Risk Preferences and the Macroeconomic Announcement Premium

Hengjie Ai and Ravi Bansal *

January 28, 2018

This paper develops a revealed preference theory for the 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. We provide a characterization theorem for the set of intertemporal preferences that generates a non-negative announcement premium. Our theory establishes that the announcement premium identifies a significant deviation from time-separable expected utility and provides asset-market-based evidence for a large class of non-expected utility models. We also provide conditions under which asset prices may rise prior to some macroeconomic announcements and exhibit a pre-announcement drift.

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, Jaroslav Borovička, V.V. Chari, Anna Cieslak, Anisha Ghosh, Lars Hansen, John Heaton, Peter Klibanoff, Lars Lochstoer, Monika Piazzesi, Jacob Sagi, Martin Schneider, Jan Werner, Mungo Wilson, and Stanley Zin; participants at the Adam Smith Asset Pricing Workshop, AFA 2017, the ASU Sonoran Winter Finance Conference, CITE 2017, NBER Neemrana, and SED 2016; and seminar participants at the Duke Finance workshop, HKU, HKUST, SAIF, SMU, NTU, NUS, the University of Minnesota Economic Theory workshop, University of Texas at Austin, and University of Toronto for their helpful comments on the paper.

1 Introduction

In this paper, we develop a revealed preference theory for the risk premium for pre-scheduled macroeconomic announcements. We demonstrate that the premium around macroeconomic announcements provides asset-market-based evidence that establishes the importance of incorporating non-expected utility analysis in macro and asset pricing models.

Macroeconomic announcements, such as the release of the employment report and the Federal Open Market Committee (FOMC) statements, resolve uncertainty about the future course of the macroeconomy, and therefore asset prices react to these announcements instantaneously. Empirically, a large fraction of the market equity premium is realized within a small number of trading days with significant macroeconomic announcements. In the 1961-2014 period, during the thirty days per year with significant macroeconomic announcements, the cumulative excess returns of the S&P 500 index averaged 3.36%, which accounts for 55% of the total annual equity premium of 6.19%. The average return on days with macroeconomic announcements is 11.2 basis points (bps), which is significantly higher than the 1.27 bps average return on non-announcement days. High-frequency-data-based evidence shows that much of this premium is realized within hourly windows around announcements, or within a few trading hours prior to the announcements.

To understand the above features of the financial markets, we develop a theoretical model that allows macroeconomic announcements to carry information about the prospect of future economic growth. In this setup, we characterize the set of intertemporal preferences for the representative consumer under which macroeconomic announcements are associated with realizations of the market equity premium.

Throughout the paper, we focus on a representative-agent model and assume that aggregate consumption does not instantaneously respond to the macroeconomic announcements, whereas asset prices do. This assumption is well motivated because the announcement returns are realized in hourly windows around the announcements and the consumption response, if any, at this frequency is not likely to be significant enough to rationalize the magnitude of the premium.[1]

We follow Strzalecki (2013) and consider intertemporal preferences that can be represented recursively as $V_t = u(C_t) + \beta \mathcal{I}[V_{t+1}]$, where u maps current-period consumption into utility, and $\mathcal{I}$ maps the next-period continuation utility into its certainty equivalent. Our main result is that announcements are associated with realizations of the premium if and only

[1]This assumption, as further discussed in section 3.1, is also consistent with the empirical findings that consumption does not co-move contemporaneously with the stock market return (e.g., see Hall (1978)).

if the certainty equivalent functional, $\mathcal{I}$, is non-decreasing with respect to second-order stochastic dominance, a property we define as generalized risk sensitivity. This theorem has two immediate implications. First, intertemporal preferences have a time-separable expected utility representation if and only if the announcement premium is zero for all assets. Second, announcement premiums must be compensation for generalized risk sensitivity in the certainty equivalent functional, $\mathcal{I}$, and not compensation for the risk aversion of the Von Neumann–Morgenstern utility function, u.

The macroeconomic announcement premium, therefore, provides asset-market-based evidence that identifies a key aspect of investors' preferences not captured by the time-separable expected utility. Non-expected utilities, such as the recursive utility (Kreps and Porteus (1978), Epstein and Zin (1989)), the maximin expected utility (Gilboa and Schmeidler (1989)), the robust control model (Hansen and Sargent (2007)), and the smooth ambiguity model (Klibanoff, Marinacci, and Mukerji (2005)), among others, are widely applied in asset pricing studies to enhance the model-implied market price of risk. We show that generalized risk sensitivity is the key property of these models that distinguishes their asset pricing implications from expected utility. The large magnitude of the announcement premium in the data can be interpreted as 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 often be interpreted as the density of a probability distortion. We demonstrate that the probability distortion component 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 of the announcement premium is informative about the relative importance of the two components of the stochastic discount factor for quantitative asset pricing models. We find that the Sharpe ratio on announcement days is significantly higher than that on non-announcement days. Therefore, a substantial fraction of the volatility of the stochastic discount factor must come from generalized risk sensitivity.

Generalized risk sensitivity is precisely the property of preferences that requires a risk compensation for news. The long-run risks literature typically uses the Epstein and Zin (1989) utility with a preference for early resolution of uncertainty to generate a risk premium for news shocks. We show that preference for early resolution of uncertainty is not a necessary condition for generalized risk sensitivity and provide examples of preferences that require a compensation for news shocks but do not exhibit a preference for early resolution.

Our theoretical framework also provides an explanation for the difference between the timing of the realization of the premiums for FOMC announcements and that for other macroeconomic announcements. Using high-frequency data, Lucca and Moench (2015) document a pre-announcement drift for FOMC announcements, but not for other macroeconomic announcements. Specifically, equity premiums start to materialize a few hours prior to the official FOMC announcements, but there is no such pattern in other announcements. Our theorem implies the existence of a pre-announcement drift if investors receive informative signals before the announcements. Based on this idea, we present a continuous-time model to account for the pre-announcement drift in FOMC announcements and its absence in other macroeconomic announcements.

Our theoretical framework does not allow for several models of time-non-separable utilities widely applied in the asset pricing literature, so we study them separately. We establish that the external habit model of Campbell and Cochrane (1999) generates a zero announcement premium, and the internal habit model of Constantinides (1990) and Boldrin, Christiano, and Fisher (2001) produces a negative announcement premium. The consumption substitutability model of Dunn and Singleton (1986) and Heaton (1993) is consistent with a positive announcement premium, although this feature of the utility function smooths the marginal utility process and has difficulty in accounting for many aspects of the asset market data, as highlighted in Gallant, Hansen, and Tauchen (1990).

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 (2013). 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 (1989); its dynamic version studied by Chen and Epstein (2002) and Epstein and Schneider (2003); the recursive preference of Kreps and Porteus (1978) and Epstein and Zin (1989); the robust control preference of Hansen and Sargent (2005, 2007) and the related multiplier preference of Strzalecki (2011); the variational ambiguity-averse preference of Maccheroni, Marinacci, and Rustichini (2006a,b); the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2005, 2009); the disappointment aversion preference of Gul (1991); and the recursive smooth ambiguity preference of Hayashi and Miao (2011). We also discuss the relationship between our notion of generalized risk sensitivity and the related decision theoretic concepts, such as 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 (2010) for a review of asset pricing studies with the maxmin expected utility model; Ju and Miao (2012) for an application of the smooth ambiguity-averse preference; Hansen and Sargent (2008) for the robust control preference; Routledge and Zin (2010) for an asset pricing model with disappointment aversion; and Bansal and Yaron (2004), Bansal (2007), and Hansen, Heaton, and Li (2008) for the long-run risks model that builds on recursive preferences. Skiadas (2009) provides an excellent textbook treatment of recursive-preferences in asset pricing theory.

Unlike the calibration methodology used in the above papers, our paper takes a revealed preference approach. Earlier work on the revealed preference approach for expected utility includes Green and Srivastava (1986) and Epstein (2000). More recently, Kubler, Selden, and Wei (2014) and Echenique and Saito (2015) develop asset-market-based characterizations of the expected utility model. None of the above papers focus on the macroeconomic announcement premium and generalized risk sensitivity.

Quantitatively, our findings are consistent with the literature that identifies large variations in marginal utilities from the asset market data (see, e.g., Hansen and Jagannathan (1991), Bansal and Lehmann (1997), and Alvarez and Jermann (2004, 2005)). Our theory implies that most of the variations in marginal utility must come from generalized risk sensitivity and not from risk aversion of the Von Neumann–Morgenstern utility function. This observation likely has sharp implications for the research on macroeconomic policies. Several recent papers study optimal policy design problems in non-expected utility models. For example, Farhi and Werning (2008) and Karantounias (2015) analyze optimal fiscal policies with recursive preferences, and Woodford (2010), Karantounias (2013), Hansen and Sargent (2012), and Kwon and Miao (2013b,a) focus on preferences that are averse to model uncertainty. In the above studies, the non-linearity in agents' certainty equivalent 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. This 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 (2013)) and internationally (Brusa, Savor, and Wilson

(2015)). Lucca and Moench (2015) find similar patterns and document a pre-FOMC announcement drift. Mueller, Tahbaz-Salehi, and Vedolin (2017) document an FOMC announcement premium on the foreign exchange market and attribute it to compensation to financially constrained intermediaries.

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 announcement days. Section 6 concludes.

2 Stylized facts

To demonstrate the significance of the equity premium for macroeconomic 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 are released at monthly or lower frequencies. Within this category, we select the top five announcements ranked by investor attention by Bloomberg users. This procedure yields, on average, thirty announcement days per year for the period of 1961-2014. We summarize our main findings below and provide details about the data construction in Appendix A.

(i) A large fraction of the market equity premium is realized on a relatively small number of trading days with pre-scheduled macroeconomic announcements.

In Table I, we report the average market excess returns on macroeconomic announcement days and non-announcement days during the 1961-2014 period. In this period, on average, thirty trading days per year have significant macroeconomic announcements. At the daily level, the average stock market excess return is 11.21 bps on announcement days and 1.27 bps on days without major macroeconomic announcements. As a result, the cumulative stock market excess return on the thirty announcement days averages 3.36% per year, accounting for about 55% of the annual equity premium (6.19%) during this period.

In Table II, we report the average market excess return on announcement days (0) and the same moments for the market return on the day before (-1) and the day after (+1) announcement days. The difference in mean returns between announcement days and non-announcement days is statistically and economically significant with a t-statistic of 3.36. This evidence is consistent with the previous literature (see e.g., Savor and Wilson (2013)).

(ii) Most of the premiums for FOMC announcements are realized in several hours prior to the announcements. Premiums for other macroeconomic announcements are realized upon the release of these announcements.

In Table III, we report the point estimates with standard errors for average hourly excess returns around announcements. We normalize the announcement time as hour zero. For $k = -5, -4, \cdots, 0, +1, +2$, the announcement window k in the table is defined as hour $k - 1$ to hour k. The hourly returns typically peak at the announcement, as reflected in row 1 of the table. The mean return during the announcement hour is economically important: 6.46 bps with a standard error of 2.71. The difference in mean excess returns in announcement hours compared to non-announcement hours, like in the daily returns data, is significant with a t-statistic of 2.06. In the case of FOMC announcements, consistent with Lucca and Moench (2015), the mean returns prior to the announcement window are statistically significant (see row 2 of Table III); this pre-announcement drift is not reflected in other macroeconomic announcements, as shown in row 3 of Table III. In Figure 1, we plot the average hourly stock market excess returns for FOMC announcements (top panel) and those for other macroeconomic announcements (bottom panel) in the hours around the announcements. There is a "pre-announcement drift" for FOMC announcements, but not for other macroeconomic announcements. The premiums for non-FOMC announcements are mainly realized at the announcement.[2]

In addition, Lucca and Moench (2015) document that there is no statistically significant pre-FOMC announcement drift for Treasury bonds in the 1994-2011 period, and Savor and Wilson (2013) present evidence of a moderate level of announcement premiums for Treasury bonds, which averages about 3 bps on announcement days during the longer sample period of 1961-2009.

[2] The evidence reported in Table III is robust to using 30-minute windows as opposed to hourly windows.

Figure 1. **Hourly return around announcements**

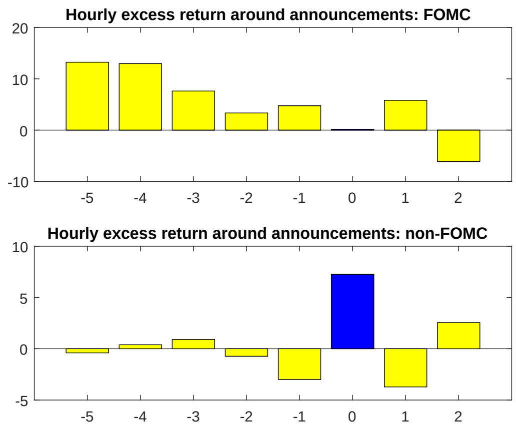


Figure 1 plots the average hourly excess returns around macro announcements for the period of 1997-2013. The top panel is for FOMC announcements and the bottom panel includes all other macro announcements. The horizontal axis marks announcement windows, and the vertical axis is the average hourly excess return for the announcement windows, measured in basis points. We normalize the announcement time to hour zero. For $k = -5,\ -4, \cdots, 0,\ 1,\ 2$, announcement window k is defined as the interval between hour $k-1$ and hour k.

3 Intuition from a two-period setup

In this section, we use a two-period setup to illustrate intuitively the conditions under which resolutions of uncertainty are associated with realizations of the equity premium and to motivate the key ingredients in the fully dynamic model, which we formally develop in Section 4.

3.1 Asset market for announcements

We consider a representative-agent economy with two periods, 0 and 1. Period 0 has no uncertainty and the aggregate consumption is a known constant, C_0. The aggregate consumption in period 1, denoted by 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 $\{C_1(n)\}_{n=1,2,\cdots N}$ and the possible realizations of asset payoff as $\{X(n)\}_{n=1,2,\cdots N}$. 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 C_1 is revealed, the pre-announcement market opens and asset prices at this point are called pre-announcement prices and are denoted by P^-. P^- cannot depend on the realization of C_1, which is unknown at this point. In period 0^+, the agent receives an announcement s that carries information about C_1. Immediately after the announcement, the post-announcement asset market opens. The post-announcement asset prices depend on s and are denoted by $P^+(s)$. In period 0^+, prices are denominated in current date-and-state-contingent consumption units, and the agent makes both optimal consumption and investment decisions given prices. In period 0^-, there is only investment decisions but no consumption decision. We denominate asset prices at 0^- in units of consumption goods delivered non-contingently in period 0^+.

For simplicity, we assume that announcements fully reveal the true state, that is, $s \in \{1, 2, \cdots, N\}$, although this assumption is not necessary in the fully dynamic model we develop in Section 4. In addition, we assume complete markets and differentiability of utility functions, so that Arrow-Debreu prices can be computed from marginal rates of substitution. In Figure 2, we illustrate the timing of information and consumption (top panel) and that of asset prices (bottom panel), assuming $N = 2$.[3]

The announcement return of an asset, denoted by $R_A(s)$, is defined as the return of a strategy that buys the asset before the pre-scheduled announcement and sells immediately afterwards (assuming no dividend payment at 0^+):

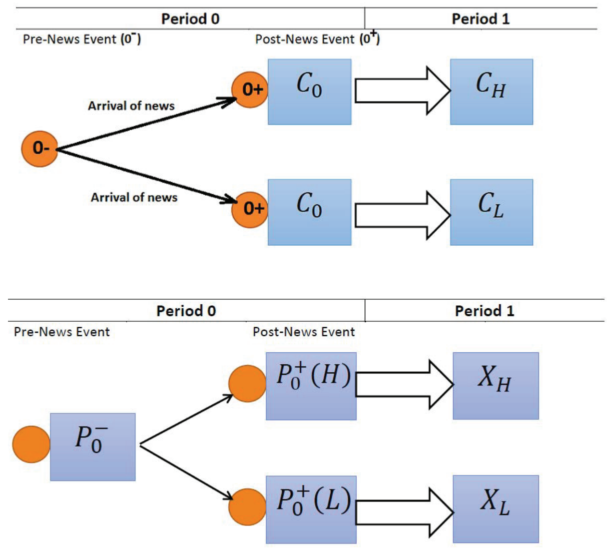
$$R_A(s) = \frac{P^+(s)}{P^-}. \tag{1}$$

The risk-free announcement return is the announcement return on an asset that delivers one unit of state-non-contingent consumption in period 0^+. Because of our choice of consumption numeraire, the risk-free announcement return must be one by no arbitrage. We say that an asset requires a positive announcement premium if $E[R_A(s)] > 1$. We also define the post-announcement return conditioning on announcement s as $R_P(X|\, s) = \frac{X(s)}{P^+(s)}$.

The assumption of the absence of a consumption decision at time 0^- and our choice of the consumption numeraire guarantees a zero risk-free announcement return and simplify our analysis. Allowing consumption at 0^- implies that the risk-free announcement return is in general different from one but does not affect our analysis of the announcement premium

[3]We provide details of the Arrow-Debreu market setup Section S. 1 of the Supplemental Material and formally establish that the Arrow-Debreu setup leads to the same asset pricing equations as the sequential market setup, which is a more convenient modeling choice for the fully dynamic model in Section 4.

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



as long as consumption at 0^+ does not depend on the content of the announcements. The key element of our assumption is that the arrival of announcements is not associated with a resolution of uncertainty about current-period consumption, but with that of future consumption.

We note two important properties of announcements in our model. First, announcements affect the conditional distribution of future consumption, but rational expectations imply that surprises in announcements must average to zero by the law of iterated expectation. Second, as mentioned above, we make the simplifying assumption that consumption does not instantaneously respond to macroeconomic announcements. This captures the well-established empirical findings of Hall (1978) and Parker and Julliard (2005), among others, that contemporaneous consumption co-move very little with stock market returns. Additionally, Bansal and Shaliastovich (2011) document that even large movements in stock prices are not associated with any significant immediate adjustment in aggregate consumption. This lack of contemporaneous covariance of stock returns and consumption implies that the contribution of the consumption covariance with asset returns over very short intervals (daily and hourly), if any, is too small to affect the observed announcement premiums discussed in Section 2.

The assumption that consumption does not instantaneously respond to macroeconomic announcements is also well motivated from the perspective of production-based models, where consumption is endogenous. In standard production-based models, the response of consumption to news is generally quantitatively small and yields a negative announcement premium. The response of consumption is quantitatively small because risk-averse agents dislike large consumption adjustments over short intervals. Further, the announcement premium that results, if at all, from the immediate response of consumption to news is, in general, negative. In reality, it is difficult to instantaneously adjust aggregate output upon announcements. Beaudry and Portier (2004, 2014) show that if output cannot respond to news, consumption and Tobin's q (and therefore asset returns) typically move in opposite directions regardless of whether the income or the substitution effect dominates. As a result, the negative co-movement of consumption and Tobin's q contributes negatively to the announcement premium. Our assumption, therefore, allows us to focus on the properties of preferences that generates a positive announcement premium.

3.2 Simple examples

Expected utility We first consider the case in which the representative agent has expected utility:[4] $E\left[u\left(C_0\left(s\right)\right)+\beta u\left(C_1\left(s\right)\right)\right]$, where u is strictly increasing and continuously differentiable.[5] The period 0^- price of one unit of period 1 consumption goods, which is measured in units of period 0^+ state-non-contingent consumption goods, can be computed from the ratio of marginal utilities: $\pi\left(s\right)\frac{\beta u'\left(C_1\left(s\right)\right)}{u'\left(C_0\right)}$.[6] Therefore, the pre-announcement price of an asset with payoff $\left\{X\left(s\right)\right\}_{s=1}^{N}$ is given by:

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

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

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

[4]We use the term "expected utility" to mean utility functions that are additively separable with respect to both time and states.

[5]Because the decision for C_0 is made at 0^+ after the announcement is made, from the agent's point of view, $C_0\left(s\right)$ is allowed to depend on s.

[6]From the agent's perspective, the marginal utility of one unit of period 0^+ state non-contingent consumption is $E\left[u'\left(C_0\left(s\right)\right)\right]$. In equilibrium, the market clearing condition implies that $C_0\left(s\right)$ cannot depend on s. Therefore, the expectation sign is not necessary: $E\left[u'\left(C_0\left(s\right)\right)\right]=u'\left(C_0\right)$. In the rest of this section, we will use the notation $C_0\left(s\right)$ when describing preference to emphasize that individual agent's consumption choice is allowed to depend on s. In the expressions of stochastic discount factors, we will impose market clearing and write C_0. Please see Section S. 1 in the Supplemental Material for a detailed derivation.

As a result, 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.

Robust control Consider an agent with the constraint robust control preference of Hansen and Sargent (2001):

$$\min_{\{m(s)\}_{s=1}^{n}} E[m(s)\{u(C_0(s)) + \beta u(C_1(s))\}], \tag{5}$$
$$\textit{subject to}: \; E[m(s)\ln m(s)] \leq \eta,$$
$$E[m(s)] = 1.$$

The above expression can also be interpreted as the maxmin expected utility of Gilboa and Schmeidler (1989). The agent treats the reference probability measure, under which the equity premium is evaluated (by econometricians), as an approximation. As a result, the agent takes into account of a class of alternative probability measures, represented by the density m, close to the reference probability measure, and evaluates utility using the worst-case probability. The inequality $E[m\ln m] \leq \eta$ requires that the relative entropy of the alternative probability models is less than η.

In this case, the pre-announcement price of an asset with payoff $\{X(s)\}_{s=1}^{N}$ is:

$$P^{-} = E\left[m^{*}(s)\frac{\beta u'(C_1(s))}{u'(C_0)} X(s)\right], \tag{6}$$

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

$$m^{*}(s) = \frac{e^{-\frac{u(C_1(s))}{\theta}}}{E\left[e^{-\frac{u(C_1(s))}{\theta}}\right]}. \tag{7}$$

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

In period 0^{+}, after the resolution of uncertainty, the agent's utility reduces to (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:

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

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

Claim 1. *Consider post-announcement prices that are co-monotone with $C_1(s)$, that is, $\forall s$ and s', $C_1(s) \geq C_1(s')$ if and only if $P(s) \geq P(s')$.*[7] *Equation (8) implies $P^{-} \leq E[P^{+}(s)]$. As a result, the announcement premium is non-negative.*

The intuition of the above result is clear. Because uncertainty is resolved after the announcement, asset prices are discounted using marginal utilities. Under the 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 expectations. 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 extra discounting to 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 call it the announcement stochastic discount factor (SDF), or A-SDF, to distinguish it from the standard SDF derived from agents' marginal rate of intertemporal substitution of consumption. In our model, there is no intertemporal consumption decision before the announcement at 0^-. The term m^* reflects investors' uncertainty aversion and identifies the probability distortion relative to rational expectation.

Recursive utility The last example we discuss here is the recursive utility of Kreps and Porteus (1978) and Epstein and Zin (1989) with constant elasticity of substitution (CES). Because all uncertainties are fully resolved after the announcement, in period 0^+, the agent first aggregates utility across time to compute continuation utility given announcement s:

$$\frac{1}{1-\frac{1}{\psi}} C_0^{1-\frac{1}{\psi}}(s) + \beta \frac{1}{1-\frac{1}{\psi}} C_1^{1-\frac{1}{\psi}}(s),$$

[7] See also equation (21) for the definition of co-monotonicity.

where ψ is the intertemporal elasticity of substitution parameter. Before the announcement, in period 0^-, the agent computes the certainty equivalent of the continuation utility:[8]

$$\left\{E\left[\left\{C_0^{1-\frac{1}{\psi}}(s)+\beta C_1^{1-\frac{1}{\psi}}(s)\right\}^{\frac{1-\gamma}{1-1/\psi}}\right]\right\}^{\frac{1}{1-\gamma}}. \tag{9}$$

Again, the period 0^- Arrow-Debreu price of one unit of period 1 consumption goods can be computed from the ratio of marginal utilities: $m^*(s)\,\beta\left[\frac{C_1(s)}{C_0}\right]^{-\frac{1}{\psi}}$, where

$$m^*(s)=\frac{\left\{C_0^{1-\frac{1}{\psi}}+\beta C_1^{1-\frac{1}{\psi}}(s)\right\}^{\frac{1/\psi-\gamma}{1-1/\psi}}}{E\left[\left\{C_0^{1-\frac{1}{\psi}}+\beta C_1^{1-\frac{1}{\psi}}(s)\right\}^{\frac{1/\psi-\gamma}{1-1/\psi}}\right]}. \tag{10}$$

can be interpreted as A-SDF as in the case of the robust control preference. Clearly, m^* is a decreasing function of continuation utility if and only if $\gamma > \frac{1}{\psi}$, which coincides with the condition for preference for early resolution of uncertainty for this class of preferences.[9]

3.3 A-SDF for general preferences

In this section, we provide an intuitive discussion of the A-SDF for general preferences in the two-period setup. Because there is no uncertainty after the announcement at time 0^+, we assume that the agent ranks consumption streams according to a time-separable utility function, and we denote the continuation utility conditioning upon announcement s by $V_s = u(C_0(s)) + \beta u(C_1(s))$. At time 0^-, prior to the announcement, the agent ranks uncertain outcomes according to a general certainty equivalent functional $\mathcal{I}[V]$, where $\mathcal{I}$ maps random variables into the real line.[10] Because there are N states of the world, we use the vector notation $V = [V_1, V_2, \cdots, V_N]$ and think of $\mathcal{I}$ as a mapping from the N-dimensional Euclidean space to the real line.

[8] Here, we choose a convenient normalization of the recursive utility so that it fits the general representation assumed in the theorems in Section 4. See also Section S. 1.3 in the Supplemental Material.

[9] Note that the announcement leads uncertainty about C_1 to resolve before its realization, which corresponds to the case of early resolution of uncertainty in Kreps and Porteus (1978). See Section S. 1.3 in the Supplemental Material for a comparison between SDF computed from consumption plans with early resolution of uncertainty and that with late resolution of uncertainty, respectively.

[10] We follow Strzalecki (2013) and call $\mathcal{I}$ the certainty equivalent functional. However, we note that $\mathcal{I}[V]$ is measured in utility terms, not in consumption terms.

To compute Arrow-Deberu prices, note that from the perspective of period 0^-, the marginal utility of one unit of state-non-contingent consumption delivered in period 0^+ is $\sum_{s=1}^{N} \frac{\partial}{\partial V_s}\mathcal{I}[V] \cdot u'(C_0)$. The marginal utility of one unit of period 1 consumption good in state s is $\frac{\partial}{\partial V_s}\mathcal{I}[V] \cdot \beta u'(C_1(s))$. The pre-announcement price of an asset can therefore be computed as the marginal utility weighted payoffs:

$$P^- = \sum_{s=1}^{N} \frac{\frac{\partial}{\partial V_s}\mathcal{I}[V] \cdot \beta u'(C_1(s))}{\sum_{s=1}^{N} \frac{\partial}{\partial V_s}\mathcal{I}[V] \cdot u'(C_0)} X(s) = E\left[m^*(s)\,\beta \frac{u'(C_1(s))}{u'(C_0)} X(s)\right], \tag{11}$$

where

$$m^*(s) = \frac{1}{\pi(s)} \frac{\frac{\partial}{\partial V_s}\mathcal{I}[V]}{\sum_{s=1}^{N} \frac{\partial}{\partial V_s}\mathcal{I}[V]}. \tag{12}$$

Clearly, the asset pricing equation (8) holds with the A-SDF m^* defined by (12).

If we focus on assets with pro-cyclical payoffs, in the sense that they are increasing functions of the representative agent's continuation utility, V_s, then, as we show in Claim 1, a sufficient condition for a non-negative announcement premium is that $m(s)$, and, equivalently, $\frac{\partial}{\partial V_s}\mathcal{I}[V]$ is co-monotone with respect to V_s. That is, for all s and s',

$$\frac{\partial}{\partial V_s}\mathcal{I}[V] \geq \frac{\partial}{\partial V_{s'}}\mathcal{I}[V] \text{ if and only if } V_s \leq V_{s'}. \tag{13}$$

Condition (13) is known to be a characterization of Schur concavity. Under the assumption that all states occur with equal probabilities, that is, $\pi(s) = \frac{1}{N}$ for $s = 1, 2, \cdots, N$, the above property is equivalent to monotonicity with respect to second-order stochastic dominance (see, e.g., Marshall, Arnold, and Olkin (2011) and Muller and Stoyan (2002)). This is key insight of our paper: non-negative announcement premiums for payoffs that are co-monotone with respect to continuation utility are equivalent to the certainty equivalent functional being increasing in second-order stochastic dominance.

In the following section, we formally develop the above results in a fully dynamic model with a continuous probability space, which allows us to dispense with the assumptions of fully revealing announcements and finite states with equal probabilities.

4 Risk preferences and the announcement premium

4.1 A dynamic model with announcements

Preferences The setup of our model follows that of Strzalecki (2013), but we extend his framework to allow for announcements. Let S be a non-atomic measurable space, and Σ the associated Borel σ-field. Let $(\Omega, \mathcal{F}) = (S, \Sigma)^{2T}$ be the product space. We index the $2T$ copies of (S, Σ) by $j = 0^+, 1^-, 1^+, 2^-, \cdots, T-1^+, T^-$ with the interpretation that t^- is the pre-announcement period at time t and t^+ is the post-announcement period at time t. A typical element in Ω is therefore denoted by $\omega = \left\{s_0^+, s_1^-, s_1^+, s_2^-, \cdots, s_{T-1}^+, s_T^-\right\}$. Let $z_{t-1}^+ = \left\{s_0^+, s_1^-, s_1^+, s_2^-, \cdots, s_{t-1}^-, s_{t-1}^+\right\}$ and $z_t^- = \left\{s_0^+, s_1^-, s_1^+, s_2^-, \cdots, s_{t-1}^+, s_t^-\right\}$ denote the history of the realizations until $t-1^+$ and until t^-, respectively. Let $\mathcal{F}_{t-1}^+ = \sigma\left(z_{t-1}^+\right)$, and $\mathcal{F}_t^- = \sigma\left(z_t^-\right)$ be the σ-fields generated by the history of realizations, for $t = 1, 2, \cdots T$. The filtration $\left\{\mathcal{F}_{t-1}^+, \mathcal{F}_t^-\right\}_{t=1}^T$ represents public information. We use $\mathbf{Z}$ to denote the set of all histories, and let $z \in \mathbf{Z}$ denote a generic element of $\mathbf{Z}$.

We endow the measurable space $(\Omega, \mathcal{F})$ with a non-atomic probability measure P, under which the distribution of $\left\{s_0^+, s_1^-, s_1^+, s_2^-, \cdots, s_{T-1}^+, s_T^-\right\}$ is stationary. The interpretation is that P is the probability measure under which all expected returns are calculated. We assume that consumption takes value in $\mathbf{Y}$, an open subset of $\mathbf{R}$, endowed with the Borel σ-field $\mathcal{B}$. Let $L^2(\Omega, \mathcal{F}, P)$ be the Hilbert space of square-integrable real-valued random variables defined on $(\Omega, \mathcal{F}, P)$. A consumption plan is an $\left\{\mathcal{F}_t^+\right\}_{t=1}^T$-adapted process $\{C_t\}_{t=1}^T$, such that C_t is a $\mathbf{Y}$-valued square-integrable random variables for all t. $\mathcal{C}$ denotes the space of all such consumption plans, and a typical element in $\mathcal{C}$ is denoted by $\mathbf{C} \in \mathcal{C}$.

The aggregate endowment of the economy, denote as $\bar{\mathbf{C}} \in \mathcal{C}$ is required to be $\left\{\mathcal{F}_t^-\right\}_{t=1}^T$-adapted. As in the two-period model, individual consumption choices are allowed to be made contingent on the announcements, $\left\{s_{t-1}^+\right\}_{t=1}^T$. However, announcements carry information about future endowments but do not affect current-period endowments. That is, $\forall t$, the aggregate consumption $\bar{C}_t$, must be $\mathcal{F}_t^-$ measurable. The above setup allows us to model announcements as revelations of public information associated with realizations of $\left\{s_{t-1}^+\right\}_{t=1}^T$, separately from the realizations of consumption. As a result, our theory is able to separate the property of preferences that requires premiums for assets with a payoff correlated with resolutions of uncertainty from the property of preferences that demands excess returns for assets that co-move with realizations of consumption.

Strzalecki (2013) shows that most of the non-expected utility models can be represented as

$$V_t = u(C_t) + \beta \mathcal{I}[V_{t+1}]. \tag{14}$$

Below, we adapt representation (14) to allow for announcements and describe a recursive procedure to construct a system of conditional preferences, $\{\succsim_z\}_{z \in \mathbf{Z}}$ on $\mathcal{C}$, such that $\mathbf{C} \succsim_z \mathbf{C}'$ if $V_z(\mathbf{C}) \geq V_z(\mathbf{C}')$. Formally, the representative agent's dynamic preference is defined by a triple $\{u, \beta, \mathcal{I}\}$, where $u : \boldsymbol{Y} \to \mathbf{R}$ is a strictly increasing Von Neumann–Morgenstern utility function, and $\beta \in (0, 1]$ is the subjective discount rate. The certainty equivalent functional $\mathcal{I}$ is a family of functions, $\{\mathcal{I}[\cdot|\, z]\}_{z \in \mathbf{Z}}$, such that $\forall z \in \mathbf{Z}$, $\mathcal{I}[\cdot|\, z] : L^2(\Omega, \mathcal{F}, P) \to \mathbf{R}$ is a (conditional) certainty equivalent functional that maps continuation utilities into the real line. Given $\{u, \beta, \mathcal{I}\}$, the agent's utility function is constructed recursively as follows.

- At the terminal time T, $V_{z_T^-}(\mathbf{C}) = u(C_T)$.
- For $t = 0, 1, 2, \cdots T-1$, given $V_{z_{t+1}^-}(\mathbf{C})$, in period t after the signal s_t is revealed, $V_{z_t^+}(\mathbf{C})$ is calculated according to:

$$V_{z_t^+}(\mathbf{C}) = u(C_t) + \beta \mathcal{I}\left[V_{z_{t+1}^-}(\mathbf{C}) \middle|\, z_t^+ \right]. \tag{15}$$

- For $t = 1, 2, 3, \cdots T-1$, given a continuation utility $V_{z_t^+}(\mathbf{C})$, in period t before the signal s_t is received, $V_{z_t^-}(\mathbf{C})$ is defined as

$$V_{z_t^-}(\mathbf{C}) = \mathcal{I}\left[V_{z_t^+}(\mathbf{C}) \middle|\, z_t^- \right]. \tag{16}$$

  Here, there is no consumption decision at z_t^- before the signal s_t is received, and we simply use the certainty equivalent functional $\mathcal{I}$ to aggregate utility across states.

In Section S. 2 of the Supplemental Material, we show that the above representation incorporates the following dynamic preferences under uncertainty and provide expressions for the A-SDF implied by these preferences:

(i) The recursive utility of Kreps and Porteus (1978) and Epstein and Zin (1989).

(ii) The maxmin expected utility of Gilboa and Schmeidler (1989). The dynamic version of this preference is studied in Epstein and Schneider (2003) and Chen and Epstein (2002).

(iii) The variational preferences of Maccheroni, Marinacci, and Rustichini (2006a), the dynamic version of which is studied in Maccheroni, Marinacci, and Rustichini (2006b).

(iv) The multiplier preferences of Hansen and Sargent (2008) and Strzalecki (2011).

(v) The second-order expected utility of Ergin and Gul (2009).

(vi) The smooth ambiguity preferences of Klibanoff, Marinacci, and Mukerji (2005), and Klibanoff, Marinacci, and Mukerji (2009).

(vii) The disappointment aversion preference of Gul (1991).

(viii) The recursive smooth ambiguity preference of Hayashi and Miao (2011) can also be represented as (14) with some additional assumptions on the intertemporal aggregator. We discuss the A-SDF for this class of preferences in Section S. 2 of the Supplemental Material.

Asset markets Because preferences are defined recursively, it is more convenient to model the asset market as one with sequential trading. We assume that asset markets open after each history $z \in \mathbf{Z}$. We interpret the realizations of $\left\{s_t^+\right\}_{t=1}^T$ as announcements, because they carry information about future consumption but are not associated with realizations of current-period consumption. Markets at period t^- are called pre-announcement markets. Here, agents can trade a vector of $J+1$ returns: $\left\{R_{A,j}\left(\cdot\middle|\, z_t^-\right)\right\}_{j=0,1,\cdots,J}$, where $R_{A,j}\left(\cdot\middle|\, z_t^-\right)$ represents an announcement-contingent return that is traded at history z_t^- and that pays off $R_{A,j}\left(s_t^+\middle|\, z_t^-\right)$ at all subsequent histories $\left\{\left(z_t^-, s_t^+\right)\right\}_{s_t^+}$ (The notation $\left\{\left(z_t^-, s_t^+\right)\right\}_{s_t^+}$ denotes the vector of histories for a fixed z_t^- and for all possible realizations of s_t^+ after z_t^-). Similarly, agents can trade a vector of post-announcement returns, $\left\{R_{P,j}\left(\cdot\middle|\, z_t^+\right)\right\}_{j=0,1,\cdots,J}$ on the post-announcement market at history z_t^+. In general, we use $R_j\left(\cdot|\, z\right)$ to denote the return on asset j traded at history $z \in \mathbf{Z}$ with the understanding that it is an announcement return if z is of the form z_t^- and a post-announcement return if z is of the form z_t^+. We adopt the convention that $R_0\left(\cdot|\, z\right)$ is the risk-free return at history $z \in \mathbf{Z}$, and we write it as $R_0\left(z\right)$ whenever convenient.

Given the recursive nature of the preferences, the optimal consumption-portfolio choice problem of the agent can be solved by backward induction. For any $z \in \mathbf{Z}$, we use $V_z\left(W\right)$ to denote the agent's continuation utility as a function of wealth at history z, and call them value functions. We denote $\xi = [\xi_0, \xi_1, \xi_2, \cdots \xi_J]$ as the vector of investment in the $J+1$ securities on the post-announcement asset market. In the last period T, agents simply consume their total wealth, and therefore $V_{z_T^-}\left(W\right) = u\left(W\right)$. For $t = 1, 2, \cdots, T-1$, given $V_{z_{t+1}^-}\left(W\right)$, the value function at the history z_t^+ that precedes z_{t+1}^- can be constructed by

$$V_{z_t^+}(W) = \max_{C,\xi} \left\{ u(C) + \beta \mathcal{I}\left[V_{z_{t+1}^-}(W') \middle| z_t^+ \right] \right\} \tag{17}$$
$$C + \sum_{j=0}^{J} \xi_j = W$$
$$W' = \sum_{j=0}^{J} \xi_j R_{P,j}\left(s_{t+1}^- \middle| z_t^+ \right), \quad all\ s_{t+1}^-.$$

Similarly, given the post-announcement value function, $V_{z_t^+}(W)$, the optimal portfolio choice problem on the pre-announcement market is

$$V_{z_t^-}(W) = \max_{\zeta} \mathcal{I}\left[V_{z_t^+}(W') \middle| z_t^- \right] \tag{18}$$
$$W' = W - \sum_{j=0}^{J} \zeta_j + \sum_{j=0}^{J} \zeta_j R_{A,j}\left(s_t^+ \middle| z_t^- \right), \quad all\ s_t^+,$$

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

Like in our two-period model, asset prices and wealth are measured in current-period consumption goods on the post-announcement market (see equation (17)). On the pre-announcement market, the agent makes portfolio allocation decisions, but not intertemporal consumption choices. Prices on the pre-announcement market at history z_t^- are denominated in units of state-non-contingent consumption goods delivered at history $\left(z_t^-, \cdot\right)$: as shown in (18), one unit of wealth at history z_t^-, if not invested, becomes one unit of wealth at $\left(z_t^-, s_t^+\right)$ for all s_t^+. Our convention implies that the return on the risk-free asset that pays one unit of consumption goods non-contingently upon announcement must be one by no arbitrage: $R_0\left(z_t^-\right) = 1$ for all z_t^-.

We assume that for some initial wealth level, W_0 and a sequence of returns $\left\{ \left\{ R_j\left(\cdot \middle| z \right) \right\}_{j=0,1,\cdots,J} \right\}_{z \in \mathbf{Z}}$, an interior competitive equilibrium with sequential trading exists in which all markets clear. For simplicity, we start with returns directly in our description of the equilibrium with the understanding that returns can always be constructed from prices. Below, we provide a formal definition of the announcement premium.

Definition 1. *Announcement premium:*

The *announcement premium for asset j at history z_t^-* is defined as

$$E\left[R_{A,j}\left(\cdot \middle| z_t^- \right) \middle| z_t^- \right] - 1.$$

4.2 The announcement SDF

To relate the announcement premium to the properties of the certainty equivalent functional, $\mathcal{I}[\cdot]$, we first provide some definitions. The certainty equivalent functional $\mathcal{I}$ is said to be monotone with respect to first-order (second-order) stochastic dominance if $X_1 \geq_{FSD} X_2$ ($X_1 \geq_{SSD} X_2$) implies that $\forall z \in \mathbf{Z}$, $\mathcal{I}[X_1|z] \geq \mathcal{I}[X_2|z]$. It is strictly monotone with respect to first-order (second-order) stochastic dominance if $X_1 >_{FSD} X_2$ ($X_1 >_{SSD} X_2$) implies that $\forall z \in \mathbf{Z}$, $\mathcal{I}[X_1|z] > \mathcal{I}[X_2|z]$, where $\geq_{FSD}$ and $\geq_{SSD}$ stand for first- and second-order stochastic dominance, respectively.[11] In what follows, we assume that $\mathcal{I}$ is normalized; that is, $\mathcal{I}[X|z] = X$ a.s. whenever X is a measurable function of z.

Conceptually, the property of asset prices imposes restrictions on the derivatives of utility functions. Our theoretical exercise amounts to recovering the property of utility functions from their derivatives and is related to the "local utility" analysis in Machina (1982), Wang (1993), and Ai (2005). In our setup, the certainty equivalent functional is a mapping from $L^2(\Omega, \mathcal{F}, P)$ into the real line. We therefore need a notion of differentiability in infinite dimensional spaces. We use $\|\cdot\|$ for the L^2 norm on $L^2(\Omega, \mathcal{F}, P)$, and $|\cdot|$ for absolute value, and we introduce the concept of Fréchet differentiability as follows.

Definition 2. *Fréchet Differentiability with Lipschitz Derivatives:*

The certainty equivalent functional $\mathcal{I}$ is *Fréchet differentiable* if $\forall z \in \mathbf{Z}$, $\forall X \in L^2(\Omega, \mathcal{F}, P)$, there exists a unique continuous linear functional, $D\mathcal{I}[X|z] \in L^2(\Omega, \mathcal{F}, P)$ such that for all $\Delta X \in L^2(\Omega, \mathcal{F}, P)$,

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

A Fréchet differentiable certainty equivalent functional $\mathcal{I}$ is said to have *Lipschitz derivatives* if $\forall X, Y \in L^2(\Omega, \mathcal{F}, P)$, $\forall z \in \mathbf{Z}$, $\|D\mathcal{I}[X|z] - D\mathcal{I}[Y|z]\| \leq K\|X - Y\|$ for some constant K.[12]

Given a (conditional) certainty equivalent functional $\mathcal{I}[\,|z]$ and $X \in L^2(\Omega, \mathcal{F}, P)$, the Fréchet derivative of $\mathcal{I}[\,|z]$ at X is a continuous linear functional on $L^2(\Omega, \mathcal{F}, P)$, which has a unique representation in $L^2(\Omega, \mathcal{F}, P)$ by the Riesz representation theorem. In what

[11] The definitions of first- and second-order stochastic dominance are standard and are provided in Appendix B.1.

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

follows, we denote $D\mathcal{I}\left[X|\, z\right]$ as (the Riesz representation of) the Fréchet derivative of $\mathcal{I}\left[\,|\, z\right]$ at X. To simplify notations, for any pre-announcement history z_t^-, and any announcement s_t^+ that follows z_t^-, we denote $V\left(s_t^+\middle|\, z_t^-\right) \equiv V_{z_t^+}\left(W_{z_t^+}\right)$, where $z_t^+ = \left(z_t^-, s_t^+\right)$, and $W_{z_t^+}$ is the equilibrium level of wealth of the representative agent at history z_t^+. That is, $V\left(s_t^+\middle|\, z_t^-\right)$ is the representative agent's equilibrium continuation utility at announcement s_t^+ following history z_t^-. The following theorem provides an existence result for the A-SDF.

Theorem 1. *(Existence of an A-SDF)* [13]

Suppose both u and $\mathcal{I}$ are Lipschitz continuous, Fréchet differentiable with Lipschitz derivatives. Suppose that u has strictly positive first-order derivatives on its domain and $\mathcal{I}$ is strictly monotone with respect to first-order stochastic dominance, then in any interior competitive equilibrium with sequential trading, $\forall z_t^- \in \mathbf{Z}$, there exists a non-negative measurable function $m^* : \mathbf{R} \to \mathbf{R}^+$ such that

$$E\left[m^*\left(V\left(\cdot|\, z_t^-\right)\right)\left\{R_{A,j}\left(\cdot|\, z_t^-\right) - 1\right\}\middle|\, z_t^-\right] = 0 \quad for \quad all \quad j = 0, 1, 2, \cdots J. \tag{19}$$

Under the regularity condition (54) in Appendix B.2, $E\left[m^*\left(V\left(\cdot|\, z_t^-\right)\right)\middle|\, z_t^-\right] = 1$ and (19) can be written as:

$$E\left[m^*\left(V\left(\cdot|\, z_t^-\right)\right)R_{A,j}\left(\cdot|\, z_t^-\right)\middle|\, z_t^-\right] = 1 \quad for \quad all \quad j = 0, 1, 2, \cdots J. \tag{20}$$

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. Let $\left\{f\left(s_t^+\middle|\, z_t^-\right)\right\}_{s_t^+}$ be an asset traded at history z_t^- with a payoff contingent on the announcement s_t^+. The payoff f is said to be co-monotone with continuation utility if

$$\left[f\left(s|\, z_t^-\right) - f\left(s'|\, z_t^-\right)\right]\left[V\left(s|\, z_t^-\right) - V\left(s'|\, z_t^-\right)\right] \geq 0 \text{ for all } s, s' \text{ almost surely.} \tag{21}$$

Intuitively, co-monotonicity captures the idea that the payoff f is non-decreasing in continuation utility V. The following theorem formalizes our discussion in Section 3.3 and provides necessary and sufficient conditions for the announcement premium.

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

[13] To avoid overly technical conditions, we assume the Fréchet differentiability of $\mathcal{I}$ in Theorems 1 and 2. The proofs in Appendix B remain valid under weaker conditions. In particular, we do not need the Fréchet derivative of $\mathcal{I}[V]$ to be unique, we only need the projection of the gradient, $D\mathcal{I}[V]$ onto $L^2(\Omega, \sigma(V), P)$ to be unique. This latter condition allows for multiple prior preferences and robust control preferences that do not satisfy Fréchet differentiability.

(i) The A-SDF $m^*(V) = 1$ for all V if and only if $\mathcal{I}$ is the expectation operator.

(ii) The following conditions are equivalent:

> *(a) The certainty equivalent functional $\mathcal{I}$ is monotone with respect to second-order stochastic dominance.*
>
> *(b) The A-SDF $m^*(V)$ is a non-increasing function of continuation utility V.*
>
> *(c) The announcement premium is non-negative for all payoffs that are co-monotone with continuation utility.*

Theorem 2 is our revealed preference characterization of the announcement premium. The presence of the announcement premium imposes restrictions on preferences because according to Theorem 1, the A-SDF that prices announcement returns is constructed from marginal utilities. Therefore, data on announcement returns impose restrictions on the marginal utilities of investors, and marginal utilities can be integrated to obtain the utility function itself. Like any revealed preference exercise, richer data allows more precise statements about preferences. Here, the assumption of non-atomic probability space is important, as it allows us to construct a rich enough set of test assets and to use the pricing information on these assets to infer the properties of investors' utility functions.

The key insight from Theorem 2 is that the announcement premium is informative about how agents aggregate continuation utilities to compute their certainty equivalent. From the examples in Section 3.2, there can be no announcement premium under expected utility. The first part of the above theorem implies that the converse of the statement is also true: if we have enough test assets and the announcement premiums for all test assets are zero, we can infer that the representative agent must be a time-separable expected utility maximizer.

The second part of the theorem provides a necessary and sufficient condition for non-negative announcement premiums. In particular, if the announcement premiums for all payoffs that are co-monotone with continuation utility are non-negative, then we can conclude that the certainty equivalent functional $\mathcal{I}$ must be monotone with respect to second-order stochastic dominance.

To conclude that $\mathcal{I}$ is increasing in second-order stochastic dominance, we only need the announcement premium to be non-negative for a relatively small class of assets, that is, assets that satisfy the co-monotonicity condition in (21). However, if we already know that $\mathcal{I}$ is increasing in second-order stochastic dominance, it is straightforward to show that the announcement premium must be non-negative for a much larger set of assets. In particular, any payoff of the form $f\left(s|\, z_t^-\right)+\varepsilon$, where $E\left[\varepsilon|\, z_t^-, s\right] = 0$ must require a non-negative announcement premium. This observation is useful in asset pricing applications in which payoffs may not be measurable functions of the continuation utility.[14]

4.3 Generalized risk-sensitive preferences

Theorem 2 motivates the following definition of generalized risk sensitivity.

Definition 3. *Generalized Risk Sensitivity:*

An intertemporal preference $\{u, \beta, \mathcal{I}\}$ is said to satisfy (strict) *generalized risk sensitivity*, if the certainty equivalent functional $\mathcal{I}$ is (strictly) monotone with respect to second-order stochastic dominance.

Under the assumptions of Theorem 1, generalized risk sensitive preferences are precisely the class of preferences that require a non-negative risk compensation for all assets with announcement payoffs co-monotone with investors' continuation utility.

Loosely speaking, generalized risk sensitivity is a "concavity" property of the certainty equivalent functional. The decision theory literature has studied related properties of the certainty equivalent functional, for example, uncertainty aversion (Gilboa and Schmeidler (1989)), and preference for early resolution of uncertainty (Kreps and Porteus (1978)). To clarify the notion of generalized risk sensitivity, in this section, we discuss its relationship with the above decision theoretic concepts. Throughout, we will assume that the intertemporal preference, $\{u, \beta, \mathcal{I}\}$ is normalized and satisfies the assumptions of Theorem 1. Also, we assume that either $u(\mathbf{Y}) = \mathbf{R}$ or $u(\mathbf{Y}) = \mathbf{R}^+$ like in Strzalecki (2013).

Generalized risk sensitivity and uncertainty aversion As in Strzalecki (2013), we define uncertainty aversion as the quasiconcavity of the certainty equivalent functional $\mathcal{I}$:

Definition 4. *Uncertainty Aversion:*

An intertemporal preference $\{u, \beta, \mathcal{I}\}$ is said to satisfy *uncertainty aversion*, if the certainty equivalent functional $\mathcal{I}$ is quasiconcave, that is, $\forall X_1, X_2 \in L^2(\Omega, \mathcal{F}, P)$, $\forall \lambda \in (0,1)$, $\mathcal{I}[\lambda X_1 + (1-\lambda) X_2] \geq \min\{\mathcal{I}[X_1], \mathcal{I}[X_2]\}$.

We make the following observations about the relationship between uncertainty aversion and generalized risk sensitivity. We provide formal proofs in Appendix C.1.

[14] We thank an anonymous referee for pointing this out.

(i) *The quasiconcavity of $\mathcal{I}$ is sufficient, but not necessary, for generalized risk sensitivity.*

A direct implication of the above result is that all uncertainty-averse preferences can be viewed as ways to formalize generalized risk sensitivity. Under the assumptions of Theorem 1, they all require 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 (1989); the second-order expected utility of Ergin and Gul (2009); the smooth ambiguity preference of Klibanoff, Marinacci, and Mukerji (2005); the variational preference of Maccheroni, Marinacci, and Rustichini (2006a); the multiplier preference of Hansen and Sargent (2008) and Strzalecki (2011); and the confidence preference of Chateauneuf and Faro (2009).

In Appendix C.1, we provide a proof for the sufficiency of quasiconcavity for generalized risk sensitivity. To illustrate that quasiconcavity is not necessary, in the same appendix, we also provide an example that satisfies generalized risk sensitivity, but not quasiconcavity.

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

The certainty equivalent functional of many intertemporal preferences takes the above form, for example, the the second-order expected utility of Ergin and Gul (2009) and the recursive preferences of Kreps and Porteus (1978) and Epstein and Zin (1989). For these preferences, generalized risk sensitivity is equivalent to the concavity of ϕ.

(iii) *Assume the ϕ function in the representation (22) below is continuous and strictly increasing. Within this class of smooth ambiguity-averse preferences, uncertainty aversion is equivalent to generalized risk sensitivity.*

The smooth ambiguity-averse preference of Klibanoff, Marinacci, and Mukerji (2005, 2009) can be represented in the form of (14) with the following choice of the certainty equivalent functional:

$$\mathcal{I}[V] = \phi^{-1}\left\{\int_{\Delta} \phi\left(E^{x}[V]\right) d\mu(x)\right\}. \tag{22}$$

Here, Δ denotes a set of probability measures indexed by x, denoted by P_x. The notation $E^x[\cdot]$ stands for expectations under the probability P_x, and $\mu(x)$ is a probability measure over Δ. In Appendix C, 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 Our definition of preference for early resolution directly follows from Strzalecki (2013). We first introduce a binary relation, $\geq_t$ on a subspace of $\mathcal{C}$ (see also definition 1 of Strzalecki (2013)). In the following definition, let $\check{C} : (S, \Sigma) \to (\mathbf{Y}, \mathcal{B})$ be a measurable function that specifies consumption as a function of the state $s \in S$. We also use $y_j \in \mathbf{Y}$ to denote a constant consumption plan that is measurable with respect to the trivial σ-field, $\{\emptyset, \Omega\}$.

Definition 5. *Early Resolution:*

Let $\mathbf{C}, \mathbf{C}' \in \mathcal{C}$, then $\mathbf{C} \geq_t^- \mathbf{C}'$ if there exists $\{y_j\}_{j \neq t+1}$ such that $C_j = C'_j = y_j$, for $j = 1, 2, \cdots, t, t+2, \cdots, T$, and $C_{t+1} = \check{C}\left(s_t^+\right)$, $C'_{t+1} = \check{C}\left(s_{t+1}^-\right)$.

Intuitively, $\mathbf{C}$ and $\mathbf{C}'$ are consumption plans that have no uncertainty other than in period $t+1$. Consumption in period $t+1$ have identical distributions under $\mathbf{C}$ and $\mathbf{C}'$, except that C_{t+1} depends on the realization of state s_t^+, whereas C'_{t+1} depends on s_{t+1}^-. In other words, under plan $\mathbf{C}'$, uncertainty in C'_{t+1} is not known until $t+1^-$, and under plan $\mathbf{C}$, uncertainty in C_{t+1} is known earlier, at t^+. Preference for early resolution of uncertainty is defined as follows.

Definition 6. *Preference for Early Resolution of Uncertainty:*

A system of conditional preferences $\{\succsim_z\}_{z \in \mathbf{Z}}$ is said to satisfy *preference for early resolution of uncertainty* if $\forall \mathbf{C}, \mathbf{C}' \in \mathcal{C}$, $\mathbf{C} \geq_t^- \mathbf{C}'$ implies $\mathbf{C} \succsim_{z_t^-} \mathbf{C}'$.

We summarize our main results for the relationship between preference for early resolution of uncertainty and generalized risk sensitivity as follows. The formal proofs for these statements can be found in Appendix C.2 of the paper.

(i) *Concavity of the certainty equivalent 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 (2013) also implies that these preferences satisfy preference for early resolution of uncertainty. As a result, Theorems 2 and 3 of Strzalecki (2013) imply that the variational preference of Maccheroni, Marinacci, and Rustichini (2006a) satisfies both generalized risk sensitivity and preference for early resolution of uncertainty.

(ii) *If $\mathcal{I}$ is of the form $\mathcal{I}[V] = \phi^{-1}\left(E\left[\phi\left(V\right)\right]\right)$ or it is the smooth ambiguity preference, $\mathcal{I}[V] = \int_\Delta \phi\left(E^x\left[V\right]\right) d\mu\left(x\right)$, 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:*

(i) $u(\mathbf{Y}) = \mathbf{R}$ and there exists $A \geq 0$ such that $-\frac{\phi''(a)}{\phi'(a)} \in [\beta A, A]$ for all $a \in \mathbf{R}$.

(ii) $u(\mathbf{Y}) = \mathbf{R}^{+}$ 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 (2013). 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 (2013), either of the above conditions implies preference for early resolution of uncertainty.

Because the CES recursive utility can be represented in the form of (14) with

$$u(C) = \frac{1}{1-\frac{1}{\psi}} C^{1-\frac{1}{\psi}}, \ \mathcal{I}[V] = \phi^{-1}(E[\phi(V)]), \tag{23}$$

where $\phi(x) = \left[\frac{1-\frac{1}{\psi}}{1-\gamma}x\right]^{\frac{1-\gamma}{1-1/\psi}}$, it follows from Condition (b) that $\mathcal{I}$ is quasiconcave and therefore requires a non-negative 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.

(iii) *In general, preference for early resolution of uncertainty is neither sufficient nor necessary for generalized risk sensitivity.*

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

(iv) *Generalized risk sensitivity and indifference toward the timing of resolution of uncertainty imply the following "maxmin expected utility" representation:*

$$\mathcal{I}[V] = \inf_{m \in M(F_V)} \int mV dP. \tag{24}$$

In the above expression, F_V stands for the distribution of V, and $M(F_V)$ is a family of densities that depends on F_V. In the maxmin expected utility of Gilboa and Schmeidler (1989), the set of priors are typically specified without referencing the distribution of

V. Therefore, the above representation (24) contains preferences not allowed in Gilboa and Schmeidler (1989). If we further require $\mathcal{I}$ to be quasiconcave, then Theorem 1 in Strzalecki (2013) implies that $M(F_V)$ cannot depend on F_V and $\mathcal{I}[V]$ is the maxmin expected utility in the sense of Gilboa and Schmeidler (1989).

In the above sections, we have set up our model in a finite-horizon setting. Extending our results to the infinite-horizon setting will be an important topic for future research. It will require imposing conditions on u and $\mathcal{I}$ so that infinite repetitions of the recursion in (15) and (16) converge to a limit in an appropriately defined functional space. One will also need to show that doing so preserves differentiability of the value function V, as we establish in our finite-horizon setting in Appendix B. Given the differentiability of the value function, Theorem 2 above can be directly applied to establish the equivalence of generalized risk sensitivity and the non-negativity of the announcement premium.

We now turn to the asset pricing implications of our theory.

4.4 Asset pricing implications

Risk compensation for news Compared to traditional consumption-based asset pricing models, our setup decomposes intertemporal returns into an announcement return and a post-announcement return. At all pre-announcement history z_t^-, we have

$$E\left[m^*\left(V\left(\cdot|\, z_t^-\right)\right) R_{A,j}\left(\cdot|\, z_t^-\right)\middle|\, z_t^-\right] = 1, \tag{25}$$

and at all post-announcement history $z_t^+ = \left(z_t^-, s_t^+\right)$, where s_t^+ is an announcement after the history z_t^-,

$$E\left[y^*\left(\cdot|\, z_t^+\right) R_{P,j}\left(\cdot|\, z_t^+\right)\middle|\, z_t^+\right] = 1. \tag{26}$$

$y^*\left(\cdot|\, z_t^+\right)$ in the above equation is an SDF that, given information at history z_t^+, prices all assets that pay off in the next period at history $\left\{\left(z_t^+, s_{t+1}^-\right)\right\}_{s_{t+1}^-}$. Using the law of iterated expectations, the above two equations can be combined to write

$$E\left[m^* y^* \overrightarrow{R}_j\left(\cdot|\, z_t^-\right)\middle|\, z_t^-\right] = 1, \tag{27}$$

where $\overrightarrow{R}_j\left(\cdot|\, z_t^-\right) = R_{A,j}\left(\cdot|\, z_t^-\right) \cdot R_{P,j}\left(\cdot|\, z_t^+\right)$ is the cumulative return for asset j on the pre- and post-announcement markets. The A-SDF only depends on the curvature of the certainty equivalent functional $\mathcal{I}$, and the SDF y^* depends on both the curvature of u and the curvature of $\mathcal{I}$.

Theorem 2 implies that generalized risk sensitivity is precisely the class of preferences under which m^* is a decreasing function of continuation utility and therefore enhances risk compensation. That is, it is the class of preferences under which news about future continuation utility requires a risk compensation. Hansen and Sargent (2008) use a risk-sensitive operator to motivate an additional component in the SDF that increases its volatility. In this sense, our theory generalizes the notion of risk sensitivity of Hansen and Sargent (2008).

In the literature, risk compensation for news about the future is often attributed to uncertainty aversion in the maxmin expected utility model and preference for early resolution of uncertainty in the CES recursive utility model. These intuitions are valid because, as we discussed in Section 4.3, uncertainty aversion and preference for early resolution of uncertainty provide sufficient conditions for generalized risk sensitivity, respectively, in the context of the maxmin expected utility model and in the context of the CES recursive utility model. In general, however, as we demonstrated in Section 4.3, preference for early resolution of uncertainty is neither necessary nor sufficient for generalized risk sensitivity. It is possible to model risk compensation for news without assuming any preference for early resolution of uncertainty, because generalized risk sensitivity is the essential property of preferences that is responsible for this asset pricing implication.

Quantifying the importance of generalized risk sensitivity Just like asset market returns are informative about the property of SDFs, announcement returns are informative about the quantitative magnitude of the A-SDF. We make the following observations:

(i) 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. This is because the A-SDF, m^*, depends only on the curvature of the certainty equivalent functional $\mathcal{I}[\cdot]$, and not that of the $u(\cdot)$ function.

(ii) The entropy bounds of Bansal and Lehmann (1997) and Alvarez and Jermann (2005) provide some insights about the contribution of m^* to equity risk premiums. To save notation, we focus on unconditional expectations here and suppress the dependence of SDF and returns on history. For any variable X, let $L(X) = \ln E[X] - E[\ln X]$ be the entropy of its distribution. The entropy bound implies that $L(m^*) \geq E[\ln R_{A,j}]$ for announcement returns, and $L(m^* y^*) \geq E\left[\ln \overrightarrow{R}_j - \ln R_0\right]$ for the cumulative returns on the pre- and post-announcement markets. Using the average annual market return in Table I, $L(m^*) \geq 3.17\%$ per annum and $L(m^* y^*) \geq 5.08\%$ per annum. The announcement returns are large and compose about 55% of the total equity premium. This clearly suggests a large contribution of m^* to the market price of risk. On a daily basis, the equity premium on announcement days is 11.2 bps, whereas the average daily return in the entire sample period is 2.5 bps. The lower bound on the entropy of the SDF on announcement days is roughly five times of that on an average trading day. This evidence implies that $L(m^*)$ is sizable and that models in which announcement returns are absent or small are mis-specified from the perspective of asset market data.

(iii) The Hansen and Jagannathan (1991) bound for the SDF's leads to a similar conclusion. Equation (13) implies that for any announcement return, $R_{A,j}$, $\sigma[m^*] \geq \frac{E[R_{A,j}]-1}{\sigma[R_{A,j}]}$.Using the Sharpe ratio for the announcement-day returns reported in Table I, we have $\sigma[m^*] \geq 9.85\%$ at the daily level. This bound is much tighter than the Hansen-Jaganathan bound derived from the annualized market returns for the SDF $m^* y^*$ during the same period: $\sigma[m^* y^*] \geq 2.55\%$.

Pre-FOMC announcement drift The theoretical notion of announcements in our model can be interpreted as pre-scheduled macroeconomic announcements or informative signals before the officially scheduled announcements. As a result, Theorem 2 is also a statement about the pre-announcement drift. That is, if the contents of announcements are communicated to the public before the pre-scheduled announcements, then these communications will be associated with realizations of risk premiums under generalized risk sensitivity. In the continuous-time example in the next section, we demonstrate our model's implications for both the announcement premium and the pre-announcement drift.

5 Continuous-time examples

In this section, we use a continuous-time setup to discuss the implications of several examples of dynamic preferences for the announcement premium. We first provide an example of generalized risk sensitivity by using the continuous-time version of the recursive preference developed by Duffie and Epstein (1992). The continuous-time model allows us to highlight the high-frequency nature of announcement returns by distinguishing between the compensation for generalized risk sensitivity that is instantaneously realized upon announcements and the risk premium that accumulates incrementally over time as shocks to consumption materialize. We also use this example to analyze the pre-FOMC announcement drift, assuming information is communicated hours before the scheduled announcement. Finally, we use the continuous-time setup to provide characterizations of the announcement premium implied by some time-non-separable utilities that our representation (14) does not allow for.

5.1 Consumption and information

We consider a continuous-time representative agent economy, where the growth rate of aggregate consumption contains a predictable component, x_t, and an i.i.d. component modeled by increments of a Brownian motion:

$$\frac{dC_t}{C_t} = x_t dt + \sigma dB_{C,t}.$$

Similar to the model of Ai (2010), we assume that x_t is a continuous-time AR(1) process (an Ornstein-Uhlenbeck process) unobservable to the agent in the economy. The law of motion of x_t is

$$dx_t = a_x \left(\bar{x} - x_t\right) dt + \sigma_x dB_{x,t}, \tag{28}$$

where $B_{C,t}$ and $B_{x,t}$ are independent standard Brownian motions.

We assume that the prior belief of the representative agent about x_0 can be represented by a normal distribution. The agent can use two sources of information to update beliefs about x_t. First, the realized consumption path contains information about x_t, and second, at pre-scheduled discrete time points $T, 2T, 3T, \cdots$, additional signals about x_t are revealed through announcements. 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 i.i.d. over time, and normally distributed with mean zero and variance σ_S^2.

Because the posterior distribution of x_t is Gaussian, it can be summarized by the first two moments. We define $\hat{x}_t = E_t\left[x_t\right]$ as the posterior mean and $q_t = E_t\left[\left(x_t - \hat{x}_t\right)^2\right]$ as the posterior variance, respectively, of x_t given information up to time t. At time $t = nT$, where n is an integer, the agent updates his beliefs using Bayes' rule:

$$\hat{x}_{nT}^{+} = \frac{1}{q_{nT}^{+}}\left[\frac{1}{\sigma_S^2} s_n + \frac{1}{q_{nT}^{-}} \hat{x}_{nT}^{-}\right]; \qquad \frac{1}{q_{nT}^{+}} = \frac{1}{\sigma_S^2} + \frac{1}{q_{nT}^{-}}, \tag{29}$$

where $\hat{x}_{nT}^{+}$ and q_{nT}^{+} are the posterior mean and variance after announcements, and $\hat{x}_{nT}^{-}$ and q_{nT}^{-} are the posterior mean and variance before announcements, respectively.

In the interior of $(nT, (n+1)\,T)$, the agent updates his beliefs based on the observed consumption process using the Kalman-Bucy filter:

$$d\hat{x}_t = a_x \left[\bar{x} - \hat{x}_t\right] dt + \frac{q(t)}{\sigma} d\tilde{B}_{C,t}, \tag{30}$$

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:

$$dq(t) = \left[\sigma_x^2 - 2a_x q(t) - \frac{1}{\sigma^2} q^2(t)\right] dt. \tag{31}$$

5.2 Generalized risk sensitive preferences

Preferences and the stochastic discount factor We first consider a simple example of generalized risk sensitivity. We assume that the representative agent has a Kreps-Porteus utility with $\gamma > \frac{1}{\psi}$, and we specify the continuous-time preference as the limit of the discrete-time recursion in (14). Over a small time interval $\Delta > 0$,

$$V_t = \left(1 - e^{-\rho\Delta}\right) u(C_t) + e^{-\rho\Delta} \mathcal{I}\left[V_{t+\Delta} \middle| \hat{x}_t, q_t\right], \tag{32}$$

where u and $\mathcal{I}\left[\cdot \middle| \hat{x}_t, q_t\right]$ are given in equation (23). To derive closed-form solutions, we focus on the limiting case of $\psi = 1$, where $u(C) = \ln C$ and $\mathcal{I}[V] = \frac{1}{1-\gamma} \ln E\left[e^{(1-\gamma)V}\right]$.

Like in previous discrete-time examples, the stochastic discount factor over a small interval $(t, t+\Delta)$ is given by

$$SDF_{t,t+\Delta} = e^{-\rho\Delta} \frac{u'(C_{t+\Delta})}{u'(C_t)} \frac{e^{(1-\gamma)V_{t+\Delta}}}{E_t\left[e^{(1-\gamma)V_{t+\Delta}}\right]}. \tag{33}$$

Clearly, the term $m^*_{t+\Delta} = \frac{e^{(1-\gamma)V_{t+\Delta}}}{E_t\left[e^{(1-\gamma)V_{t+\Delta}}\right]}$ is a density and can be interpreted as a probability distortion.

Announcement premiums We assume that the stock market is the claim to the following dividend process:

$$\frac{dD_t}{D_t} = \left[\bar{x} + \phi(x_t - \bar{x})\right] dt + \phi\sigma dB_{C,t}, \tag{34}$$

and we allow the leverage parameter $\phi > 1$ so that dividends are more risky than consumption, as in Bansal and Yaron (2004).

In the interior of $(nT,(n+1)T)$, all state variables, C_t, $\hat{x}_t$, and q_t in (33) are continuous functions of t. As a result, as $\Delta \to 0$, $SDF_{t,t+\Delta} \to 1$ and the equity premium on any asset must converge to zero. In fact, using a first-order approximation, we show in Section S. 3.1 of the Supplemental Material that the equity premium over the interval $(t, t+\Delta)$ must be proportional to the holding period Δ:

$$\left[\gamma\sigma + \frac{\gamma-1}{a_x+\rho}\frac{q_t}{\sigma}\right]\left[\phi\sigma + \frac{\phi-1}{a_x+e^{-\bar{\varrho}}}\frac{q_t}{\sigma}\right]\Delta, \tag{35}$$

where $\bar{\varrho}$ is the steady-state log price-to-dividend ratio.

In contrast, let $t = T - \frac{1}{2}\Delta$, so that the interval $(t, t+\Delta)$ always contains an announcement. As $\Delta \to 0$, the term $e^{-\rho\Delta}\frac{u'(C_{t+\Delta})}{u'(C_t)}$ converges to 1, but $m^*_{t+\Delta}$ does not. Because the value function $V_{t+\Delta}$ depends on the announcement, and $E_t\left[e^{(1-\gamma)V_{t+\Delta}}\right]$ does not, the probability distortion does not disappear as $\Delta \to 0$. In Section S. 3.1 of the Supplemental Material, we show that in the limit, the announcement premium can be approximated by

$$\frac{\gamma-1}{a_x+\rho}\frac{\phi-1}{a_x+e^{-\bar{\varrho}}}\left(q_T^- - q_T^+\right). \tag{36}$$

We make the following two observations.

(i) As $\Delta \to 0$, the market equity premium vanishes without announcements, but stays strictly positive if an announcement is made during the interval $(t, t+\Delta)$, as shown in Equations (35) and (36).

Figure 3. **Average hourly return around announcements**

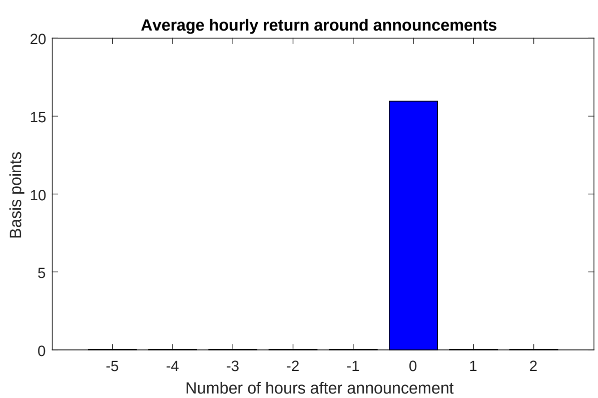


Figure 3 plots the model implied average hourly return around pre-scheduled announcements.

In Figure 3, we plot the average hourly market return around announcements. We choose standard parameters used in the long-run risks literature, the details of which are provided in Section S. 3.2 of the Supplemental Material. The premium realized during the announcement hour is 17 bps, whereas the average return during non-announcement hours is much smaller by comparison. This pattern is consistent with the bottom panel of Figure 1.

(ii) As shown in Equation (36), the magnitude of the announcement premium is proportional to the amount of uncertainty reduction, $q_T^- - q_T^+$, and is increasing in the persistence of the x_t process. Increases in the persistence of x_t, which is inversely related to a_x, have two effects.[15] First, they imply that revelations of x_t have a stronger impact on continuation utility $V_{t+\Delta}$ and therefore $m^*_{t+\Delta}$. Second, more persistence in the expected growth rate of cash flow is also associated with a stronger effect of announcements on the price-to-dividend ratio of the equity. Together, they imply that the announcement premium must increase with the persistence of x_t. The above observation implies that the heterogeneity in the magnitude of the premium for different macroeconomic announcements can be potentially explained by the differences in their informativeness and the significance of their welfare implications.

(iii) In our endowment economy model, although instantaneous consumption C_T does not respond to the announcement made at time T, future consumption does, as x_T is the expected consumption growth rate. Our results below for the announcement premium and the pre-announcement drift will continue to hold in neoclassical production economies, where x_t is interpreted as expected productivity growth and consumption is allowed to respond instantaneously to announcements about x_t. As we remarked earlier, in production economies, the instantaneous reaction of consumption to announcements contributes to a small and negative premium, but the overall announcement premium is positive as long as we allow for significant generalized risk sensitivity in the preferences.[16]

Pre-FOMC announcement drift As discussed earlier in the paper, the announcement in our theory represents a resolution of macroeconomic uncertainty. It can be due to pre-scheduled macroeconomic announcements, informal communications from Fed officials to the public, or information leakage. In our continuous-time model, we simply assume that the agent receives informative signals prior to the FOMC announcements and

[15] The autocorrelation between x_t and $x_{t+\Delta}$ is roughly $1 - a_x\Delta$ for small values of Δ.

[16] We have solved a model with neoclassical production technology and obtained similar results for announcement premiums and the pre-announcement drift. The results are available upon request.

Figure 4. **Pre-FOMC announcement drift**

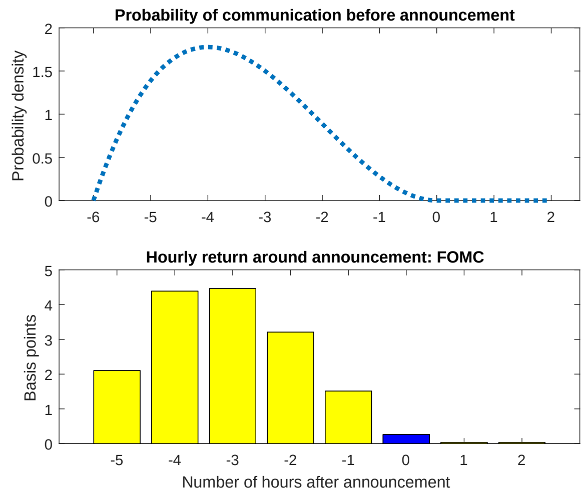


Figure 4 plots the probability density of communication before announcement (top panel) and the average hourly return around announcements (bottom panel).

explore its implications for the pre-announcement drift. Recent research provides suggestive evidence of information leakage as a plausible channel for these signals assumed in our model. Cieslak, Morse, and Vissing-Jorgensen (2015) provide evidence of systematic informal communication of Fed officials with the media and financial sector as the information transmission channel. In a similar vein, Bernile, Hu, and Tang (2016) find evidence consistent with informed trading during a very short window (approximately 30 minutes) of news embargoes prior to the FOMC scheduled announcements and not of other macroeconomic announcements.[17] While the timing of the above empirical evidence of leaks does not exactly match the timing of the pre-announcement drift reported in Lucca and Moench (2015), it is generally indicative of the possibility of information leaks.

In Figure 4, we plot the implication of our model for the pre-FOMC announcement drift, assuming communications occur before announcements. For simplicity, we assume that communication, whenever it occurs, fully reveals x_t, and, we plot the probability density of communication at time t (y-axis) as a function of t (x-axis) in the top panel of Figure 4. In the bottom panel, we plot the model-implied average hourly market return around

[17]In the transcripts of the October 15, 2010 FOMC conference call, then Chairman Ben Bernanke also expressed concerns about information leaks to market participants. See https://www.federalreserve.gov/monetarypolicy/files/FOMC20101015confcall.pdf.

announcements. We provide the details of the calculation of the pre-announcement drift in Section S. 3.2 of the Supplemental Material. Note that the magnitude of the announcement premium is proportional to the probability of the occurrence of communication. The announcement premium peaks during hours with the highest probability of communication and converges to zero as $t \to 0$, because communication occurs with probability one before the pre-scheduled announcement time. This pattern of the pre-announcement drift implied by our model is very similar to its empirical counterpart in Figure 1.

To evaluate the dynamics of nominal bond returns and the announcement premium in the bond markets, we also solve a more extensive model related to Piazzesi and Schneider (2006) and Bansal and Shaliastovich (2013) with growth and inflation dynamics. We show that consistent with the data, the bond announcement premium in our calibrated model is about 3 bps. We also demonstrate that in small samples comparable to those used in earlier empirical work, the pre-announcement drift is present in the equity markets but statistically absent in bond returns, because the announcement premium for bonds is substantially smaller in magnitude than is that for equity.[18]

5.3 Time-non-separable preferences

In this section, we analyze several examples of time-non-separable preferences that are not allowed by representation (14). We continue to use the specification of consumption and information structure in Section 5.1. We assume that the representative agent ranks intertemporal consumption plans according to the following utility function:

$$E\left[\int_0^\infty e^{-\rho t} u\left(C_t + bH_t\right) dt\right], \tag{37}$$

for some appropriately defined habit process $\{H_t\}_{t=0}^\infty$, which we specify below. The above representation includes the external habit model of Campbell and Cochrane (1999); the internal habit model of Constantinides (1990) and Boldrin, Christiano, and Fisher (2001); and the consumption substitutability model (see Dunn and Singleton (1986) and Heaton (1993)) as special cases. In this section, we denote the marginal utility of C_t as

$$\Lambda_t = \frac{\partial}{\partial C_t} E_t\left[\int_0^\infty e^{-\rho(t+s)} u\left(C_{t+s} + bH_{t+s}\right) ds\right].$$

[18]This evidence is available from the authors on request.

The sign of the announcement premium depends on how the marginal utility, Λ_t reacts to the announcement at time t. We make the following observations and provide the detailed proofs in Section S. 3.2 of the Supplemental Material.

(i) *The external habit model has zero announcement premium.*

Suppose $b \in (-1, 0)$ and H_t is a habit process defined as

$$H_t = \left(1 - \int_0^t \xi(t, s)\, ds\right) H_0 + \int_0^t \xi(t, s)\, C_s ds, \tag{38}$$

where $\{\xi(t, s)\}_{s=0}^t$ is a non-negative weighting function that satisfies the regularity conditions (S. 3.16)-(S. 3.18) in Section S. 3.2 of the Supplemental Material. In the external habit model, the consumption, $\{C_s\}_{s=0}^t$ in equation (38), is interpreted as aggregate consumption, which is exogenous to the choice of the agent. Our specification is therefore a generalization of the Campbell and Cochrane (1999) model in continuous time. Because the habit stock H_t is exogenous, like in expected utility models, marginal utilities depend on current-period consumption only:

$$\Lambda_t = e^{-\rho t} u'(C_t + bH_t).$$

Clearly, news about the future does not affect Λ_t and the announcement premium must be zero.

Strictly speaking, the external habit model is time-separable. Because individuals take the habit stock as given, the external habit preference is essentially an expected utility with time-varying risk aversions. It has a zero announcement premium for the same reason that the expected utility model does.

(ii) *The internal habit model generates a negative announcement premium.*

We continue to assume $b \in (-1, 0)$ and (38), except that $\{C_s\}_{s=0}^t$ in equation (38) is interpreted as the agent's own consumption choice. This model is a generalized version of the Constantinides (1990) model. The marginal utility of C_t for the internal habit model can be written as:

$$\Lambda_t = e^{-\rho t}\left\{u'(C_t + bH_t) + bE\left[\int_0^\infty e^{-\rho s}\xi(t+s, t)\, u'(C_{t+s} + bH_{t+s})\, ds \middle| \hat{x}_t, q_t\right]\right\}. \tag{39}$$

We show in Section S. 3.2 of the Supplemental Material that Λ_t in (39) is an increasing function of $\hat{x}_t$. Therefore, the internal habit model implies a negative premium for any return positively correlated with $\hat{x}_t$.

The marginal utility Λ_t increases with $\hat{x}_t$ because good news about the future lowers the negative impact of accumulating habit stock. While investors with an external habit preference take the habit process as exogenous to their choices, internal habit utility maximizers take into account the effect of current-period consumption on future habit stocks when computing marginal utilities. As shown in equation (39), an increase in C_t has a positive effect on the current-period utility, which is $u'(C_t + bH_t)$, but a negative impact on utility in all future periods, because it raises the level of H_{t+s} for all $s \geq 0$. The negative impact of accumulating the habit stock is captured by the expectation of marginal utilities in the future: $E\left[\int_0^\infty e^{-\rho s}\xi(t+s,t)\, u'(C_{t+s} + bH_{t+s})\, ds \middle| \hat{x}_t, q_t\right]$ in equation (39). Good news about consumption growth leaves the current-period marginal utility, $u'(C_t + bH_t)$, unchanged, but lowers the marginal utility in all future periods. As $b < 0$, this reduction in future marginal utilities in response to positive innovations in $\hat{x}_t$ raises the overall marginal utility, Λ_t.

(iii) *The consumption substitutability model produces a positive announcement premium.*

Suppose the agent's preference is defined by (37) and (38) with $b > 0$. In this case, past consumption increases current-period utility. Opposite of the internal habit model, the marginal utility (39) is a decreasing function of $\hat{x}_t$. Therefore, the announcement premium is positive for any asset with a return positively correlated with $\hat{x}_t$. Even though the presence of consumption substitutability produces a positive announcement premium, it lowers the agent's effective risk aversion and exacerbates the equity premium puzzle, as emphasized by Gallant, Hansen, and Tauchen (1990).

6 Conclusion

Motivated by the fact that a large fraction of the market equity premium is realized on a small number of trading days with significant macroeconomic announcements, in this paper, we provide a revealed preference analysis of the equity premium for macroeconomic announcements. Assuming that consumption does not respond instantaneously to announcements, we show that a non-negative announcement premium is equivalent to generalized risk sensitivity; that is, investors' certainty equivalent functional is monotone in second-order stochastic dominance. We demonstrate that generalized risk sensitivity is exactly the class of preferences that demands a risk compensation for news that affects continuation utilities, or “long-run risks.” As a result, our theoretical framework implies that the announcement premium can be interpreted as asset-market-based evidence for a broad class of non-expected utility models that have this feature.

Because of its high-frequency nature, continuous-time models are particularly convenient for studying the announcement premium and the pre-announcement drift in the FOMC announcements. We show in a continuous-time model that the pre-announcement drift can arise in environments in which information about announcements is communicated to the market prior to the scheduled announcement.

We assume a representative agent throughout the paper; however, some of our results may extend to more general setups. For example, the result that the expected utility implies a zero announcement premium on all assets should also hold in complete-market economies where agents’ preferences are heterogenous, but all have an expected utility representation. Standard aggregation results imply that asset prices in these economies are observationally equivalent to a representative-agent economy with time-separable expected utility. This observation should also apply to the external habit model.

Several related topics may provide promising directions for future research. A natural extension of the current paper is to provide a characterization for generalized risk sensitivity in the continuous-time framework. Such conditions may bear interesting connections with Skiadas (2013), who provides a continuous-time analysis of certainty equivalent functionals for non-expected utilities. Another idea worth careful exploration is to evaluate whether asset market frictions related to liquidity or slow-moving capital, as emphasized in Duffie (2010), may contribute to the announcement premium and the pre-announcement drift. Finally, our theory has several implications that may be tested empirically. For example, our analysis implies that the magnitude of the announcement premium is determined by how informative an announcement is about the future course of the economy. In addition, a sizable literature documents significant excess returns at the firm level around earnings announcements (see e.g., Chari, Jagannathan, and Ofer (1988) and Ball and Kothari (1991)). To the extent that these earnings announcements carry news about the macroeconomy, premiums associated with earnings announcements can be consistent with our theory. Further exploration of this issue may be an interesting direction for future research.

APPENDICES

The following appendices provide details of the data construction for the stylized facts in Section 2 and the proofs of the main results in Section 4. Appendix A is the data appendix. Appendix B contains the proofs of Theorem 1 and Theorem 2. Appendix C provides the omitted proofs for the results on the relationship between generalized risk sensitivity, uncertainty aversion, and preference for early resolution in Section 4.3.

Appendix A Data Description

Macroeconomic announcements We focus on the top five macroeconomic news ranked by investor attention among all macroeconomic announcements at the monthly or lower frequencies. They are unemployment/non-farm payroll (EMPL/NFP) and the producer price index (PPI) published by the U.S. Bureau of Labor Statistics (BLS), the FOMC statements, gross domestic product (GDP) reported by the U.S. Bureau of Economic Analysis, and the Institute for Supply Management's Manufacturing Report (ISM) released by Bloomberg.[19]

The EMPL/NFL and the PPI are both published monthly 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 starting date for our EMPL/NFL announcements sample. The PPI data series starts in 1971.[20] 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 began 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.[21] Finally, ISM is a monthly

[19]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.

[20]While the CPI data are 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 is already known to the market at the time of its release. For this reason, we opt in favor of using the PPI.

[21]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 focus only on the four advance estimate release dates every year.

announcement with dates coming from Bloomberg starting from 1997. Our sample ends in 2014.

High-frequency returns In Table III and Figure 1, we report the average stock market excess returns over one-hour intervals before and after news announcements in event time. Here, we use high-frequency data for the S&P 500 SPDR that runs from 1997 to 2013 and comes from the TAQ database. For each second, the median price of all transactions occurring in that second is computed. Prices at lower frequency intervals (e.g. hourly prices) are then constructed as the price for the last (most recent) second in that interval when transactions were observed. The exact times at which the announcements are released are reported by Bloomberg.

Appendix B Proof of Theorems 1 and 2

B.1 Preliminaries

We first state the definition of a non-atomic probability space, which is an assumption maintained throughout Section 4.

Definition B.1. *Non-atomic probability space:*

A probability space $(\Omega, \mathcal{F}, P)$ is said to be *non-atomic* if $\forall \omega \in \Omega$, $P(\omega) = 0$.

Next, we state the definition of first-order stochastic dominance (FSD) and second-order stochastic dominance (SSD).

Definition B.2. *First-order stochastic dominance:*

X_1 *first-order stochastic dominates* X_2, or $X_1 \geq_{FSD} X_2$, if there exists a random variable $Y \geq 0$ a.s. such that X_1 has the same distribution as $X_2 + Y$. X_1 strictly *first-order stochastic dominates* X_2, or $X_1 >_{FSD} X_2$, if $P(Y > 0) > 0$ in the above definition.

Definition B.3. *Second-order stochastic dominance:*

X_1 *second order stochastic dominates* X_2, or $X_1 \geq_{SSD} X_2$, if there exists a random variable Y such that $E[Y|X_1] = 0$ and X_2 has the same distribution as $X_1 + Y$. X_1 *strictly second order stochastic dominates* X_2, or $X_1 >_{SSD} X_2$, if $P(Y \neq 0) > 0$ in the above definition.[22]

[22]Our definition of SSD is the same as the standard concept of increasing risk (see Rothschild and Stiglitz (1970) and Werner (2009)). However, it is important to note that in our model, the certainty equivalent functional $\mathcal{I}$ is defined on the space of continuation utilities rather than consumption.

FSD and SSD are typically defined as stochastic orders on the space of distributions. Here, it is more convenient to define FSD and SSD as binary relations on the space of random variables. Our definitions are equivalent to the standard definitions of FSD and SSD due to the assumption of a non-atomic probability space (See Muller and Stoyan (2002)).

Our strategy for proving Theorem 1 and 2 consists of two steps. First, we apply the envelope theorems in Milgrom and Segal (2002) to establish the differentiability of the value functions. Second, we compute the derivatives of $\mathcal{I}$ to construct the A-SDF and use derivatives of $\mathcal{I}$ to integrate back to recover the certainty equivalent functional.[23]

Most of our analysis below is on the conditional certainty equivalent functional $\mathcal{I}[\cdot|\, z]$. To save notation, whenever it does not cause confusion, we suppress the dependence of $\mathcal{I}[\cdot|\, z]$ on z and simply write $\mathcal{I}[\cdot]$. We often use the following operation to relate the certainty equivalent functional $\mathcal{I}$ 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

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

where $D\mathcal{I}[X + t(Y - X)]$ is understood as the representation of the Fréchet derivative of $\mathcal{I}[\cdot]$ evaluated at $X + t(Y - X)$. The Riesz representation theorem implies that $D\mathcal{I}[X + t(Y - X)]$ is an element of $L^2(\Omega, \mathcal{F}, P)$, and $D\mathcal{I}[X + t(Y - X)]$ applied to $(Y - X)$ can be computed as the dot product, $\int_\Omega D\mathcal{I}[X + t(Y - X)](Y - X)\, dP$.

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 (1997)). To see that $g'(t)$ is continuous, note that

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

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

The Lipschitz continuity of $D\mathcal{I}$ implies that, for some positive constant K,

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

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 apply (40) to write for some $\hat{t} \in (0, 1)$,

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

Much of our analysis below relies on the theory of differentiability for nonlinear operators defined on infinite dimensional spaces, for example, in Tapia (1971) and Luenberger (1997).

B.2 Existence of A-SDF

In this section, we provide a proof for Theorem 1 and establish the existence of A-SDF.

Differentiability of value function We establish the differentiability of value functions recursively. In particular, we show that the value functions are elements of $\mathcal{D}$, which is defined as:

Definition B.4. $\mathcal{D}$ is the set of differentiable functions on the real line, denoted by f, that satisfy the following two properties.

(i) f is Lipschitz continuous and $f'(x) > 0$.

(ii) $\forall x \in \mathbf{R}$, as $h \to 0$, $\frac{1}{h}[f(x + h - a) - f(x - a)]$ converges uniformly to $f'(x - a)$ in a. That is, $\forall \varepsilon > 0$, there exists $\delta > 0$ such that $|h| < \delta$ implies that $\left|\frac{1}{h}[f(x + h - a) - f(x - a)] - f'(x - a)\right| < \varepsilon$ for all $a \in \mathbf{R}$.

We first introduce some notations. 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, by:

$$f_v(W, \xi) = u\left(W - \sum\nolimits_{j=0}^{J} \xi_j\right) + \beta\mathcal{I}\left[v\left(\sum\nolimits_{j=0}^{J} \xi_j R_j\right)\right], \tag{42}$$

$$g_v(W, \xi) = \mathcal{I}\left[v\left(W + \sum\nolimits_{j=0}^{J} \xi_j (R_j - 1)\right)\right]. \tag{43}$$

Because $R_j \in L^2(\Omega, \mathcal{F}, P)$ and v is Lipschitz continuous, for a fixed ξ, $v\left(\sum_{j=0}^{J} \xi_j R_j\right)$ and $v\left(W + \sum_{j=0}^{J} \xi_j (R_j - 1)\right)$ are both square-integrable and equations (42) and (43) are well-defined.

We define two operators on $\mathcal{D}$. For any $v \in \mathcal{D}$, let T^+v and T^-v be defined by:

$$\left[T^+ v\right](W) = \sup_{\xi} f_v(W, \xi), \quad (44)$$

$$\left[T^- v\right](W) = \sup_{\xi} g_v(W, \xi). \quad (45)$$

Clearly, the value functions $V_{z_t^+}(W)$ and $V_{z_t^-}(W)$ can be constructed recursively as $V_{z_t^+}(W) = \left[T^+ V_{z_{t+1}^-}\right](W)$, and $V_{z_t^-}(W) = \left[T^- V_{z_t^+}\right](W)$ (with the understanding that the certainty equivalent functionals in the definition of $f_v(W, \xi)$ and $g_v(W, \xi)$ are appropriately chosen conditional certainty equivalent functionals). Because we start with the assumption of the existence of an interior equilibrium, the maximization problems (44) and (45) are well defined, and the maximums are achieved.

Below, we prove that $V_{z_t^+}$ and $V_{z_t^-}$ are elements of $\mathcal{D}$ in two steps. First, Lemma B.1 below establishes the equi-differentiability of the family of functions $\{f_v(W, \xi)\}_\xi$ and $\{g_v(W, \xi)\}_\xi$ so that we can apply the envelope theorem in Milgrom and Segal (2002). Second, in Lemma B.2, we apply the envelope theorem repeatedly to show that the operators T^+ and T^- map $\mathcal{D}$ into itself.

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

Proof: *First, as $h \to 0$,*

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

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

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

and the convergence is uniform for all ξ. Note that

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

and

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

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

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

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

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

Proof: *It follows from Lemma B.1 that for any* $v\in\mathcal{D}$*, we can apply Theorem 3 in Milgrom and Segal (2002) and establish that both* T^+v *and* T^-v *are differentiable, and*

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

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

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

$$f_v(W_1,\xi(W_2))-f_v(W_2,\xi(W_2))\le T^+v(W_1)-T^+v(W_2)\le f_v(W_1,\xi(W_1))-f_v(W_2,\xi(W_1)). \qquad (48)$$

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

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

Note that $\forall\xi$,

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

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

In addition, equations (48) and (49) can be used to show that the family of functions $\{T^{+}v(W-a)\}_a$ *and* $\{T^{-}v(W-a)\}_a$ *are equi-differentiable. For example, let* $W_1\to W_2$,

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

converges uniformly by Lemma B.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.*

Finally, we note that if $v'(x)>0$ *for all* $x\in\mathbf{R}$*, then* $[T^{+}v](W)$ *and* $[T^{-}v](W)$ *must satisfy the same property by the envelope theorem.*

Proof of Theorem 1 In this section, we establish the existence of SDF as stated in Theorem 1. To save notation, whenever convenient, we denote $R_j(z)$ to be the one-period return of asset j that payoff at history z. That is, if $z=z_t^{+}=\left(z_t^{-},s_t^{+}\right)$ is a post-announcement history, then $R_j(z)\equiv R_{A,j}\left(s_t^{+}\middle|z_t^{-}\right)$ is an announcement return, and if z is of the form $z=z_{t+1}^{-}=\left(z_t^{+},s_{t+1}^{-}\right)$, then $R_j(z)\equiv R_{P,j}\left(s_{t+1}^{-}\middle|z_t^{+}\right)$ is a post-announcement return. We write the portfolio selection problem at z_t^{-} as

$$\max_{\zeta}\mathcal{I}\left[V_{z_t^{+}}\left(W+\sum_{j=0}^{J}\zeta_j\left(R_j\left(z_t^{+}\right)-1\right)\right)\middle|z_t^{-}\right]. \tag{50}$$

Clearly, no arbitrage implies that the risk-free announcement return $R_0\left(z_t^{+}\right)=1$. Because $V_{z_t^{+}}$ and $\mathcal{I}\left[\cdot\middle|z_t^{-}\right]$ are (Fréchet) differentiable, $\mathcal{I}\left[V_{z_t^{+}}\left(W+\sum_{j=0}^{J}\zeta_j\left(R_j\left(z_t^{+}\right)-1\right)\right)\middle|z_t^{-}\right]$ is differentiable in ζ.[24] Therefore, the first order condition with respect to ζ_j implies that

$$E\left[D\mathcal{I}\left[V_{z_t^+}\left(W'\right)\right] \frac{d}{dW'} V_{z_t^+}\left(W'\right)\left(R_j\left(z_t^+\right)-1\right) \middle| z_t^- \right] = 0, \tag{51}$$

where we denote $W' = W + \sum_{j=0}^{J} \hat{\zeta}_j \left(R_j\left(z_t^+\right)-1\right)$ and $\hat{\zeta}$ is the optimal portfolio choice.

The value function $V_{z_t^+}(\cdot)$ in (50) is determined by the the agent's portfolio choice problem at z_t^+ after the announcement s_t^+ is made:

$$V_{z_t^+}(W) = \max_{\xi} \left\{ u\left(W - \sum_{j=0}^{J} \xi_j\right) + \beta \mathcal{I}\left[V_{z_{t+1}^-}\left(\sum_{j=0}^{J} \xi_j R_j\left(z_{t-1}^-\right)\right) \middle| z_t^+ \right] \right\}. \tag{52}$$

The envelope condition for (52) implies

$$\frac{d}{dW} V_{z_t^+}(W) = u'\left(W - \sum_{j=0}^{J} \xi_j\right) = u'\left(C_t\right) = u'\left(\bar{C}_t\right),$$

where the last equality uses the market clearing condition. Because consumption at time t must equal to total endowment, $\bar{C}_t$, and because $\bar{C}_t$ must be z_t^- measurable, so must $\frac{d}{dW} V_{z_t^+}(W)$.

By our results in Appendix B.2, $\frac{d}{dW} V_{z_t^+}(W) = u'\left(\bar{C}_t\right) > 0$ is z_t^- measurable, as a result, (51) implies:

$$E\left[D\mathcal{I}\left[V_{z_t^+}(W)\right] \left(R_j\left(z_t^+\right)-1\right) \middle| z_t^- \right] = 0. \tag{53}$$

As we show in Lemma B.4 in the next section, monotonicity of $\mathcal{I}$ guarantees that $D\mathcal{I} \geq 0$ with probability one. To derive an expression for A-SDF, we need to assume the following slightly stronger regularity condition:

$$D\mathcal{I}[X] > 0 \text{ with strictly positive probability for all } X.^{25} \tag{54}$$

In this case, the A-SDF can be constructed as:

$$m^*\left(s_t^+ \middle| z_t^-\right) = \frac{D\mathcal{I}\left[V_{z_t^+}\left(W_{z_t^-, s_t^+}\right)\right]}{E\left[D\mathcal{I}\left[V_{z_t^+}\left(W_{z_t^-, s_t^+}\right)\right] \middle| z_t^- \right]}, \tag{55}$$

[24] This is a version of the chain rule. See Proposition 1 in Chapter 7 of Luenberger (1997).

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

where W_z denote the equilibrium wealth of the agent at history z. Because $E\left[m^*\left(s_t^+\middle|\, z_t^-\right)\middle|\, z_t^-\right] = 1$, we can write (53) as an asset pricing equation with A-SDF:

$$E\left[m^*\left(\cdot|\, z_t^-\right) R_{A,j}\left(\cdot|\, z_t^-\right)\middle|\, z_t^-\right] = 1.$$

Now we constructed the A-SDF as the Fréchet derivative of the certainty equivalent 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 1, we only need to show that $m^*\left(s_t^+\middle|\, z_t^-\right)$ can be represented as a measurable function of continuation utility: $m^*\left(s_t^+\middle|\, z_t^-\right) = m^* \circ V_{z_t^+}\left(W_{z_t^-, s_t^+}\right)$ for some measurable function $m^* : \mathbf{R} \to \mathbf{R}$.[26] That is, $m^*\left(s_t^+\middle|\, z_t^-\right)$ 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 that $m^*\left(s_t^+\middle|\, z_t^-\right)$ is measurable with respect to the σ-field generated by $V_{z_t^+}\left(W_{z_t^-, s_t^+}\right)$.

Lemma B.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 $X \in L^2(\Omega, \mathcal{F}, P)$, to prove that $D\mathcal{I}[X]$ is a measurable function of X, it is enough to show that $D\mathcal{I}[X]$ is measurable with respect to the $\sigma-$field generated by X (which we denote as $\sigma(X)$). Let T be a measure-preserving transformation such that the invariant $\sigma-$field of T differ from $\sigma(X)$ only by measure zero sets (The assumption of a non-atomic probability space guarantees the existence of such measure-preserving transformations. See exercise 17.43 in Kechris (1995)). Below, we show that $D\mathcal{I}[X]$ is measurable with respect to the invariant $\sigma-$field of T by deomonstrating $D\mathcal{I}[X] \circ T = D\mathcal{I}[X]$ with probability one.*[27]

Because the Fréchet derivative of $\mathcal{I}[X]$ is unique, to establish $D\mathcal{I}[X] = D\mathcal{I}[X] \circ T$, we show that $D\mathcal{I}[X] \circ T$ is also a Fréchet derivative of $\mathcal{I}[\cdot]$ at 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,

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

[26] In general, m^* may depend on z_t^-. Here, with a slight abuse of notation, we denote m^* both as the A-SDF, which is an element of L^2, and as a measurable function $\mathbf{R} \to \mathbf{R}$.

[27] By Proposition 6.17 of Brieman (1992), the statement that $D\mathcal{I}[X]$ is measurable with respect to the invariant $\sigma-$field of T is equivalent to $D\mathcal{I}[X] \circ T = D\mathcal{I}[X]$ with probability one.

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

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

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

where the last equality uses the fact that $[D\mathcal{I}[X] \cdot (Y \circ T^{-1})] \circ T = D\mathcal{I}[X] \circ T \cdot Y$ *have the same distribution with* $D\mathcal{I}[X] \times Y \circ T^{-1}$. *This proves (56).*

B.3 Generalized Risk Sensitivity and the Announcement Premium

We prove Theorem 2 in this section. Part 1 is straightforward given our results in the proof of Theorem 1 in Appendix B.2. From equation (55), if $\mathcal{I}$ is expected utility, then $m^*\left(s_t^+ \middle| z_t^-\right)$ must be a constant. Conversely, if $m^*\left(s_t^+ \middle| z_t^-\right)$ is a constant, then $\mathcal{I}$ is linear and must have an expected utility representation.

We prove part 2) of theorem 2 in three steps. First, we use Lemma B.4 to establish that $m^*\left(V_{z_t^+}\right)$ is non-negative if and only if $\mathcal{I}$ is monotone with respect to FSD. Second, we prove the equivalence between (a) and (b). Lemma B.5 and B.6 jointly establish that generalized risk sensitivity of $\mathcal{I}$ is equivalent to $m^*\left(V_{z_t^+}\right)$ being a non-increasing function of $V_{z_t^+}$. Finally, we use Lemma B.7 to establish the equivalence between (b) and (c).

Lemma B.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., using (40), we have:*

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

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 $m^*\left(V_{z_t^+}\right)$ is non-increasing in $V_{z_t^+}$. We first prove the following lemma.

Lemma B.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 $\mathcal{G} \subseteq \mathcal{F}$,*

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

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*

$$\mathcal{I}(X) \geq \mathcal{I}(X + Y).$$

Using (40),

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

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 $\varepsilon > 0$, $\forall t \in (0, \varepsilon)$,

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

Therefore,

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

However, $(1-\varepsilon)X+\varepsilon E\left[X|\mathcal{G}\right]\geq_{SSD}X$*, a contradiction.*[28]

Due to Lemma B.3, $D\mathcal{I}\left[X\right]$ can be represented by a measurable function of X, we denote $D\mathcal{I}\left[X\right]=\eta\left(X\right)$. To establish the equivalence between monotonicity with respect to SSD and the negative monotonicity of $m^{*}\left(V_{z_t^+}\right)$, we only need to prove that condition (57) is equivalent to $\eta\left(\cdot\right)$ being a non-increasing function, which is Lemma B.6 below.

Lemma B.6. *Condition (57) is equivalent to* $\eta\left(X\right)$ *being a non-increasing function of* X *with probability one.*

Proof: *First, we assume* $\eta\left(X\right)$ *is non-increasing with probability one. To prove (57), note that* $E\left[X|\mathcal{G}\right]$ *is measurable with respect to* $\sigma\left(X\right)$*, and we can use the law of iterated expectation to write:*

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

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

Second, to prove the converse of the above statement by contradiction, we assume (57) is true, but $\eta\left(x\right)$ *is not non-decreasing with probability one. That is, there exist* $x_1<x_2$*, both occur with positive probability such that* $\eta\left(x_1\right)<\eta\left(x_2\right)$*. Under this assumption, we construct a random variable* Y*:*

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

[28] An easy way to prove the statement, $(1-\varepsilon)X+\varepsilon E\left[X|\mathcal{G}\right]\geq_{SSD}X$ is to observe that an equivalent definition of SSD is $X_1\geq_{SSD}X_2$ if $E\left[\phi\left(X_1\right)\right]\geq E\left[\phi\left(X_2\right)\right]$ for all concave functions ϕ (see Rothschild and Stiglitz (1970) and Werner (2009)). To see $(1-\varepsilon)X+\varepsilon E\left[X|\mathcal{G}\right]\geq_{SSD}X$, take any concave function ϕ, we have $\phi\left((1-\varepsilon)X+\varepsilon E\left[X|\mathcal{G}\right]\right)\geq(1-\varepsilon)\phi\left(X\right)+\varepsilon\phi\left(E\left[X|\mathcal{G}\right]\right)$. Taking conditional expectation on both sides, $E\left[\phi\left((1-\varepsilon)X+\varepsilon E\left[X|\mathcal{G}\right]\right)|\mathcal{G}\right]\geq(1-\varepsilon)E\left[\phi\left(X\right)|\mathcal{G}\right]+\varepsilon\phi\left(E\left[X|\mathcal{G}\right]\right)$. Note that $(1-\varepsilon)E\left[\phi\left(X\right)|\mathcal{G}\right]+\varepsilon\phi\left(E\left[X|\mathcal{G}\right]\right)=E\left[\phi\left(X\right)|\mathcal{G}\right]+\varepsilon\left\{\phi\left(E\left[X|\mathcal{G}\right]\right)-E\left[\phi\left(X\right)|\mathcal{G}\right]\right\}\geq E\left[\phi\left(X\right)|\mathcal{G}\right]$, because ϕ is concave. Therefore, $E\left[\phi\left((1-\varepsilon)X+\varepsilon E\left[X|\mathcal{G}\right]\right)|\mathcal{G}\right]\geq E\left[\phi\left(X\right)|\mathcal{G}\right]$. Taking unconditional expectation on both sides, we have $E\left[\phi\left((1-\varepsilon)X+\varepsilon E\left[X|\mathcal{G}\right]\right)\right]\geq E\left[\phi\left(X\right)\right]$, as needed.

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

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

which contradict condition (57).

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

Lemma B.7. *That* $m^*(V)$ *is a non-increasing function of* V *is equivalent to (c).*

Proof: *If* $m^*(\cdot)$ *is a non-decreasing function, then for any payoff* $f\left(\cdot|\, z_t^-\right)$ *that is co-monotone with* $V\left(\cdot|\, z_t^-\right)$, *we have*

$$
E\left[m^*\left(V\left(\cdot|\, z_t^-\right)\right) f\left(\cdot|\, z_t^-\right)\right] \leq E\left[m^*\left(V\left(\cdot|\, z_t^-\right)\right)\right] E\left[f\left(\cdot|\, z_t^-\right)\right] = E\left[f\left(\cdot|\, z_t^-\right)\right],
$$

because $m^*\left(V\left(\cdot|\, z_t^-\right)\right)$ *and* $f\left(\cdot|\, z_t^-\right)$ *are negatively correlated.*[29]

We prove that (c) implies (b) by contradiction. Suppose that the announcement premium is non-negative for all payoffs that are co-monotone with $V\left(\cdot|\, z_t^-\right)$, *but* $m^*(v_1) < m^*(v_2)$ *for some* $v_1 < v_2$, *both of which occur with positive probability. Consider the payoff* $g\left(V\left(\cdot|\, z_t^-\right)\right)$, *where* g *is a function defined on the real line:*

$$
g(v) = \begin{cases} 1 & if \quad v = v_2 \\ -1 & if \quad v = v_1 \\ 0 & otherwise \end{cases}.
$$

Note that $g\left(V\left(\cdot|\, z_t^-\right)\right)$ *is co-monotone with* $V\left(\cdot|\, z_t^-\right)$ *and yet* $E\left[m^*\left(V\left(\cdot|\, z_t^-\right)\right) g\left(V\left(\cdot|\, z_t^-\right)\right)\right] > E\left[g\left(V\left(\cdot|\, z_t^-\right)\right)\right]$, *contradicting a non-negative premium for* $g\left(V\left(\cdot|\, z_t^-\right)\right)$.

[29]Note that the same argument implies that if $m^*(\cdot)$ is a non-decreasing function, then the announcement premium must be non-negative for the following more general class of payoffs: $f\left(s_t^+|\, z_t^-\right) + \varepsilon$, where $E\left[\varepsilon|\, z_t^-, s_t^+\right] = 0$.

Appendix C Generalized Risk-Sensitive Preferences

C.1 Generalized risk sensitivity and uncertainty aversion

In this section, we provide proofs for results for the relationship between generalized risk sensitivity and uncertainty aversion discussed in Section 4.3 of the paper.

- Quasiconcavity implies generalized risk sensitivity.

  The following lemma formalizes the above statement.

  Lemma C.1. *Suppose* $\mathcal{I} : L^2(\Omega, \mathcal{F}, P) \to \mathbf{R}$ *is continuous and 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 and the assumption of a non-atomic probability space, there exists 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 (1995)), 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 $\mathcal{I}$ then implies $\mathcal{I}[X_1] \geq \mathcal{I}[X_1 + Y] = \mathcal{I}[X_2]$, that is, $\mathcal{I}$ satisfies generalized risk sensitivity. □

- Qusiconcavity is not necessary for generalized risk sensitivity.

  It is clear from Lemma C.1 that under continuity, the following condition is sufficient for generalized risk sensitivity:

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

Clearly, this condition is weaker than quasiconcavity.

Here, 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 equivalent functional defined on X such that

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

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 measures 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 $\mathcal{I}$ satisfies 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, $\mathcal{I}\left[\lambda X+(1-\lambda)Y\right]\geq\mathcal{I}\left[X\right]$. 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 $\lambda\in\left[\frac{1}{2},1\right]$,

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

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

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

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

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

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.

- For second-order expected utility, the concavity of ϕ is equivalent to generalized risk sensitivity.

  Proof. certainty equivalent 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 (2009). 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 $\mathcal{I}[\cdot]$ is quasiconcave and satisfies generalized risk sensitivity by Lemma C.1. Conversely, suppose $\mathcal{I}[\cdot]$ satisfies generalized risk sensitivity then $E\left[\phi(X)\right] \geq E\left[\phi(Y)\right]$ whenever $X \geq_{SSD} Y$. By remark B on page 240 of Rothschild and Stiglitz (1970), ϕ is concave. □

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

  Proof. Using the results in Klibanoff, Marinacci, and Mukerji (2005, 2009), it straightforward to show that for the class of smooth ambiguity preferences, concavity of ϕ is equivalent to the quasiconcavity of $\mathcal{I}$. As a result, quasiconcavity implies generalized risk sensitivity by Lemma C.1. 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 $\mathcal{I}[E[V]] \geq \mathcal{I}[V]$, for all $V \in L^2(\Omega, \mathcal{F}, P)$. That is,

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

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. □

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

Below, we provide detailed examples and proofs for the discussions on the relationship between preference for early resolution of uncertainty and generalized risk sensitivity in Section 4.3.

- An example that satisfies generalized risk sensitivity but strictly prefers late resolution of uncertainty.

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

$$u(C) = C - b, \quad \textit{where} \;\; b = 2; \quad \mathcal{I}(X) = \left(E\sqrt{X}\right)^2; \; \textit{and} \;\; \beta = 1.$$

It straight forward to check that $\mathcal{I}$ is quasiconcave 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:

$$V^E = \mathcal{I}[u(C_0) + u(C_1)].$$

It is straightforward to show that:

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

Therefore,

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

The utility for late resolution of uncertainty is given by:

$$V^L = u(C_0) + \mathcal{I}[u(C_1)] = 0.1025.$$

- An example of $\mathcal{I}$ that prefers early resolution of uncertainty but is strictly decreasing in second order stochastic dominance.

Example 2. *Consider the following preference:*

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

Because X^2 is a strictly convex function, the certainty equivalent 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

$$V^E = \mathcal{I}[u(C_0) + u(C_1)] = 0.1485,$$

and the utility for later resolution is:

$$V^L = u(C_0) + \mathcal{I}[u(C_1)] = 0.11.$$

- Generalized risk sensitivity and indifference toward the timing of resolution of uncertainty implies representation (24).

Proof. By Lemma 1 and the proof of Theorem 1 in Strzalecki (2013), indifference between timing of resolution of uncertainty implies that $\mathcal{I}$ satisfies that for all $X \in L^2(\Omega, \mathcal{F}, P)$, all $a \geq 0$, $\mathcal{I}[a + X] = a + \mathcal{I}[X]$. Take derivatives with respect to a and evaluate at $a = 0$, we have:

$$\int D\mathcal{I}[X]\, dP = 1. \tag{61}$$

Note that because $\mathcal{I}$ is normalized, $\mathcal{I}[0] = 0$. Therefore, $\forall X \in L^2(\Omega, \mathcal{F}, P)$,

$$\begin{aligned}
\mathcal{I}[X] &= \mathcal{I}[X] - \mathcal{I}[0] \\
&= \int_0^1 \int D\mathcal{I}[tX] X dP dt \\
&= \int \int_0^1 D\mathcal{I}[tX] dt X dP.
\end{aligned}$$

Note that $\int \int_0^1 D\mathcal{I}[tX] dt dP = 1$ is a density, because of (61). In addition, generalized risk sensitivity implies that for each t,

$$[D\mathcal{I}[tX](\omega) - D\mathcal{I}[tX](\omega')] [X(\omega) - X(\omega')] \leq 0. \tag{62}$$

Therefore, $\int_0^1 D\mathcal{I}[tX] dt$ must satisfy (62) as well. By the result of Carlier and Dana (2003), $\int \int_0^1 D\mathcal{I}[tX] dt X dP$ can be represented by minimization with respect to the core of a convex distortion of P. $\square$

References

Ai, H. (2005): "Smooth Nonexpected Utility with State Independence," *Federal Reserve Bank of Minneapolis Working Paper 637.*

——— (2010): "Information Quality and Long-Run Risk: Asset Pricing Implications," *Journal of Finance*, LXV(4), 1333–1367.

Alvarez, F., and U. Jermann (2004): "Using Asset Prices to Measure the Cost of Business Cycles," *Journal of Political Economy*, 112(6), 1223–1256.

——— (2005): "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica*, 73(6), 1977–2016.

Ball, R., and S. P. Kothari (1991): "Security Returns around Earnings Announcements," *The Accounting Review*, 66(4), 718–738.

Bansal, R. (2007): "Long Run Risks and Financial Markets," *The Review of the St. Louis Federal Reserve Bank*, 89(4), 1–17.

Bansal, R., and B. N. Lehmann (1997): "Growth-Optimal Portfolio Restrictions on Asset Pricing Models," *Macroeconomic Dynamics*, 1, 333–354.

Bansal, R., and I. Shaliastovich (2011): "Learning and Asset Price Jumps," *Review of Financial Studies*, 24(8), 2738–2780.

——— (2013): "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," *Review of Financial Studies*, 26(1), 1–33.

Bansal, R., and A. Yaron (2004): "Risk for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59(4), 1481–1509.

Beaudry, P., and F. Portier (2004): "An exploration into Pigou's theory of cycles," *Journal of Monetary Economics*, 51, 1183–1216.

——— (2014): "News-Driven Business Cycles: Insights and Challenges," *Journal of Economic Literature*, 52(4), 993–1074.

Bernile, G., J. Hu, and Y. Tang (2016): "Can information be locked up? Informed trading ahead of macro-news announcements," *Journal of Financial Economics*, 52(4), 993–1074.

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

Brieman, L. (1992): *Probability.* SIAM: Society for Industrial and Applied Mathematics.

Brusa, F., P. Savor, and M. Wilson (2015): "One Central Bank to Rule Them All," Working paper, Temple University.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107(2), 205–255.

Carlier, G., and R.-A. Dana (2003): "Core of Convex Distortions of a Probability," *Journal of Economic Theory*, 113, 199–222.

Chari, V., R. Jagannathan, and A. R. Ofer (1988): "Seasonalities in security returns: The case of earnings announcements," *Journal of Financial Economics*, 21(1), 101–121.

Chateauneuf, A., and J. H. Faro (2009): "Ambiguity through confidence functions," *Journal of Mathematical Economics*, 45(9-10), 535–558.

Chen, Z., and L. Epstein (2002): "Ambiguity, Risk, and Asset Returns in Continuous Time," *Econometrica*, 70(4), 1403–1443.

Cieslak, A., A. Morse, and A. Vissing-Jorgensen (2015): "Stock Returns over the FOMC Cycle," Working paper, UC Berkeley.

Constantinides, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *The Journal of Political Economy*, 98(3), 519–543.

Duffie, D. (2010): "Presidential Address: Asset Price Dynamics with Slow-Moving Capital," *The Journal of Finance*, 65(4), 1237–1267.

Duffie, D., and L. G. Epstein (1992): "Stochastic Differential Utility," *Econometrica*, 60(2), 353–394.

Dunn, K., and K. Singleton (1986): "Modeling The Term Structure of Interest Rates under Non-Separable Utility and Durability of Goods," *Journal of Financial Economics*, 17, 27–55.

Echenique, F., and K. Saito (2015): "Savage in the Market," *Econometrica*, 83(4).

Epstein, L., and M. Schneider (2003): "Recursive multiple-priors," *Journal of Economic Theory*, 113(1), 1–31.

——— (2010): "Ambiguity and Asset Markets," *Annual Reviews of Financial Markets*, 2, 315–334.

Epstein, L., and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57(4), 937–969.

Epstein, L. G. (2000): "Are Probabilities Used in Markets," *Journal of Economic Theory*, 91(1), 1469–1479.

Ergin, H., and F. Gul (2009): "A Theory of Subjective Compound Lotteries," *Journal of Economic Theory*, 144, 899–929.

Farhi, E., and I. Werning (2008): "Optimal Savings Distortions with Recursive Preferences," *Journal of Monetary Economics, Carnegie Rochester Series*, 55(1), 21–42.

Gallant, A. R., L. P. Hansen, and G. Tauchen (1990): "Using Conditional Moments of Asset Payoffs to Infer The Volatility of Intertemporal Marginal Rates of Substitution," *Journal of Econometrics*, 45, 141–179.

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18(2), 141–153.

Green, R. C., and S. Srivastava (1986): "Expected utility maximization and demand behavior," *Journal of Economic Theory*, 38(2), 313–323.

Gul, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59(3), 667–686.

Hall, R. E. (1978): "Stochastic Implications of the Life Cycle–Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86(6), 971–987.

Hansen, L. P., J. C. Heaton, and N. Li (2008): "Consumption strikes back? Measuring long-run risk," *Journal of Political Economy*, 116(2), 260–302.

Hansen, L. P., and R. Jagannathan (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99(2), 225–262.

Hansen, L. P., and T. Sargent (2008): *Robustness*. Princeton, New Jersey: Princeton University Press.

Hansen, L. P., and T. 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(2), 60–66.

——— (2005): "Robust Estimation and Control under Commitment," *Journal of Economic Theory*, 124(2), 258–301.

——— (2007): "Recursive Robust Estimation and Control without Commitment," *Journal of Economic Theory*, 136(1), 1–27.

——— (2012): "Three Types of Amniguity," Working paper, New York University.

Hayashi, T., and J. Miao (2011): "Intertemporal Substitution and Recursive Smooth Ambiguity Preferences," *Theoretical Economics*, 121(3), 496–520.

Heaton, J. (1993): "The Interaction between Time-Nonseparable Preferences and Time Aggregation," *Econometrica*, 61(2), 353–385.

Ju, N., and J. Miao (2012): "Ambiguity, Learning, and Asset Returns," *Econometrica*, 80(2), 559–591.

Karantounias, A. G. (2013): "Managing pessimistic expectations and fiscal policy," *Theoretical Economics*, 8(1), 193–231.

——— (2015): "Optimal Fiscal Policy with Recursive Preferences," *Federal Reserve Bank of Atlanta, Working Paper 2013-7.*

Kechris, A. S. (1995): *Classical Descriptive Set Theory.* Springer-Verlag, NY, Inc.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73(6), 1849–1892.

——— (2009): "Recursive Smooth Ambiguity Preferences," *Journal of Economic Theory*, 144(3), 930–976.

Kreps, D. M., and E. L. Porteus (1978): "Temporal Resolution of Uncertainty. and Dynamic Choice Theory," *Econometrica*, 46(1), 185–200.

Kubler, F., L. Selden, and X. Wei (2014): "Asset Demand Based Tests of Expected Utility Maximization," *American Economic Review*, 104(11), 3459–3480.

Kwon, H., and J. Miao (2013a): "Three Types of Robust Ramsey Problem in a Linear-Quadratic Framework," *Working paper, Boston University.*

——— (2013b): "Woodford's Approach to Robust Policy Analysis in a Linear-Quadratic Framework," *Working paper, Boston University.*

Lucca, D. O., and E. Moench (2015): "The Pre-FOMC Announcement Drift," *The Journal of Finance*, LXX(1).

Luenberger, D. G. (1997): *Optimization by Vector Space Methods.* Wiley-Interscience; 1969 edition.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006a): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74(6), 1447–1498.

——— (2006b): “Dynamic variational preferences,” *Journal of Economic Theory*, 128(1), 4–44.

Machina, M. J. (1982): “”Expected Utility” Analysis without the Independence Axiom,” *Econometrica*, 50(2), 277–323.

Marshall, A. W., B. C. Arnold, and I. Olkin (2011): *Inequalities: Theory of Majorization and Its Applications*. Springer.

Milgrom, P., and I. Segal (2002): “Envelope Theorems for Arbitrary Choice Sets,” *Econometrica*, 70(2), 583–601.

Mueller, P., A. Tahbaz-Salehi, and A. Vedolin (2017): “Exchange Rates andMonetary Policy Uncertainty,” *Journal of Finance*, Accepted Author Manuscript. doi:10.1111/jofi.12499.

Muller, A., and D. Stoyan (2002): *Comparison Methods for Stochastic Models and Risks*. Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd.

Parker, J. A., and C. Julliard (2005): “Consumption Risk and the Cross Section of Expected Returns,” *Journal of Political Economy*, 113(1), 185–222.

Piazzesi, M., and M. Schneider (2006): “Equilibrium Yield Curves,” *NBER Macroeconomics Annual*.

Rothschild, M., and J. E. Stiglitz (1970): “Increasing Risk: A Definition,” *Journal of Economic Theory*, (2), 225–243.

Routledge, B. R., and S. E. Zin (2010): “Generalized Disappointment Aversion and Asset Prices,” *The Journal of Finance*, 65(4), 1303–1332.

Savor, P., and M. Wilson (2013): “How Much Do Investors Care About Macroeconomic Risk? Evidence from Scheduled Economic Announcement,” *Journal of Financial and Quantitative Analysis*, 48(2), 343–375.

Skiadas, C. (2009): *Asset pricing theory*. Princeton University Press.

——— (2013): “Smooth Ambiguity Aversion toward Small Risks and Continuous-Time Recursive Utility,” *Journal of Political Economy*, 121(4).

Strzalecki, T. (2011): “Axiomatic Foundations of Multiplier Preferences,” *Econometrica*, 79(1), 47–73.

——— (2013): “Temporal resolution of uncertainty and recursive models ambiguity aversion,” *Econometrica*, 81(3), 1039–1074.

Tapia, R. A. (1971): "The Differentiation and Integration of Nonlinear Operators," in *Nonlinear Functional Analysis and Applications*, ed. by L.B.Rall, pp. 45–101. Academic Press, New York.

Wang, T. (1993): "Lp-Fréchet Differentiable Preference and "Local Utility" Analysis," *Journal of Economic Theory*, 61(1), 139–159.

Werner, J. (2009): "Risk and Risk Aversion When States of Nature Matter," *Economic Theory*, 41, 231–246.

Woodford, M. (2010): "Robustly Optimal Monetary Policy with Near-Rational Expectations," *American Economic Review*, 100(1), 274–303.

Table I

Market Return on Announcement and Non-announcement Days

	# days p. a.	daily prem.	daily std.	premium p.a.
Market	252	2.46 *bps*	98.2 *bps*	6.19%
Announcement	30	11.21 *bps*	113.8 *bps*	3.36%
No Announcement	222	1.27 *bps*	95.9 *bps*	2.82%

This table documents the mean and the standard deviation of the market excess return during the 1961-2014 period. The column "# days p.a." is the average number of trading days per annum, the second column is the daily market equity premium on all days, that on announcement days, and that on days with no announcement. The column "daily std." is the standard deviation of daily returns. The column "premium p.a." is the cumulative market excess returns within a year, which is computed by multiplying the daily premium by the number of event days and converting it into percentage points.

Table II

Average Daily Return around Announcements (Basis Points)

	-1	0	+1
All Announcements	1.77	11.21	0.84
	(2.86)	(2.96)	(3.22)
All w/o FOMC	0.69	9.28	0.99
	(2.78)	(3.05)	(3.24)
No Announcement	− − −	1.27	− − −
		(0.91)	

This table documents the average daily return during the 1961-2014 period in basis points on event days (column "0"), that before event days (column "-1"), and that after event days (column "+1") with standard errors of the point estimates in parenthesis. The row "All announcements" is the average event day return on all announcement days; "All w/o FOMC" is the average event day return on all announcement days except FOMC announcement days; and "No announcements" is the average daily return on non-announcement days.

Table III
Average hourly return around announcements

Announcement window	−5	−4	−3	−2	−1	0	+1	+2
All Announcements	0.78	3.25	2.00	−0.17	−1.51	6.16	−2.32	2.11
	(0.26)	(2.34)	(1.85)	(0.02)	(−1.64)	(1.64)	(−1.24)	(0.90)
FOMC	13.35	13.54	7.65	3.37	4.78	0.19	5.84	−5.1
	(2.43)	(2.45)	(3.08)	(1.43)	(2.92)	(0.20)	(0.82)	(−1.08)
All w/o FOMC	−0.37	0.42	0.94	−0.69	−2.96	6.88	−3.22	2.72
	(−0.16)	(0.72)	(0.37)	(−0.30)	(−2.53)	(1.26)	(−1.43)	(2.56)

This table reports the average hourly excess return around announcements during the 1997-2013 period, with standard errors of the point estimates in parenthesis. The announcement time is normalized as hour zero. For $k = -5,\ -4, \cdots,\ 0,\ +1,\ +2$, announcement window k stands for the interval between hour $k-1$ and hour k. The row "All announcements" is the average hourly return on all announcement days; "FOMC" is the average hourly return on FOMC announcement days, and "All w/o FOMC" is the average hourly return on all announcement days except FOMC announcement days.

Appendix: Risk Preferences and The Macro Announcement Premium

Hengjie Ai and Ravi Bansal

A The two-period model

In this section, we provide a formal derivation of the A-SDF in the two-period model. We also establish the equivalence between Arrow-Debreu markets and sequential markets in the context of our model. We show that both formulations lead to the same set of asset pricing equations.

A.1 The Arrow-Debreu market

We use $\left\{\bar{C}_0, \left\{\bar{C}_1(s)\right\}_{s=1}^N\right\}$ to denote aggregate endowment in our two-period model and use $\left\{C_0(s),\ C_1(s)\right\}_{s=1}^N$ as the consumption choice of the agent. From an individual agent's perspective, the decision for C_0 is made after the announcement, and therefore can depend on s. Trading on the Arrow-Debreu market happens in period 0^-. Let $q_0(s)$ be the period 0^- price of an Arrow-Debreu security that delivers one unit of consumption good in period 0^+ and state s, for $s = 1, 2, \cdots, N$. Similarly, let $q_1(s)$ be the Arrow-Debreu price of one unit of consumption good in period one and state s. Because markets are complete, the utility maximization problem of the representative agent can be written as:

$$\max \mathcal{I}\left[u\left(C_0(s)\right) + \beta u\left(C_1(s)\right)\right]$$

$$\textit{subject to} \quad : \quad \sum_{s=1}^{N}\left[q_0(s)\,C_0(s) + q_1(s)\,C_1(s)\right] \leq \sum_{s=1}^{N}\left[q_0(s)\,\bar{C}_0 + q_1(s)\,\bar{C}_1(s)\right]$$

In the above setup, because the announcement is made at time 0^+, from the agent's perspective, consumption at time 0^+ is allowed to depend on s, which we write as $C_0(s)$. To save notation, as in the paper, we denote $V_s = u\left(C_0(s)\right) + \beta u\left(C_1(s)\right)$. Optimality implies that,

$$q_0(s) = \lambda \frac{\partial \mathcal{I}[V]}{\partial V_s} u'\left(C_0(s)\right), \quad q_1(s) = \lambda \frac{\partial \mathcal{I}[V]}{\partial V_s} \beta u'\left(C_1(s)\right),$$

where λ is the Lagrangian multiplier of the budget constraint. In equilibrium, market clearing implies that $C_0(s) = \bar{C}_0$ for all s. If we normalize the price of one unit state-non-contingent consumption at time 0^+ to be one, that is, $\sum_{s=1}^{N} q_0(s) = 1$; then, for all s,

$$q_0(s) = \frac{\frac{\partial \mathcal{I}[V]}{\partial V_s}}{\sum_{s=1}^{N} \frac{\partial \mathcal{I}[V]}{\partial V_s}}, \tag{A.1}$$

and $\frac{q_1(s)}{q_0(s)} = \beta \frac{u'(C_1(s))}{u'(C_0)}$. That is, we can simply use ratios of marginal utilities to compute Arrow-

Debreu prices. Clearly, (A.1) implies the expression of the A-SDF in equation (12) of the paper.

A.2 The sequential market

Here, we show that the two-period version of the sequential market setup described in Section 4 leads to the same asset pricing equation, (12). In period 0^-, there is no consumption decision and the agent chooses investment in a vector of announcement returns to maximize:

$$\begin{aligned} & \max_{\{\xi_j\}_{j=1}^{J}} \mathcal{I}\left[V\left(W'\right)\right] \\ \textit{subject to} \quad : \quad & W^{'} = W - \sum_{j=1}^{J} \xi_j + \sum_{j=1}^{J} \xi_j R_{A,j}(s)\,, \quad \textit{all } s, \end{aligned} \tag{A.2}$$

where $V(W) = \{V_s(W)\}_{s=1}^{N}$ is a vector of value functions. For each s, the value function $V_s(W)$ is defined by the optimal portfolio choice problem on the post-announcement market:

$$\begin{aligned} V_s(W) \quad &= \quad \max_{C_0, C_1} u(C_0) + \beta u(C_1) \\ \textit{subject to} \quad &: \quad C_1 = (W - C_0) R_{P,s}. \end{aligned} \tag{A.3}$$

Note that $R_{P,s}$ is the return from period 0^+ to period 1 after announcement s. Because the announcement fully reveals the true state of the world, $R_{P,s}$ is a risk-free return.

The first order condition for (A.2) with respect to ξ_j implies that for any announcement returns $R_{A,j}$,

$$\sum_{s=1}^{N} \frac{\partial}{\partial V_s} \mathcal{I}\left[V\left(W'\right)\right] \frac{\partial V_s\left(W_s'\right)}{\partial W_s^{'}} \left[R_{A,j}(s) - 1\right] = 0, \tag{A.4}$$

where W_s' denote the equilibrium wealth of the agent in period 0^+ after announcement s. The envelope condition for (A.3) implies that $\frac{\partial V_s(W_s')}{\partial W_s^{'}} = u'(C_0(s)) = u'\left(\bar{C}_0\right)$, where the second equality uses the market clearing condition. As $u' > 0$, equation (A.4) implies

$$\sum_{s=1}^{N} \frac{\frac{\partial}{\partial V_s}\mathcal{I}\left[V\left(W'\right)\right]}{\sum_{s=1}^{N} \frac{\partial}{\partial V_s}\mathcal{I}\left[V\left(W'\right)\right]} R_{A,j}(s) = 1,$$

as in equation (11) of the paper.

A.3 The example of recursive utility

Here, we provide details of the computation of the A-SDF for the recursive utility in Section 3.2 of the paper. We illustrate that because the announcement in our model leads uncertainty to resolve before the realization of consumption shocks, the computation of utilities and therefore, marginal utilities differ from that in models in which resolution of uncertainty happens at the same time of the realization of the consumption shocks.

Figure 1: **Early and Late Resolution of Uncertainty**

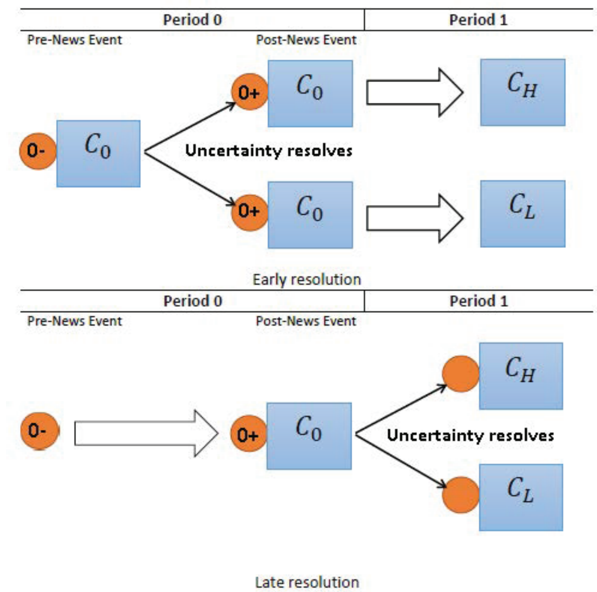


Figure 1 plots a consumption plan with early resolution of uncertainty (top panel) and a consumption plan with late resolution of uncertainty.

Figure 1 illustrates a two-period model with announcement and one without announcement. The top panel is the same as that in Figure 2 in our main text, where the announcement at time 0^+ fully reveals the true state and leads to early resolution of uncertainty. In the bottom panel of Figure 1, due to the absence of announcement, the uncertainty is resolved in period 1 when consumption is realized; that is, it is a case of late resolution of uncertainty.[1]

We denote the utility at 0^- in the case of early resolution as $V^E\left(C_0, \{C_1(s)\}_{s=1}^n\right)$. In the case of late resolution, 0^- and 0^+ have the same utility level, which we denote as $V^L\left(C_0, \{C_1(s)\}_{s=1}^n\right)$. In the case of early resolution, because there is no uncertainty in period 0^+, we first aggregate over time to compute the continuation utility as $\frac{1}{1-\frac{1}{\psi}} C_0^{1-\frac{1}{\psi}} + \beta \frac{1}{1-\frac{1}{\psi}} C_1^{1-\frac{1}{\psi}}$, and then aggregate over uncertain realizations of the announcement to compute its certainty equivalent at 0^- as:

$$V^E = \left\{ \sum_{s=1}^{n} \pi(s) \left[\left\{ C_0^{1-\frac{1}{\psi}} + \beta C_1^{1-\frac{1}{\psi}}(s) \right\}^{\frac{1-\gamma}{1-1/\psi}} \right] \right\}^{\frac{1}{1-\gamma}}. \tag{A.5}$$

In the case of late resolution, we first aggregate over uncertain period 1 consumption to compute

[1]The comparison between early and late resolution of uncertainty here is the same as that in Figure 2 of Kreps and Porteus [14]. Our top panel corresponds to node $d_0(a)$ and the bottom panel corresponds to node $d_0(b)$ in that Figure.

its certainty equivalent: $\left\{E\left[C_1^{1-\gamma}\right]\right\}^{\frac{1}{1-\gamma}}$, and then aggregate over time to compute V^L as:

$$V^L = \left\{\frac{1}{1-\frac{1}{\psi}} C_0^{1-\frac{1}{\psi}} + \beta \frac{1}{1-\frac{1}{\psi}} \left\{\sum_{s=1}^{n} \pi(s)\left[C_1^{1-\gamma}(s)\right]\right\}^{\frac{1-1/\psi}{1-\gamma}}\right\}^{\frac{1}{1-1/\psi}}. \tag{A.6}$$

The Arrow-Debreu price for one unit of consumption in period one measured in period-0 consumption numeraire can be computed as follows.[2] In the case of early resolution, the marginal rate of substitution between $C_1(s)$ and C_0 is:

$$\frac{\frac{\partial V^E}{\partial C_1(s)}}{\frac{\partial V^E}{\partial C_0}} = \pi(s)\beta\left(\frac{C_1}{C_0}\right)^{-\frac{1}{\psi}} \frac{\left\{C_0^{1-\frac{1}{\psi}} + \beta C_1^{1-\frac{1}{\psi}}(s)\right\}^{\frac{1/\psi-\gamma}{1-1/\psi}}}{\sum_{s=1}^{n} \pi(s)\left[\left\{C_0^{1-\frac{1}{\psi}} + \beta C_1^{1-\frac{1}{\psi}}(s)\right\}^{\frac{1/\psi-\gamma}{1-1/\psi}}\right]}. \tag{A.7}$$

In the case of late resolution,

$$\frac{\frac{\partial V^L}{\partial C_1(s)}}{\frac{\partial VL}{\partial C_0}} = \pi(s)\beta\left(\frac{C_1(s)}{C_0}\right)^{-\frac{1}{\psi}} \left\{\frac{C_1(s)}{\left[\sum_{s=1}^{n} \pi(s) C_1^{1-\gamma}(s)\right]^{\frac{1}{1-\gamma}}}\right\}^{\frac{1}{\psi}-\gamma}. \tag{A.8}$$

Clearly, the SDF for the early resolution case, (A.7) can be decomposed into the m^* in equation (10) and an SDF that discounts period 1 cash flow into period 0^+ consumption units: $\beta\left(\frac{C_1}{C_0}\right)^{-\frac{1}{\psi}}$. The SDF in (A.8) takes a familiar from as in many consumption-based asset pricing models where uncertainty is assumed to resolve at the same time of the realization of consumption shocks. In general, the term $\left\{\frac{C_1(s)}{\left[\sum_{s=1}^{n} \pi(s) C_1^{1-\gamma}(s)\right]^{\frac{1}{1-\gamma}}}\right\}^{\frac{1}{\psi}-\gamma}$ does not integrate to one unless in the special case of unit IES.

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 (14). We also provide an expression for the implied A-SDF.[3]

- The recursive utility of Kreps and Porteus [14] and Epstein and Zin [5]. The recursive

[2] As in the paper, in the case with announcement, the period-0 consumption numeraire is interpreted as one unit of period-0 consumption delivered non-contingently at time 0^+.

[3] Depending on the model, additional conditions may be needed so that the assumptions of Theorem 1 can be verified. We provide the expressions for A-SDF assuming appropriate conditions on the primitive utility functions can be imposed to guarantee its existence.

preference can be generally represented as:

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

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,

$$V = u\,(U)\,, \tag{B.2}$$

the recursive relationship for V can be written in the form of (14) with the same u function in equation (B.1) and the certainty equivalence functional:

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

where $\phi = h \circ u^{-1}$. the A-SDF can be written as:

$$m^*\,(V) \propto \phi'\,(V)\,, \tag{B.3}$$

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

- The maxmin expected utility of Gilboa and Schmeidler [7]. The dynamic version of this preference is studied in Epstein and Schneider [4] and Chen and Epstein [2]. This preference can be represented as the special case of (14) 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 convex and closed in the weak* topology. As we show in Section 3.2 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 [17], the dynamic version of which is studied in Maccheroni, Marinacci, and Rustichini [18], 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 [8] and Strzalecki [22] 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 [6] can be written as (14) 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 preference of Klibanoff, Marinacci, and Mukerji [12] and Klibanoff, Marinacci, and Mukerji [13] can be represented as:

$$\mathcal{I}(V) = \phi^{-1}\left(\int_M \phi\left(\int_\Omega mV\,dP\right) d\mu(m)\right), \tag{B.4}$$

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

$$m^*(\omega) \propto \int_M \phi'\left(\int mV\,dP\right) m(\omega)\, d\mu(m). \tag{B.5}$$

- The certainty equivalence functional $\mathcal{I}$ for the disappointment aversion preference is implicitly defined as $\mathcal{I}[V] = \mu$, 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 concave function. The A-SDF can be written as:

$$m^*(V) = \begin{cases} \frac{\phi'(V)}{\phi'(\mu)[1+\theta P(V\leq\mu)]} & if\ V > \mu \\ \frac{(1-\theta)\phi'(V)}{\phi'(\mu)[1+\theta P(V\leq\mu)]} & f\ V \leq \mu \end{cases},$$

whenever $\mathcal{I}[V]$ is differentiable at V.

- Hayashi and Miao [10] develop a class of generalized recursive smooth ambiguity model that takes the following form:[4]

$$\bar{V}_t = u^{-1}\left\{(1-\beta)\, u(C_t) + \beta\left[u \circ \nu^{-1}\right]\left(\int_M \left[\nu \circ \phi^{-1}\right]\left(\int m\phi\left(\bar{V}_{t+1}\right) dP\right) d\mu(m)\right)\right\}, \tag{B.6}$$

where u, v, and ϕ are all smooth and monotone functions. As in the Klibanoff, Marinacci,

[4] The model in Hayashi and Miao [10] is more general than (B.6) and may not permit a representation of the form $V_t = u(C_t) + \beta\mathcal{I}[V_{t+1}]$. However, the applied examples of this preference are often special cases of (B.6). See also the related generalized recursive multiple-priors model of Hayashi [9], which can be obtained as a limiting case of (B.6).

and Mukerji [12] model, M is a set of probability densities that represent ambiguous beliefs, and μ is a measure on the set of densities. With a monotonic transformation, $V_t = u\left(\bar{V}_t\right)$, the above can be written in the form of (14) with

$$\mathcal{I}(V) = \left[u \circ \nu^{-1}\right]\left(\int_M \left[\nu \circ \phi^{-1}\right]\left(\int m\left[\phi \circ u^{-1}\right](V)\, dP\right) d\mu(m)\right).$$

The A-SDF for this class of preferences can be written as:

$$m^*(\omega) \propto \int_M \left[\nu \circ \phi^{-1}\right]'\left(\int m\left[\phi^{-1} \circ u\right] V dP\right) m(\omega)\left[\phi \circ u^{-1}\right]'(V(\omega))\, d\mu(m).$$

C Proofs for Theorem 1 and 2

C.1 Preliminaries

In this section, we provide formal definitions of some relevant concepts and introduce the basic methodology to prove Theorem 1 and 2. We first state the definition of a non-atomic probability space, which is an assumption maintained throughout Section 4.

Definition A.1. Non-atomic probability space: *A probability space* $(\Omega, \mathcal{F}, P)$ *is said to be non-atomic (or continuous) if* $\forall \omega \in \Omega,\ P(\omega) = 0$.

Next, we state the definition of first-order stochastic dominance (FSD) and second-order stochastic dominance (SSD).

Definition A.2. First-order stochastic dominance: X_1 *first-order stochastic dominates* X_2*, or* $X_1 \geq_{FSD} X_2$*, if there exists 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 A.3. Second-order stochastic dominance: X_1 *second order stochastic dominates* X_2*, or* $X_1 \geq_{SSD} X_2$*, if there exists 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.*[5]

FSD and SSD are typically defined as stochastic orders on the space of distributions. Here, it is more convenient to define FSD and SSD as binary relations on the space of random variables. Our definitions are equivalent to the standard definitions of FSD and SSD due to the assumption of a non-atomic probability space (See Muller and Stoyan [20]).

Our strategy for proving Theorem 1 and 2 consists of two steps. First, we apply the envelope theorems in Milgrom and Segal [19] to establish the differentiability of the value functions. Second,

[5] Our definition of SSD is the same as the standard concept of increasing risk (see Rothschild and Stiglitz [21] and Werner [25]). 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.

we 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.[6]

Most of our analysis below is on the conditional certainty equivalence functional $\mathcal{I}[\cdot|\, z]$. To save notation, whenever it does not cause confusion, we suppress the dependence of $\mathcal{I}[\cdot|\, z]$ on z and simply write $\mathcal{I}[\cdot]$. We often use the following operation to relate the certainty equivalence functional $\mathcal{I}$ 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

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

where $D\mathcal{I}[X + t(Y-X)]$ is understood as the representation of the Fréchet derivative of $\mathcal{I}[\cdot]$ evaluated at $X + t(Y-X)$. The Riesz representation theorem implies that $D\mathcal{I}[X + t(Y-X)]$ is an element of $L^2(\Omega, \mathcal{F}, P)$, and $D\mathcal{I}[X + t(Y-X)]$ applied to $(Y-X)$ can be computed as the dot product, $\int_\Omega D\mathcal{I}[X + t(Y-X)](Y-X)\, dP$.

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 [16]). 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 of $D\mathcal{I}$ 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 (C.1).

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

$$\mathcal{I}[Y] - \mathcal{I}[X] = \int_\Omega D\mathcal{I}\left[X + \hat{t}(Y-X)\right](Y-X)\, dP. \tag{C.2}$$

Much of our analysis below relies on the theory of differentiability for nonlinear operators defined on infinite dimensional spaces, for example, in Tapia [24] and Luenberger [16].

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

C.2 Existence of A-SDF

In this section, we provide a proof for Theorem 1 and establish the existence of A-SDF.

Differentiability of value function We establish the differentiability of value functions recursively. In particular, we show that the value functions are elements of $\mathcal{D}$, which is defined as:

Definition A.4. *$\mathcal{D}$ is the set of differentiable functions on the real line, denoted by f, that satisfy the following two properties.*

1. *f is Lipschitz continuous and $f'(x) > 0$.*
2. *$\forall x \in \mathbf{R}$, as $h \to 0$, $\frac{1}{h}[f(x+h-a) - f(x-a)]$ converges uniformly to $f'(x-a)$ in a. That is, $\forall \varepsilon > 0$, there exists $\delta > 0$ such that $|h| < \delta$ implies that $\left|\frac{1}{h}[f(x+h-a) - f(x-a)] - f'(x-a)\right| < \varepsilon$ for all $a \in \mathbf{R}$.*

We first introduce some notation. 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, by:

$$f_v(W, \xi) = u\left(W - \sum_{j=0}^{J} \xi_j\right) + \beta \mathcal{I}\left[v\left(\sum_{j=0}^{J} \xi_j R_j\right)\right], \tag{C.3}$$

$$g_v(W, \xi) = \mathcal{I}\left[v\left(W + \sum_{j=0}^{J} \xi_j (R_j - 1)\right)\right]. \tag{C.4}$$

Because $R_j \in L^2(\Omega, \mathcal{F}, P)$ and v is Lipschitz continuous, for a fixed ξ, $v\left(\sum_{j=0}^{J} \xi_j R_j\right)$ and $v\left(W - \sum_{j=0}^{J} \xi_j (R_j - 1)\right)$ are both square integrable and equations (C.3) and (C.4) are well-defined.

We define two operators on $\mathcal{D}$. For any $v \in \mathcal{D}$, let $T^+ v$ and $T^- v$ be defined by:

$$\left[T^+ v\right](W) = \sup_{\xi} f_v(W, \xi), \tag{C.5}$$

$$\left[T^- v\right](W) = \sup_{\xi} g_v(W, \xi). \tag{C.6}$$

Clearly, the value functions $V_{z_t^+}(W)$ and $V_{z_t^-}(W)$ can be constructed recursively as $V_{z_t^+}(W) = \left[T^+ V_{z_{t+1}^-}\right](W)$, and $V_{z_t^-}(W) = \left[T^- V_{z_t^+}\right](W)$ (with the understanding that the certainty equivalence functionals in the definition of $f_v(W, \xi)$ and $g_v(W, \xi)$ are appropriately chosen conditional certainty equivalence functionals). Because we start with the assumption of the existence of an interior equilibrium, the maximization problems (C.5) and (C.6) are well defined, and the maximums are achieved.

Below, we prove that $V_{z_t^+}$ and $V_{z_t^-}$ are elements of $\mathcal{D}$ in two steps. First, Lemma A.1 below establishes the equi-differentiability of the family of functions $\{f_v(W, \xi)\}_{\xi}$ and $\{g_v(W, \xi)\}_{\xi}$ so that we can apply the envelope theorem in Milgrom and Segal [19]. Second, in Lemma A.2, we apply the envelope theorem repeatedly to show that the operators T^+ and T^- map $\mathcal{D}$ into itself.

Lemma A.1. *Suppose $u, v \in \mathcal{D}$, as $h \to 0$, 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

$$\frac{1}{h}\left[g_v\left(W+h,\xi\right)-g_v\left(W,\xi\right)\right]\to\frac{\partial}{\partial W}g_v\left(W,\xi\right) \tag{C.7}$$

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

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

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 (C.2). 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 (C.7) can be written as $\int_\Omega D\mathcal{I}\left[\bar{v}\left(0\right)\right]\bar{v}'\left(0\right)dP$, we have:

$$\begin{aligned} & \left|\frac{1}{h}\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\left(\bar{v}\left(1\right)-\bar{v}\left(0\right)\right)dP-\int_\Omega D\mathcal{I}\left[\bar{v}\left(0\right)\right]\bar{v}'\left(0\right)dP\right| \\ = & \left|\begin{array}{c} \frac{1}{h}\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\left(\bar{v}\left(1\right)-\bar{v}\left(0\right)\right)dP-\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\bar{v}'\left(0\right)dP \\ +\int_\Omega D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\bar{v}'\left(0\right)dP-\int_\Omega D\mathcal{I}\left[\bar{v}\left(0\right)\right]\bar{v}'\left(0\right)dP \end{array}\right| \\ \leq & \int_\Omega\left|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\right|\left|\frac{1}{h}\left(\bar{v}\left(1\right)-\bar{v}\left(0\right)\right)-\bar{v}'\left(0\right)\right|dP+\int_\Omega\left|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]-D\mathcal{I}\left[\bar{v}\left(0\right)\right]\right|\left|\bar{v}'\left(0\right)\right|dP \\ \leq & \left\|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]\right\|\left\|\frac{1}{h}\left(\bar{v}\left(1\right)-\bar{v}\left(0\right)\right)-\bar{v}'\left(0\right)\right\|+\left\|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]-D\mathcal{I}\left[\bar{v}\left(0\right)\right]\right\|\left\|\bar{v}'\left(0\right)\right\| \end{aligned} \tag{C.8}$$

Because $v\in\mathcal{D}$, for h small enough, $\left|\frac{1}{h}\left(\bar{v}\left(1\right)-\bar{v}\left(0\right)\right)-\bar{v}'\left(0\right)\right|\leq\varepsilon$ with probability one and $\left\|\frac{1}{h}\left(\bar{v}\left(1\right)-\bar{v}\left(0\right)\right)-\bar{v}'\left(0\right)\right\|\leq\varepsilon$. Also, because $D\mathcal{I}$ is Lipschitz continuous, $\left\|D\mathcal{I}\left[\bar{v}\left(\hat{t}\right)\right]-D\mathcal{I}\left[\bar{v}\left(0\right)\right]\right\|\leq K\left\|\bar{v}\left(1\right)-\bar{v}\left(0\right)\right\|\leq K^2h$, where the second inequality is due to the Lipschitz continuity of v. This proves the uniform convergence of (C.8).

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

Proof: *It follows from Lemma A.1 that for any $v \in \mathcal{D}$, we can apply Theorem 3 in Milgrom and Segal [19] and establish that both $T^{+}v$ and $T^{-}v$ are differentiable, and*

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

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

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

$$f_{v}\left(W_{1},\xi\left(W_{2}\right)\right)-f_{v}\left(W_{2},\xi\left(W_{2}\right)\right)\leq T^{+}v\left(W_{1}\right)-T^{+}v\left(W_{2}\right)\leq f_{v}\left(W_{1},\xi\left(W_{1}\right)\right)-f_{v}\left(W_{2},\xi\left(W_{1}\right)\right). \tag{C.9}$$

Because $\forall\xi$, $\left|f\left(W_{1},\xi\right)-f\left(W_{2},\xi\right)\right| = \left|u\left(W_{1}-\sum_{j=0}^{J}\xi_{j}\right)-u\left(W_{2}-\sum_{j=0}^{J}\xi_{j}\right)\right| \leq K\left|W_{1}-W_{2}\right|$, where K is a Lipschitz constant for u, $\left|Tv\left(W_{1}\right)-Tv\left(W_{2}\right)\right|\leq K\left|W_{1}-W_{2}\right|$. We can prove that $T^{-}v\left(W\right)$ is Lipschitz continuous in a similar way:

$$g_{v}\left(W_{1},\xi\left(W_{2}\right)\right)-g_{v}\left(W_{2},\xi\left(W_{2}\right)\right)\leq T^{-}v\left(W_{1}\right)-T^{-}v\left(W_{2}\right)\leq g_{v}\left(W_{1},\xi\left(W_{1}\right)\right)-g_{v}\left(W_{2},\xi\left(W_{1}\right)\right). \tag{C.10}$$

Note that $\forall\xi$,

$$\begin{aligned} \left|g_{v}\left(W_{1},\xi\right)-g_{v}\left(W_{2},\xi\right)\right| &= \left|\mathcal{I}\left[v\left(W_{1}+\sum\nolimits_{j=0}^{J}\xi_{j}\left(R_{j}-1\right)\right)\right]-\mathcal{I}\left[v\left(W_{2}+\sum\nolimits_{j=0}^{J}\xi_{j}\left(R_{j}-1\right)\right)\right]\right| \\ &\leq K\left\|v\left(W_{1}+\sum\nolimits_{j=0}^{J}\xi_{j}\left(R_{j}-1\right)\right)-v\left(W_{2}+\sum\nolimits_{j=0}^{J}\xi_{j}\left(R_{j}-1\right)\right)\right\| \\ &\leq K^{2}\left|W_{1}-W_{2}\right|, \end{aligned}$$

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

In addition, equations (C.9) and (C.10) can be used to show that the family of functions $\left\{T^{+}v\left(W-a\right)\right\}_{a}$ and $\left\{T^{-}v\left(W-a\right)\right\}_{a}$ are equi-differentiable. For example, let $W_{1}\to W_{2}$,

$$\frac{1}{W_{1}-W_{2}}\left[f_{v}\left(W_{1},\xi\right)-f_{v}\left(W_{2},\xi\right)\right]$$

converges uniformly by Lemma A.1, and by equation (C.9), $\frac{1}{W_{1}-W_{2}}\left[T^{+}v\left(W_{1}\right)-T^{+}v\left(W_{2}\right)\right]$ must also converge uniformly.

Finally, we note that if $v'\left(x\right)>0$ for all $x\in\mathbf{R}$, then $\left[T^{+}v\right]\left(W\right)$ and $\left[T^{-}v\right]\left(W\right)$ must satisfy the same property by the envelope theorem.

Proof of Theorem 1 In this section, we establish the existence of SDF as stated in Theorem 1. To save notation, whenever convenient, we denote $R_{j}\left(z\right)$ to be the one-period return of asset j that payoff at history z. That is, if $z=z_{t}^{+}=\left(z_{t}^{-},s_{t}^{+}\right)$ is a post-announcement history, then

$R_j(z) \equiv R_{A,j}\left(s_t^+ \middle| z_t^-\right)$ is an announcement return, and if z is of the form $z = z_{t+1}^- = \left(z_t^+, s_{t-1}^-\right)$, then $R_j(z) \equiv R_{P,j}\left(s_{t+1}^- \middle| z_t^+\right)$ is a post-announcement return. We write the portfolio selection problem at z_t^- as

$$\max_{\zeta} \mathcal{I}\left[V_{z_t^+}\left(W + \sum_{j=0}^{J} \zeta_j \left(R_j\left(z_t^+\right) - 1\right)\right) \middle| z_t^-\right]. \tag{C.11}$$

Clearly, no arbitrage implies that the risk-free announcement return $R_0\left(z_t^+\right) = 1$. Because $V_{z_t^+}$ and $\mathcal{I}\left[\cdot \middle| z_t^-\right]$ are (Fréchet) differentiable, $\mathcal{I}\left[V_{z_t^+}\left(W + \sum_{j=0}^{J} \zeta_j \left(R_j\left(z_t^+\right) - 1\right)\right) \middle| z_t^-\right]$ is differentiable in ζ.[7] Therefore, the first order condition with respect to ζ_j implies that

$$E\left[D\mathcal{I}\left[V_{z_t^+}\left(W'\right)\right] \frac{d}{dW} V_{z_t^+}\left(W'\right)\left(R_j\left(z_t^+\right) - 1\right) \middle| z_t^-\right] = 0, \tag{C.12}$$

where we denote $W' = W + \sum_{j=0}^{J} \hat{\zeta}_j \left(R_j\left(z_t^+\right) - 1\right)$ and $\hat{\zeta}$ is the optimal portfolio choice.

The value function $V_{z_t^+}(W)$ in (C.11) is determined by the the agent's portfolio choice problem at z_t^+ after the announcement s_t^+ is made:

$$V_{z_t^+}(W) = \max_{\xi}\left\{u\left(W - \sum_{j=0}^{J} \xi_j\right) + \beta \mathcal{I}\left[V_{z_{t+1}^-}\left(\sum_{j=0}^{J} \xi_j R_j\left(z_{t-1}^-\right)\right) \middle| z_t^+\right]\right\}. \tag{C.13}$$

The envelop condition for (C.13) implies

$$\frac{d}{dW} V_{z_t^+}(W) = u'\left(W - \sum_{j=0}^{J} \xi_j\right) = u'\left(C_t\right) = u'\left(\bar{C}_t\right),$$

where the last equality uses the market clearing condition. Because consumption at time t must equal to total endowment, $\bar{C}_t$, and because $\bar{C}_t$ must be z_t^- measurable, so must $\frac{d}{dW} V_{z_t^+}(W)$.

By our results in Section C.2, $\frac{d}{dW} V_{z_t^+}(W) = u'\left(\bar{C}_t\right) > 0$. Because $\frac{d}{dW} V_{z_t^+}(W)$ is z_t^- measurable, (C.12) implies:

$$E\left[D\mathcal{I}\left[V_{z_t^+}(W)\right]\left(R_j\left(z_t^+\right) - 1\right) \middle| z_t^-\right] = 0. \tag{C.14}$$

As we show in Lemma A.4 in the next section, monotonicity of $\mathcal{I}$ guarantees that $D\mathcal{I} \geq 0$ with probability one. To derive an expression for A-SDF, we need to assume a slightly stronger condition:

$$D\mathcal{I}[X] > 0 \text{ with strictly positive probability for all } X.^{8} \tag{C.15}$$

In this case, the A-SDF can be constructed as:

$$m^*\left(s_t^+ \middle| z_t^-\right) = \frac{D\mathcal{I}\left[V_{z_t^+}\left(W_{z_t^-, s_t^+}\right)\right]}{E\left[D\mathcal{I}\left[V_{z_t^+}\left(W_{z_t^-, s_t^+}\right)\right] \middle| z_t^-\right]}, \tag{C.16}$$

[7] This is a version of the chain rule. See Proposition 1 in Chapter 7 of Luenberger [16].

[8] Note that monotonicity with respect to FSD implies that $D\mathcal{I}[X] \geq 0$ with probability one for all X. If condition (C.15) 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 (C.15), our result implies that the A-SDF exists generically.

where W_z denote the equilibrium wealth of the agent at history z. Because $E\left[m^*\left(s_t^+\middle|\, z_t^-\right)\middle|\, z_t^-\right]=1$, we can write (C.14) as an asset pricing equation with A-SDF:

$$E\left[m^*\left(\cdot|\, z_t^-\right)R_{A,j}\left(\cdot|\, z_t^-\right)\middle|\, z_t^-\right]=1.$$

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 1, we only need to show that $m^*\left(s_t^+\middle|\, z_t^-\right)$ can be represented as a measurable function of continuation utility: $m^*\left(s_t^+\middle|\, z_t^-\right)=m^*\circ V_{z_t^+}\left(W_{z_t^-,s_t^+}\right)$ for some measurable function $m^*:\mathbf{R}\to\mathbf{R}$.[9] That is, $m^*\left(s_t^+\middle|\, z_t^-\right)$ 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 that $m^*\left(s_t^+\middle|\, z_t^-\right)$ is measurable with respect to the σ-field generated by $V_{z_t^+}\left(W_{z_t^-,s_t^+}\right)$.

Lemma A.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 $X\in L^2(\Omega,\mathcal{F},P)$, 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 (The assumption of a non-atomic probability space guarantees the existence of such measure-preserving transformations. See exercise 17.43 in Kechris [11]). Let $D\mathcal{I}[X]$ be the $L^2(\Omega,\mathcal{F},P)$ representation of the Fréchet derivative of the certainty equivalence functional $\mathcal{I}$ at X. Below, we first show that $D\mathcal{I}[X]\circ T$ must also be a Fréchet derivative of $\mathcal{I}$ at X. Because the Fréchet derivative is unique, we must have $D\mathcal{I}[X]=D\mathcal{I}[X]\circ T$ with probability one; therefore, $D\mathcal{I}[X]$ must be measurable with respect to the invariant σ-field of T and therefore, also measurable with respect to $\sigma(X)$.*

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

$$\lim_{\alpha\to 0}\frac{1}{\alpha}\left[V(X+\alpha Y)-V(X)\right]=\int\left(D\mathcal{I}[X]\circ T\right)\cdot Y dP \tag{C.17}$$

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

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 $[X\circ T+\alpha Y]\circ T^{-1}=$

[9] In general, m^* may depend on z_t^-. Here, with a slight abuse of notation, we denote m^* both as the A-SDF, which is an element of L^2, and as a measurable function $\mathbf{R}\to\mathbf{R}$.

$X + \alpha Y \circ T^{-1}$ has the same distribution with $X \circ T + \alpha Y$. As a result,

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

where the last equality uses the fact that $\left[D\mathcal{I}\left[X\right]\cdot Y\circ T^{-1}\right]\circ T = D\mathcal{I}\left[X\right]\circ T\cdot Y$ have the same distribution with $D\mathcal{I}\left[X\right]\times Y\circ T^{-1}$. This proves (C.17).

C.3 Generalized Risk Sensitivity and the Announcement Premium

We prove Theorem 2 in this section. Part 1 is straightforward given our results in the last section. From equation (C.16), if $\mathcal{I}$ is expected utility, then $m^*\left(s_t^+\middle|\, z_t^-\right)$ must be a constant. Conversely, if $m^*\left(s_t^+\middle|\, z_t^-\right)$ is a constant, then $\mathcal{I}$ is linear and must have an expected utility representation.

We prove part 2) of theorem 2 in three steps. First, we use Lemma A.4 to establish that $m^*\left(V_{z_t^+}\right)$ is non-negative if and only if $\mathcal{I}$ is monotone with respect to FSD. Second, we prove the equivalence between (a) and (b). Lemma A.5 and A.6 jointly establish that generalized risk sensitivity of $\mathcal{I}$ is equivalent to $m^*\left(V_{z_t^+}\right)$ being a non-increasing function of $V_{z_t^+}$. Finally, we use Lemma A.7 to establish the equivalence between (b) and (c).

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

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

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

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

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

contradicting monotonicity with respect to FSD.

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

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

$$\int D\mathcal{I}\left[X\right]\cdot\left(X-E\left[X\middle|\,\mathcal{G}\right]\right)dP\leq 0. \tag{C.18}$$

Proof: *Suppose condition (C.18) 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 (C.1),

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

where the last inequality uses (C.18).

Conversely, assuming $\mathcal{I}$ *is increasing in SSD, we prove (C.18) by contradiction. if (C.18) 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.*[10]

Due to Lemma A.3, $D\mathcal{I}[X]$ can be represented by a measurable 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 $m^*\left(V_{z_t^+}\right)$, we only need to prove that condition (C.18) is equivalent to $\eta(\cdot)$ being a non-increasing function, which is Lemma A.6 below.

Lemma A.6. *Condition (C.18) is equivalent to* $\eta(X)$ *being a non-increasing function of* X.

Proof: *First, we assume* $\eta(X)$ *is non-increasing. To prove (C.18), note that* $E[X|\mathcal{G}]$ *is*

[10] 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 [21] and Werner [25]). If $E[Z|V_1] = 0$, then for any concave function ϕ, $\phi(V_1 + \lambda Z_1) \geq \lambda\phi(V_1 + Z) + (1-\lambda)\phi(V_1)$. Therefore, $E[\phi(V_1 + \lambda Z_1)] \geq \lambda E[\phi(V_1 + Z)] + (1-\lambda)E[\phi(V_1)] \geq E[\phi(V_1 + Z)]$, where the last inequality is true because $E[\phi(V_1)] \geq E[\phi(V_1 + Z_1)]$.

measurable with respect to $\sigma(X)$*, and we can use the law of iterated expectation to write:*

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

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

Second, to prove the converse of the above statement by contradiction, we assume (C.18) is true, but there exist $x_1<x_2$*, both occur with positive probability such that* $\eta(x_1)<\eta(x_2)$*. Under this assumption, we construct a random variable* Y*:*

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

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

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

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

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

Lemma A.7. *That* $m^*\left(V_{z_t^+}\right)$ *is a non-increasing function of* $V_{z_t^+}$ *is equivalent to (c).*

Proof: *If* $m^*\left(V_{z_t^+}\right)$ *is a non-decreasing function, then for any payoff* f *that is co-monotone with* $V_{z_t^+}$*, we have*

$$E\left[m^*\left(V_{z_t^+}\right)f\left(V_{z_t^+}\right)\right]\leq E\left[m^*\left(V_{z_t^+}\right)\right]E\left[f\left(V_{z_t^+}\right)\right]=E\left[f\left(V_{z_t^+}\right)\right],$$

because $m_t^*\left(V_{z_t^+}\right)$ *and* $f\left(V_{z_t^+}\right)$ *are negatively correlated.*[11]

We prove that (c) implies (b) by contradiction. Suppose that the announcement premium is non-negative for all payoffs that are co-monotone with $V_{z_t^+}$*, but* $m^*(v_1)<m^*(v_2)$ *for some* $v_1<v_2$*,*

[11] Note that the same argument implies that if $m^*\left(V_t^+\right)$ is a non-decreasing function, then the announcement premium must be non-negative for the following more general class of payoffs: $f\left(s|\,z_t^-\right)+\varepsilon$, where $E\left[\varepsilon|\,z_t^-,s\right]=0$.

both of which occur with positive probability. Consider the following payoff $f(\cdot)$:

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

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

D Generalized Risk-Sensitive Preferences

D.1 Generalized risk sensitivity and uncertainty aversion

In this section, we provide proofs for results for the relationship between generalized risk sensitivity and uncertainty aversion discussed in Section 4.3 of the paper.

- Quasiconcavity implies generalized risk sensitivity.

  The following lemma formalizes the above statement.

  Lemma A.8. *Suppose $\mathcal{I} : L^2(\Omega, \mathcal{F}, P) \to \mathbf{R}$ is continuous and 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 [11]), then quasiconcavity implies that

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

  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 $\mathcal{I}$ then implies $\mathcal{I}[X_1] \geq \mathcal{I}[X_1 + Y] = \mathcal{I}[X_2]$, that is, $\mathcal{I}$ satisfies generalized risk sensitivity. □

- Qusiconcavity is not necessary for generalized risk sensitivity.

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

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

Clearly, this condition is weaker than quasiconcavity.

Here, 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

$$\forall X\in\mathbf{X},\ \ \mathcal{I}\left[X\right]=\phi^{-1}\left\{\min_{m\in M}E\left[m\phi\left(X\right)\right]\right\},\quad with\ \ \phi\left(x\right)=e^{x}, \tag{D.2}$$

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 measures and the parameter $\eta\geq e^{B}$. Note that $\mathcal{I}$ defined in (D.2) is not concave because $\phi(x)$ is a strictly convex function. Below we show that $\mathcal{I}$ satisfies generalized risk sensitivity, but is not quasiconcavity.

Using (D.1), 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, $\mathcal{I}\left[\lambda X+(1-\lambda)Y\right]\geq\mathcal{I}\left[X\right]$. 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 $\lambda\geq\frac{1}{2}$,

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

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

$$\frac{d}{d\lambda}\phi\left(\mathcal{I}\left[\lambda X+(1-\lambda)Y\right]\right)\leq 0. \tag{D.3}$$

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

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

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

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

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 (D.3). Similarly, one can prove $\mathcal{I}[\lambda X + (1-\lambda) Y] \geq \mathcal{I}[Y]$ 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.

- For second-order expected utility, the concavity of ϕ is equivalent to generalized risk sensitivity.

  Proof. Certainty equivalence functionals of the form $\mathcal{I}[V] = \phi^{-1}(E[\phi(V)])$, where ϕ is strictly increasing is called second-order expected utility in Ergin and Gul [6]. 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 $\mathcal{I}[\cdot]$ is quasiconcave and satisfies generalized risk sensitivity by Lemma A.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 [21], ϕ is concave. □

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

  Proof. Using the results in Klibanoff, Marinacci, and Mukerji [12, 13], it straightforward to show that for the class of smooth ambiguity preferences, concavity of ϕ is equivalent to the quasiconcavity of $\mathcal{I}$. As a result, quasiconcavity implies generalized risk sensitivity by Lemma A.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 $\mathcal{I}[E[V]] \geq \mathcal{I}[V]$, for all $V \in L^2(\Omega, \mathcal{F}, P)$. That is,

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

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 detailed examples and proofs for the discussions on the relationship between preference for early resolution of uncertainty and generalized risk sensitivity in Section 4.3.

- An example that satisfies generalized risk sensitivity but strictly prefers late resolution of uncertainty.

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

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

  It straight forward to check that $\mathcal{I}$ is quasiconcave 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:

  $$V^E = \mathcal{I}\left[u(C_0) + u(C_1)\right].$$

  It is straightforward to show that:

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

  Therefore,

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

  The utility for late resolution of uncertainty is given by:

  $$V^L = u(C_0) + \mathcal{I}\left[u(C_1)\right] = 0.1025.$$

- An example of $\mathcal{I}$ that prefers early resolution of uncertainty but is strictly decreasing in second order stochastic dominance.

  Example A.2. *Consider the following preference:*

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

  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 A.1. It *is straightforward to verify that the utility for early resolution of uncertainty is*

$$V^E = \mathcal{I}\left[u\left(C_0\right) + u\left(C_1\right)\right] = 0.1485,$$

and the utility for later resolution is:

$$V^L = u\left(C_0\right) + \mathcal{I}\left[u\left(C_1\right)\right] = 0.11.$$

- Generalized risk sensitivity and indifference toward the timing of resolution of uncertainty implies representation (24).

  Proof. By Lemma 1 and the proof of Theorem 1 in Strzalecki [23], indifference between timing of resolution of uncertainty implies that $\mathcal{I}$ satisfies that for all $X \in L^2\left(\Omega, \mathcal{F}, P\right)$, all $a \geq 0$, $\mathcal{I}\left[a + X\right] = a + \mathcal{I}\left[X\right]$. Take derivatives with respect to a and evaluate at $a = 0$, we have:

$$\int D\mathcal{I}\left[X\right] dP = 1. \tag{D.4}$$

  Note that because $\mathcal{I}$ is normalized, $\mathcal{I}\left[0\right] = 0$. Therefore, $\forall X \in X \in L^2\left(\Omega, \mathcal{F}, P\right)$,

$$\begin{aligned} \mathcal{I}\left[X\right] &= \mathcal{I}\left[X\right] - \mathcal{I}\left[0\right] \\ &= \int_0^1 \int D\mathcal{I}\left[tX\right] X dP dt \\ &= \int \int_0^1 D\mathcal{I}\left[tX\right] dt X dP. \end{aligned}$$

  Note that $\int \int_0^1 D\mathcal{I}\left[tX\right] dt dP = 1$ is a density, because of (D.4). In addition, generalized risk sensitivity implies that for each t,

$$\left[D\mathcal{I}\left[tX\right]\left(\omega\right) - D\mathcal{I}\left[tX\right]\left(\omega'\right)\right]\left[X\left(\omega\right) - X\left(\omega'\right)\right] \leq 0. \tag{D.5}$$

  Therefore, $\int_0^1 D\mathcal{I}\left[tX\right] dt$ must satisfy (D.5) as well. By the result of Carlier and Dana [1], $\int \int_0^1 D\mathcal{I}\left[tX\right] dt X dP$ can be represented by minimization with respect to the core of a convex distortion of P. □

E Details of the Continuous-time model

E.1 Asset Pricing in the Learning Model

Value function of the representative agent Because announcements fully reveal the value of x_t at nT, $q_{nT}^+ = 0$. We start from $q_0 = 0$. In the interior of $(0, T)$, the standard optimal filtering implies that the posterior mean and variance of x_t are given by equations (30) and (31). Here q_t

has a closed form solution:

$$q\left(t\right)=\frac{\sigma_{x}^{2}\left(1-e^{-2\hat{a}t}\right)}{\left(\hat{a}-a\right)e^{-2\hat{a}t}+a+\hat{a}},\tag{E.1}$$

where $\hat{a}=\sqrt{a^{2}+\left(\sigma_{x}/\sigma\right)^{2}}$. In general, we can write $q_{t}=q\left(t \bmod T\right)$ for all t.

Using the results from Duffie and Epstein [3], 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 [3]), and denote $\bar{f}$ the normalized aggregator. Under this normalization, $\bar{f}(C,V)$ is:

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

Due to homogeneity, the value function is of the form

$$\bar{V}\left(\hat{x}_{t},t,C_{t}\right)=\frac{1}{1-\gamma}H\left(\hat{x}_{t},t\right)C_{t}^{1-\gamma},\tag{E.2}$$

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

$$\begin{aligned}&-\frac{\rho}{1-\gamma}\ln H\left(\hat{x},t\right)H\left(x,t\right)+\left(\hat{x}-\frac{1}{2}\gamma\sigma^{2}\right)H\left(\hat{x},t\right)+\frac{1}{1-\gamma}H_{t}\left(\hat{x}_{t},t\right)\\&+\left[\frac{1}{1-\gamma}a_{x}\left(\bar{x}-\hat{x}\right)+q_{t}\right]H_{x}\left(\hat{x},t\right)+\frac{1}{2}\frac{1}{1-\gamma}H_{xx}\left(\hat{x},t\right)\frac{q_{t}^{2}}{\sigma^{2}}=0,\end{aligned}\tag{E.3}$$

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

$$H\left(\hat{x}_{nT}^{-},nT\right)=E\left[H\left(\hat{x}_{nT}^{+},nT\right)\middle|\,\hat{x}_{nT}^{-},q_{nT}^{-}\right].\tag{E.4}$$

A monotonic transformation of $\bar{V}$, $V_{t}=\frac{1}{1-\gamma}\ln\left[\left(1-\gamma\right)\bar{V}_{t}\right]$ has the representation of (32).

The solution to the partial differential equation (PDE) (E.3) together with the boundary condition (E.4) is separable and given by:

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

where $h\left(t\right)$ satisfy the following ODE:

$$-\rho h\left(t\right)+h'\left(t\right)+f\left(t\right)=0,\tag{E.5}$$

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 (E.5) is of the form on $(0,T)$:

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

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 (E.4) implies $h(T) = h(0) + \frac{1}{2}\left(\frac{1-\gamma}{a_x+\rho}\right)^2 q(T^-)$.

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.

We denote $p(\hat{x}_t, t)$ as the price-to-dividend ratio. 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

$$\pi_t D_t + \lim_{\Delta\to 0}\frac{1}{\Delta}\left\{E_t\left[\pi_{t+\Delta} p(\hat{x}_{t+\Delta}, t+\Delta) D_{t+\Delta}\right] - \pi_t p(\hat{x}_t, t) D_t\right\} = 0. \tag{E.6}$$

Equation (E.6) can be used to show that the price-to-dividend ration function must satisfy the following PDE:

$$1 - p(\hat{x}, t)\varpi(\hat{x}, t) + p_t(\hat{x}, t) - p_x(\hat{x}, t)\nu(\hat{x}, t) + \frac{1}{2} p_{xx}(\hat{x}, t)\frac{q^2(t)}{\sigma^2} = 0, \tag{E.7}$$

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 (E.6) can be used to derive the following boundary condition for $p(\hat{x},t)$:

$$p\left(\hat{x}_T^-,T^-\right)=\frac{E\left[e^{\frac{1-\gamma}{a_x+\rho}\hat{x}_T^+}p\left(\hat{x}_T^+,T^+\right)\Big|\,\hat{x}_T^-,q_T^-\right]}{e^{\frac{1-\gamma}{a_x+\rho}\hat{x}_T^-+\frac{1}{2}\left(\frac{1-\gamma}{a_x+\rho}\right)^2\left[q_T^--q_T^+\right]}}. \tag{E.8}$$

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 (E.7) 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,

$$\mu_{R,t}dt=\frac{1}{p(\hat{x}_t,t)D_t}\left\{D_tdt+E_t\ d\left[p(\hat{x}_t,t)D_t\right]\right\}. \tag{E.9}$$

In the interior of $(nT,(n+1)T)$, the instantaneous risk premium, $\mu_{R,t}-r(\hat{x},t)$ can be computed as

$$\left[\mu_{R,t}-r(\hat{x},t)\right]dt=-Cov_t\left[\frac{d\left[p(\hat{x}_t,t)D_t\right]}{p(\hat{x}_t,t)D_t},\frac{d\pi_t}{\pi_t}\right].$$

We have:

$$\mu_{R,t}-r(\hat{x},t)=\left[\gamma\sigma+\frac{\gamma-1}{a_x+\rho}\frac{q(t)}{\sigma}\right]\left[\phi\sigma+\frac{p_x(\hat{x},t)}{p(\hat{x},t)}\frac{q(t)}{\sigma}\right]. \tag{E.10}$$

To gain a better understanding on how the risk premium and the announcement premium depend on the parameters, let $\varrho(\hat{x},t)=\ln p(\hat{x},t)$, then equation (E.7) can be written as:

$$e^{-\varrho(\hat{x},t)}-\varpi(\hat{x},t)+\varrho_t(\hat{x},t)-\varrho_x(\hat{x},t)\nu(\hat{x},t)+\frac{1}{2}\left[\varrho_{xx}(\hat{x},t)+\varrho_x^2(\hat{x},t)\right]\frac{q^2(t)}{\sigma^2}=0. \tag{E.11}$$

Note that $\hat{x}_t$ is itself an Ornstein-Uhlenbeck process with steady state $\bar{x}$. Using a log-linear approximation around $\hat{x}=\bar{x}$, we can replace the term $e^{-\varrho(\hat{x},t)}$ with $e^{-\varrho(\hat{x},t)}\approx e^{-\bar{\varrho}}-e^{-\bar{\varrho}}\left[\varrho(\hat{x},t)-\bar{\varrho}\right]$, where we denote $\bar{\varrho}\equiv\varrho(\bar{x},t)$, and write

$$e^{-\bar{\varrho}}\left[1+\bar{\varrho}-\varrho(\hat{x},t)\right]-\varpi(\hat{x},t)+\varrho_t(\hat{x},t)-\varrho_x(\hat{x},t)\nu(\hat{x},t)+\frac{1}{2}\left[\varrho_{xx}(\hat{x},t)+\varrho_x^2(\hat{x},t)\right]\frac{q^2(t)}{\sigma^2}=0. \tag{E.12}$$

We conjecture that $\varrho(\hat{x},t)=A\hat{x}+B(t)$, and equation (E.12) can be used to solve for A and $B(t)$ by the method of undetermined coefficients to get $A=\frac{\phi-1}{a_x+e^{-\bar{\varrho}}}$.

Using the log-linearization result to evaluate equation (E.10) at $\hat{x}=\bar{x}$, we obtain (35). In addition, using $p\left(\hat{x}_T^+,T^+\right)\approx e^{A\hat{x}_T^++B\left(T^+\right)}$, we can compute the expectation in (E.8) explicitly and obtain (36).

Numerical Solutions To solve the PDE (E.7) with the boundary condition (E.8), we consider the following auxiliary problem:

$$p\left(x_t,t\right)=E\left[\int_t^T e^{-\int_t^s \varpi(x_u,u)du}ds+e^{-\int_t^T \varpi(x_u,u)du}p\left(x_T,T\right)\right], \tag{E.13}$$

where the state variable x_t follows the law of motion;

$$dx_t=-\nu\left(\hat{x},t\right)dt+\frac{q\left(t\right)}{\sigma}dB_t. \tag{E.14}$$

Note that the solution to (E.13) and (E.8) satisfies the same PDE. Given an initial guess of the pre-announcement price-to-dividend ratio, $p^-\left(x_\tau,\tau\right)$, we can solve (E.13) by the Markov chain approximation method (Kushner and Dupuis [15]):

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

2. We construct a locally consistent Markov chain approximation of of the diffusion process (E.14) as follows. We choose a small dx, let $Q=\left|\nu\left(\hat{x},t\right)\right|dx+\left(\frac{q(t)}{\sigma}\right)^2$, and define the time increment $\Delta=\frac{dx^2}{Q}$ be a function of dx. Define the following Markov chain on the space of x:

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

   One can verify that as $dx\to 0$, the above Markov chain converges to the diffusion process (E.14) (In the language of Kushner and Dupuis [15], this is a Markov chain that is locally consistent with the diffusion process (E.14)).

3. With the initial guess of $p\left(x_T,T\right)$, for $t=T-\Delta$, $T-2\Delta$, etc, we use the Markov chain approximation to compute the discounted problem in (E.13) recursively:

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

   until we obtain $p\left(x,0\right)$.

4. Compute an updated pre-announcement price-to-dividend ratio function, $p\left(x_T,T\right)$ using (E.8):

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

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

Our numerical example is based on the following choice of parameter values:

Choice of parameter values The numerical example we presented in the paper uses parameter values in the standard long-run risk model:

ρ	γ	ψ	$\bar{x}$	a_x	σ_θ	σ	h_0	ϕ	σ_S^2
0.01	10	2	1.8%	0.10	0.026%	3%	5	3	0

All parameters are annual. We assume that announcements are made at the monthly frequency, that is, $T = \frac{1}{12}$.

Pre-announcement drift The density of communication in the top panel of Figure 4 is generated from a Beta distribution with parameter $\alpha = 2$, $\delta = 3$ on $[-6, 0]$ hours before announcement. The density of the Beta distribution is

$$f\left(y \middle| \alpha, \delta\right) = B\left[\sigma, \delta\right]^{-1} y^{\alpha-1}\left(1-y\right)^{\delta-1}, \quad for\ y \in (0,1),$$

where $B\left[\sigma, \delta\right]$ is the Beta function. In our example, the density of the occurrence of a communication h hours before announcement is $f\left(1 - \frac{h}{6} \middle| \alpha, \delta\right)$.

During a small interval dt, the expected return of the dividend claim is $\mu_{R,t}dt$ if the announcement does not occur. The expected return is $\frac{E\left[p\left(\hat{x}_T^+, T^+\right) \middle| \hat{x}_T^-, q_T^-\right]}{p\left(\hat{x}_T^-, T^-\right)}$ if the announcement return occurs during dt. Given that the probability of an announcement during hour $(k-h, k)$ is $\int_{k-h}^{k} f\left(1 - \frac{t}{6} \middle| \alpha, \delta\right) dt$, the expected return of the dividend claim during hour $(k-h, k)$ can be written as

$$E\left[\int_{k-h}^{k} \left\{ f\left(1 - \frac{t}{6} \middle| \alpha, \delta\right) \frac{E\left[p\left(\hat{x}^+_{T+\frac{t}{2880}}, \left(T + \frac{t}{2880}\right)^+\right) \middle| \hat{x}^-_{T+\frac{t}{2880}}, q^-_{T+\frac{t}{2880}}\right]}{p\left(\hat{x}^-_{T+\frac{t}{2880}}, \left(T + \frac{t}{2880}\right)^-\right)} + \mu_{R,T+\frac{t}{2880}} \right\} dt\right]. \tag{E.15}$$

The above calculation assumes that there are 360 days per year and 8 hours per day. Because t is measured in hours, it needs to be divided by $360 \times 8 = 2880$ to translate into annual unit. Numerically, because the pre-announcement drift happens within hours before T, replacing $T + \frac{t}{2880}$ with T does not make any material difference in the evaluation of (E.15). In addition, the term $\int_{k-h}^{k} \mu_{R,T+\frac{t}{2880}} dt$ is negligible. We can therefore approximate the average return during hour $(k-h, k)$ as

$$E\left[\int_{k-h}^{k} f\left(1 - \frac{t}{6} \middle| \alpha, \delta\right) dt\right] \times E\left[\frac{E\left[p\left(\hat{x}_T^+, T^+\right) \middle| \hat{x}_T^-, q_T^-\right]}{p\left(\hat{x}_T^-, T^-\right)}\right].$$

E.2 Time-non-separable Utilities

To guarantee that the model is well defined, we make the following assumptions on the weighting function $\{\xi(t,s)\}_{s=0}^{t}$.

$$\int_0^t \xi(t,s)\,ds \leq 1, \text{ for all } t > 0. \tag{E.16}$$

$$\int_0^\infty \xi(t+s,t)\,ds < \infty, \quad \text{for all } t > 0. \tag{E.17}$$

$$\left(1 - \int_0^t \xi(t,s)\,ds\right) H_0 + \int_0^t \xi(t,s)\,C_s ds < C_t, \quad \text{for all } t > 0. \tag{E.18}$$

The first assumption requires that $\{\xi(t,s)\}_{s=0}^{t}$ is an appropriate weighting function, that is, total weights is less than one. The second assumption implies that the contribution of C_t to future habit stock is finite, and the last assumption ensures $C_t - H_t > 0$ so that the utility function is well defined.

External habit Under the assumption of complete markets, the state-price density can be constructed from the marginal utility of the representative agent. In the external habit model,

$$\pi_t = e^{-\beta t} u'(C_t + bH_t).$$

Internal habit In this case, the calculation of the state price density must take into account of the impact of C_t on future habit stock. Therefore, the state price density is given by (39). Because announcement fully reveals x_t, we need to show that

$$E\left[\int_0^\infty e^{-\beta s}\xi(t+s,t)\,u'(C_{t+s} + bH_{t+s})\,ds \middle| x_t = x\right] \tag{E.19}$$

is a decreasing function of x. Without loss of generality, we assume $t = 0$ in the following lemma.

Lemma A.9. *Fixing the path of Brownian motions $\{B_{C,s}, B_{x,s}\}_{s=0}^{\infty}$,*

$$\frac{\partial}{\partial x_0}[C_t + bH_t] > 0 \quad for \quad all \quad t > 0. \tag{E.20}$$

Proof. Using the law of motion of C_t, we have

$$\ln C_t = \ln C_0 - \frac{1}{2}\sigma^2 t + \int_0^t \sigma dB_{C,s} + \int_0^t x_s ds.$$

Since x_t is an Ornstein-Uhlenbeck process, we can solve $\int_0^t x_s ds$ explicitly:

$$\int_0^t x_s ds = (x_0 - \bar{x})\frac{1}{a_x}\left[1 - e^{-a_x t}\right] + \bar{x}t + \frac{1}{a_x}\int_0^t \left[1 - e^{a_x(s-t)}\right]\sigma_x dB_{x,s}.$$

Therefore, for given realizations of the Brownian motion paths,

$$\frac{\partial}{\partial x_0} C_t = C_t \frac{1}{a_x} \left[1 - e^{-a_x t}\right],$$

and

$$\begin{aligned}
\frac{\partial}{\partial x_0} H_t &= \int_0^t \xi(t,s) \frac{\partial C_s}{\partial x_0} ds \\
&= \int_0^t \xi(t,s) C_s \frac{1}{a_x} \left[1 - e^{-a_x t}\right] ds \\
&< \int_0^t \xi(t,s) C_s ds \frac{1}{a_x} \left[1 - e^{-a_x t}\right] \\
&< C_t \frac{1}{a_x} \left[1 - e^{-a_x t}\right],
\end{aligned}$$

where the first inequality is true because $s < t$, and the second is due to the fact that $\int_0^t \xi(t,s) C_s ds \leq H_t < C_t$. The inequality (E.20) follows because $b < 1$. □

Consider two initial conditions, $x_0 = x$ and $x_0 = x'$. The above lemma implies that $x > x'$ implies that $C_{t+s} + bH_{t+s}$ first order stochastic dominate $C'_{t+s} + bH'_{t+s}$. Because $u'(\cdot)$ is a strictly decreasing function, we conclude that (E.19) must be a decreasing function of x.

Consumption substitutability Because (E.19) is decreasing function of x, with $b > 0$, the state price density in (39) must be a decreasing function of x_t as well. As a result, the announcement premium must be positive for any payoff that is increasing in x_t.

References

[1] Carlier, Guillaume, and Rose-Anne Dana, 2003, Core of Convex Distortions of a Probability, *Journal of Economic Theory* 113, 199–222.

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

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

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

[5] 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.

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

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

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

[9] Hayashi, Takashi, 2005, Intertemporal substitution, risk aversion and ambiguity aversion, *Economic Theory* 25, 933–956.

[10] Hayashi, Takashi, and Jianjun Miao, 2011, Intertemporal Substitution and Recursive Smooth Ambiguity Preferences, *Theoretical Economics* 121, 496–520.

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

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

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

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

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

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

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

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

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

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

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

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

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

[24] Tapia, Richard A., 1971, The Differentiation and Integration of Nonlinear Operators, in L.B.Rall, eds.: *Nonlinear Functional Analysis and Applications* (Academic Press, New York,).

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