + \frac{1}{2}\frac{1}{1-\gamma} H_{xx}\,(\hat{x}, t)\, \frac{q_t^2}{\sigma^2} = 0, \end{aligned} \tag{65}$$

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

$$H\,(\hat{x}_{nT}^-, nT) = E\left[\, H\,(\hat{x}_{nT}^+, nT) \middle|\, \hat{x}_{nT}^-, q_{nT}^-\right]. \tag{66}$$

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) (65) together with the boundary condition (66) 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:

$$-\rho h(t) + h'(t) + f(t) = 0, \tag{67}$$

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 (67) 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 (66) 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

$$r(\hat{x}, t) = \beta + \hat{x} - \gamma\sigma^2 + \frac{1-\gamma}{a_x+\rho} q_t$$

is the risk-free interest rate, and

$$\sigma_\pi(t) = \gamma\sigma + \frac{\gamma-1}{a_x+\rho}\frac{q_t}{\sigma}$$

is the market price of the Brownian motion risk.

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

$$p(\hat{x}_t, t) D_t = E_t \left[\int_t^{(n+1)T} \frac{\pi_s}{\pi_t} D_s ds + \frac{\pi_{(n+1)T}}{\pi_t} p\left(\hat{x}^-_{(n+1)T}, (n+1)T^-\right) D_{(n+1)T} \right].$$

The above present value relationship implies that

$$\pi_t D_t + \lim_{\Delta \to 0} \frac{1}{\Delta} \left\{ E_t \left[\pi_{t+\Delta} p(\hat{x}_{t+\Delta}, t+\Delta) D_{t+\Delta} \right] - \pi_t p(\hat{x}_t, t) D_t \right\} = 0. \tag{68}$$

Equation (68) can be used to show that the price-to-dividend ration function must satisfy the following PDE:

$$1 - p(\hat{x}, t)\varpi(\hat{x}, t) + p_t(\hat{x}, t) - p_x(\hat{x}, t)\nu(\hat{x}, t) + \frac{1}{2} p_{xx}(\hat{x}, t) \frac{q^2(t)}{\sigma^2} = 0, \tag{69}$$

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

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

Alos, equation (34) can be used to derive the following boundary condition for $p(\hat{x}, t)$:

$$p\left(\hat{x}_T^-, T^-\right) = \frac{E\left[e^{\frac{1-\gamma}{a_x+\rho}\hat{x}_T^+}p\left(\hat{x}_T^+, T^+\right)\middle|\,\hat{x}_T^-, q_T^-\right]}{e^{\frac{1-\gamma}{a_x+\rho}\hat{x}_T^- + \frac{1}{2}\left(\frac{1-\gamma}{a_x+\rho}\right)^2\left[q_T^- - q_T^+\right]}}. \tag{70}$$

Again, we focus on the steady-state and denote $p(\hat{x}, 0) = p(\hat{x}, nT^+)$, and $p(\hat{x}, T) = p(\hat{x}, nT^-)$. Under this condition PDE (69) 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,

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

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

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

E.2 Numerical Solutions

To solve the PDE (69) with the boundary condition (70), we consider the following auxiliary problem:

$$p(x_t, t) = E\left[\int_t^T e^{-\int_t^s \varpi(x_u, u)du}ds + e^{-\int_t^T \varpi(x_u, u)du}p(x_T, T)\right], \tag{71}$$

where the state variable x_t follows the law of motion;

$$dx_t = -\nu(\hat{x}, t)\,dt + \frac{q(t)}{\sigma}dB_t. \tag{72}$$

Note that the solution to (71) and (70) satisfies the same PDE. Given an initial guess of the pre-new price-to-dividend ratio, $p^-(x_\tau, \tau)$, we can solve (71) by the Markov chain approximation method (Kushner and Dupuis [49]):

1. We first start with an initial guess of a pre-announcement price to dividend ratio function, $p(x_T, T)$.
2. We construct a locally consistent Markov chain approximation of of the diffusion process (72) as follows. We choose a small dx, let $Q = |\nu(\hat{x}, t)|\,dx + \left(\frac{q(t)}{\sigma}\right)^2$, 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:

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

One can verify that as $dx \to 0$, the above Markov chain converges to the diffusion process (72) (In the language of Kushner and Dupuis [49], this is a Markov chain that is locally consistent with the diffusion process (72)).

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 (71) 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-announcement price to dividend ratio function, $p(x_T, T)$ using (70):

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

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

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] Alvarez, Fernando, and Urban Jermann, 2004, Using Asset Prices to Measure the Cost of Business Cycles, *Journal of Political Economy* 112, 1223–1256.

[4] Alvarez, Fernando, and Urban Jermann, 2005, Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth, *Econometrica* 73, 1977–2016.

[5] Backus, David, Mikhail Chernov, and Stanley Zin, 2014, Sources of Entropy in Representative Agent Models, *Journal of Finance* LXIX, 51–99.

[6] Bansal, Ravi, 2007, Long Run Risks and Financial Markets, *The Review of the St. Louis Federal Reserve Bank* 89, 1–17.

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

[8] Bansal, Ravi, and Ivan Shaliastovich, 2011, Learning and Asset Price Jumps, *Review of Financial Studies* 24, 2738–2780.

[9] 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.

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

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

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

[13] Brusa, Francesca, Pavel Savor, and Mungo Wilson, 2015, One Central Bank to Rule Them All, Working paper, Temple University.

[14] 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.

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

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

[17] Chew, Soo Hong, and Mei Hui Mao, 1995, A Schur Concave Characterization of Risk Aversion for Non-expected Utility Preferences, *Journal of Economic Theory* 67, 402–435.

[18] Cieslak, Anna, Adair Morse, and Annette Vissing-Jorgensen, 2015, Stock Returns over the FOMC Cycle, Working paper, UC Berkeley.

[19] Colacito, Riccardo, and Mariano Massimiliano Croce, 2011, Risks for the Long Run and the Real Exchange Rate, *The Journal of Political Economy* 119, 153–181.

[20] Constantinides, George M., 1990, Habit Formation: A Resolution of the Equity Premium Puzzle, *The Journal of Political Economy* 98, 519–543.

[21] Croce, Mariano Massimiliano, 2014, Long-run productivity risk: A new hope for production-based asset pricing?, *Journal of Monetary Economics* 66, 13–31.

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

[23] 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.

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

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

[26] Epstein, Larry, and Martin Schneider, 2010, Ambiguity and Asset Markets, *Annual Reviews of Financial Markets* 2, 315–334.

[27] 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.

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

[29] Farhi, Emmanuel, and Ivan Werning, 2008, Optimal Savings Distortions with Recursive Preferences, *Journal of Monetary Economics, Carnegie Rochester Series* 55, 21–42.

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

[31] Gul, Faruk, 1991, A Theory of Disappointment Aversion, *Econometrica* 59, 667–686.

[32] Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, Consumption strikes back? Measuring longrun risk, *Journal of Political Economy* 116, 260–302.

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

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

[35] 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.

[36] Hansen, Lars Peter, and Thomas J. Sargent, 2005, Robust Estimation and Control under Commitment, *Journal of Economic Theory* 124, 258–301.

[37] Hansen, Lars Peter, and Thomas J. Sargent, 2007, Recursive Robust Estimation and Control without Commitment, *Journal of Economic Theory* 136, 1–27.

[38] Hansen, Lars Peter, and Thomas J. Sargent, 2012, Three Types of Amniguity, Working paper, New York University.

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

[40] Ju, Nengjiu, and Jianjun Miao, 2012, Ambiguity, Learning, and Asset Returns, *Econometrica* 80, 559–591.

[41] Kaltenbrunner, Georg, and Lars A. Lochstoer, 2010, Long-Run Risk through Consumption Smoothing, *The Review of Financial Studies* 23, 3190–3224.

[42] Karantounias, Anastasios G., 2013, Managing pessimistic expectations and fiscal policy, *Theoretical Economics* 8, 193–231.

[43] Karantounias, Anastasios G., 2015, Optimal Fiscal Policy with Recursive Preferences, *Federal Reserve Bank of Atlanta, Working Paper 2013-7.*

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

[45] Kelly, Michael A., and Steven P. Clark, 2011, Returns in trading versus non-trading hours: The difference is day and night, *Journal of Asset Management* 12, 132–145.

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

[47] Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2009, Recursive Smooth Ambiguity Preferences, *Journal of Economic Theory* 144, 930–976.

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

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

[50] Kwon, Hyosung, and Jianjun Miao, 2013a, Three Types of Robust Ramsey Problem in a Linear-Quadratic Framework, *Working paper, Boston University.*

[51] Kwon, Hyosung, and Jianjun Miao, 2013b, Woodford's Approach to Robust Policy Analysis in a Linear-Quadratic Framework, *Working paper, Boston University.*

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

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

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

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

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

[57] Muller, Alfred, and Dietrich Stoyan, 2002, *Comparison Methods for Stochastic Models and Risks.* (Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd).

[58] Polk, Christopher, Dong Lou, and Spyros Skouras, 2016, A Tug of War: Overnight vs. Intraday Expected Returns, Working paper, London School of Economics.

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

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

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

[62] Routledge, Bryan R., and Stanley E. Zin, 2010, Generalized Disappointment Aversion and Asset Prices, *The Journal of Finance* 65, 1303–1332.

[63] 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.

[64] Strzalecki, Tomasz, 2011, Axiomatic Foundations of Multiplier Preferences, *Econometrica* 79, 47–73.

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

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

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

[68] Woodford, Michael, 2010, Robustly Optimal Monetary Policy with Near-Rational Expectations, *American Economic Review* 100, 274–303.

Table 1

Market Return on Announcement and Non-announcement Days

1961-2014

	# per Year	Mean Ex	Std	Sharpe R.
Market	252	6.19	15.59	0.40
All	30	3.36	6.23	0.54
EMPL/NFP	12	0.74	3.42	0.22
PPI	12	0.91	3.84	0.24
FOMC	8	2.69	3.36	0.80
GDP	8	1.20	3.41	0.35
ISM	12	2.21	5.06	0.44
None	222	2.82	14.29	0.20

1997-2014

	# per Year	Mean Ex	Std	Sharpe R.
Market	232	7.44	20.14	0.37
All	50	8.24	9.37	0.88
EMPL/NFP	12	1.85	4.45	0.42
PPI	12	0.90	4.62	0.20
FOMC	8	2.91	3.54	0.82
GDP	8	1.20	3.41	0.35
ISM	12	2.21	5.06	0.44
None	202	−0.78	17.79	−0.04

This table documents the mean excess return of the market, its standard deviation and Sharpe ratio on announcement and non-announcement days. The mean excess return is computed as the average daily market excess return on event days multiplied by the average number of events per year. The first column is the average number of events per year during the sample period. The release dates for unemployment/non-farm payroll (EMPL/NFP) and producer price index (PPI) come from the BLS with data starting in 1961 and 1971 respectively. The dates of Federal Open Market Committee (FOMC) meetings are taken from the Federal Reserve's website and begin in 1994. Gross domestic product (GDP) release dates come from the BEA's website and Institute for Supply Management's Manufacturing Report (ISM) announcement dates come from Bloomberg. Both are available after 1997.

Table 2

Average Daily Return around Announcements (Basis Points)

1961-2014

	# of events	t-1	t	t+1
All	1,624	1.77 (0.66)	11.21(3.99)	0.84(0.27)
EMPL/NFP	648	4.24 (0.11)	6.18(1.59)	−2.47(−0.65)
PPI	527	−3.09 (−0.73)	7.61(1.57)	−1.05(−0.18)
FOMC	168	10.57 (0.99)	33.60(3.68)	1.10(0.10)
GDP	144	−0.08 (−0.01)	14.99(1.50)	13.64(1.32)
EMPL	216	−0.60 (−0.08)	18.42(1.86)	5.76(0.68)
None	11,968		1.27(1.40)	

1997-2014

	# of events	t-1	t	t+1
All	897	−0.12(−0.03)	16.48(3.76)	0.21(0.05)
EMPL/NFP	216	−1.09(−0.12)	15.44(1.77)	−7.15(−0.87)
PPI	216	−8.38(−1.10)	7.53(0.83)	−6.20(−0.68)
FOMC	144	−13.62(1.11)	36.35(3.49)	−1.34(−0.11)
GDP	144	−0.08(−0.01)	14.99(1.50)	13.64(1.32)
EMPL	216	−0.13(−0.02)	18.42(1.86)	5.76(0.68)
None	3,633		−0.36(−0.18)	

This Table documents the average daily return on the trading before announcements (t-1), at announcements (t), and after announcements (t+1). Returns are measured in basis points and T-stats based on Newey-West (5 lags) standard errors are included in parenthesis. The first column is the total number of events during the sample period. The release dates for unemployment/non-farm payroll (EMPL/NFP) and producer price index (PPI) come from the BLS with data starting in 1961 and 1971 respectively. The dates of Federal Open Market Committee (FOMC) meetings are taken from the Federal Reserve's website and begin in 1994. Gross domestic product (GDP) release dates come from the BEA's website and Institute for Supply Management's Manufacturing Report (ISM) announcement dates come from Bloomberg. Both are available after 1997.

Table 3

Intraday and Overnight Return with and without Announcement

Intrady Returns			
	# of Events	Mean (StErr)	Std
All Intraday	4278	−0.55 (1.67)	109
Announcement	336	17.0 (6.43)	118
FOMC	136	23.2 (10.0)	117
ISM	204	12.1 (8.26)	118
No Announcement	3942	−2.05 (1.72)	108

Overnight Returns			
	# of Events	Mean (StErr)	Std
All Overnight	4277	3.52 (1.06)	69.4
Announcement	544	9.32 (3.38)	78.8
NFP	204	16.2 (5.47)	78.1
PPI	204	−2.17 (5.86)	83.7
GDP	136	16.2 (6.02)	70.2
No Announcement	3733	2.67 (1.11)	67.9

This table decomposes intraday and overnight returns into announcement day returns and non-announcement day returns. The first column is the total number of events during the sample period of 1997-2014. The mean return on event days is measured in basis points with standard error of the point estimate in parenthesis.

Table 4

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 4 presents the calibrated parameters of our learning model.

Table 5

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 5 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.